

Direction des bibliothèques

AVIS

Ce document a été numérisé par la Division de la gestion des documents et des archives de l'Université de Montréal.

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

This document was digitized by the Records Management & Archives Division of Université de Montréal.

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal

**Étude de la fonction du promoteur fœtal $A\gamma$ dans la régulation
de la commutation de l'hémoglobine fœtale à adulte**

par

Hugues Beauchemin

Programme de biologie moléculaire

Faculté de Médecine

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

Philosophiæ Doctor (Ph.D.)
en Biologie Moléculaire

Mars, 2008

© Hugues Beauchemin, 2008



Université de Montréal
Faculté de Médecine

Cette thèse intitulée :

**Étude de la fonction du promoteur foetal $A\gamma$ dans la régulation de la commutation de
l'hémoglobine foetale à adulte**

présentée par :

Hugues Beauchemin

a été évaluée par un jury composé des personnes suivantes :

Dr Éric Milot, président-rapporteur

Dre Marie Trudel, directrice de recherche

Dr Tarik Möröy, membre du jury

Dr Michel Sadelain, examinateur externe

Dr Harry Bard, représentant du doyen de la faculté de médecine

Résumé

La commutation de l'hémoglobine se caractérise par l'expression séquentielle des cinq gènes du locus humain de la β -globine (ϵ , γ^G , γ^A , δ et β) lors du développement: ϵ -globine est exprimé au stade embryonnaire, γ^G -globine et γ^A -globine au stade fœtal, β -globine et δ -globine après la naissance. Ces gènes sont sous le contrôle de leur promoteur respectif et de la région de contrôle du locus. Toutefois, les mécanismes qui régulent cette commutation ne sont pas connus. L'objectif de cette thèse était de déterminer le rôle du promoteur de γ^A dans la régulation de la commutation de l'hémoglobine dans le contexte du locus entier de la β -globine, contenu dans un YAC.

Pour évaluer le rôle du promoteur γ^A dans la commutation de l'hémoglobine, la région -838 à +49 du gène γ^A du YAC a été remplacée par la région équivalente (-815 à +72) du promoteur érythroïde-spécifique et stade-indépendant de la porphobilinogène désaminase (PBGD). Tel qu'attendu du promoteur PBGD, les souris transgéniques (PBGD γ^A -YAC) ont démontré une expression stable de γ^A tout au long du développement. Également, au stade fœtal l'expression de γ^G était augmentée significativement en comparaison au locus sauvage, surpassant l'expression des gènes PBGD γ^A et β . Même après la naissance, l'expression de γ^G demeurerait toujours détectable alors que dans le locus sauvage, les gènes γ n'étaient presque plus exprimés dès e14,5. À l'inverse, la cinétique d'activation de β , était sévèrement diminuée, atteignant son expression maximal après la naissance alors que le maximum pour le locus sauvage était atteint dès e14,5, résultant en un délai de la commutation de l'hémoglobine fœtale à adulte. Des analyses de méthylation de l'ADN et de ChIP n'ont démontré qu'une variation marginale au niveau de la méthylation de l'ADN, rejetant l'hypothèse d'un phénomène épigénétique. À l'inverse, une capture de conformation chromosomique a démontré que le retard dans la commutation était dû au recrutement du LCR au domaine fœtal par le promoteur PBGD et à son exclusion consécutive du domaine adulte. Ces résultats révélaient que le LCR interagissait

simultanément avec les deux gènes d'un domaine et activait leur expression de façon concomitante, probablement via différentes régions. Je propose ainsi le modèle d'une commutation de l'hémoglobine impliquant la coopération bigénique de gènes fonctionnant en paires pour recruter le LCR via un mécanisme de bicompetition où chaque gène d'un domaine est en compétition avec un gène spécifique de l'autre domaine pour leurs sites respectifs.

Une analyse fine du promoteur γ n'a pu permettre toutefois d'identifier les régions impliquées dans cette fonction. En effet, les délétions des régions -842 à -384 et -384 à -202 du promoteur de γ n'ont démontré aucun impact sur l'expression des gènes du locus. Cependant, une analyse détaillée de la délétion -842/-384 qui était sensée éliminer la fonction répressive d'un site GATA-1 à -570, a permis de découvrir que cette délétion avait repositionné à proximité un site GATA-1 distal, reconstituant cette fonction et indiquant que la nature du site n'était pas aussi décisive que sa position relative au gène.

Mots-clés: gamma-globine, promoteur, région de contrôle du locus, coopération bigénique, compétition génique, génétique, épigénétique, régulation génique, souris transgénique.

Abstract

Hemoglobin switching is characterized by a sequential expression of the five genes of the human β -globin locus (ϵ , γ^G , γ^A , δ and β) during the development: ϵ -globin is expressed during the embryonic stage, γ^G -globin and γ^A -globin during the fetal stage and β -globin and δ -globin after birth. These genes are under the control of their respective promoter as well as the locus control region, but the mechanisms underlying the switching regulation are still unknown. The objective of this thesis was to determine the role of the γ^A promoter in the hemoglobin switching in the context of a YAC containing the entire locus.

To investigate the role of γ^A promoter on hemoglobin switching, the region -838 to +49 of γ^A was replaced by the equivalent region (-815 to +72) of the erythroid-specific stage-independent promoter of the porphobilinogen deaminase (PBGD) in the YAC. The transgenic lines (PBGD γ^A -YAC) displayed a stable γ^A expression throughout the development as predicted by the PBGD promoter's specificity. Also, during the fetal stage, γ^G expression was markedly increased compared to wild type locus, exceeding both PBGD γ^A and β -globin expression. Even after birth, γ^G expression was readily detectable whereas in the wild-type locus, the γ genes were no longer expressed from e14.5. Conversely, the kinetic of β activation was severely impaired, reaching its maximal expression after birth whereas the maximum in the wild-type locus was attained as soon as e14.5, resulting in delayed fetal to adult hemoglobin switch. DNA methylation and ChIP analyses showed only minor variations in DNA methylation, rejecting the hypothesis of an epigenetic phenomenon. In contrast, a chromosome conformation capture assay showed that the delay in the hemoglobin switch was due to the LCR being recruited by the PBGD promoter to the fetal domain at the expense of the adult domain. These results revealed that the LCR interacted and promoted concurrent activation of the two genes in a given domain, possibly via different regions. Thus, I propose a hemoglobin switching model that involves pairs of genes that cooperate to recruit the LCR through a bicompetitive mechanism in

which each gene of a domain competes with a specific gene of the other domain for their respective sites.

However, a refined analysis of the γ promoter was unable to identify the regions involved in this function. The deletion of the regions -842 to -384 and -384 to -202 of the γ promoter failed to show any impact on globin genes expression, but a detailed analysis of the deletion -842/-384 that was expected to abolish the silencer activity of a GATA-1 binding site at -570, showed that the deletion moved a distal GATA-1 site closer, reconstituting this activity, indicating that the nature of the GATA-1 site was not as critical as its position relative to the gene.

Keywords: gamma-globin, promoter, locus control region, bigenic cooperation, gene competition, genetics, epigenetics, gene regulation, transgenic mice.

Table des matières

Résumé	iii
Abstract.....	v
Table des matières.....	vii
Liste des tableaux	xiii
Liste des figures	xiv
Liste des équations	xvii
Liste des abréviations et des symboles	xviii
Remerciements	xxiv
Avant-propos	1
Chapitre 1 Le locus de la β-globine	2
1.1 L'hémoglobine.....	2
1.2 La β-globine: l'histoire de... cinq gènes	7
1.2.1 Le pigment respiratoire	7
1.2.2 La respiration fœtale	9
1.2.3 L'ère de la génétique	11
1.3 Le locus humain de la β-globine	14
1.3.1 La commutation de l'hémoglobine.....	14
1.3.2 β - et δ -globine.....	17
1.3.3 $A\gamma$ - et $G\gamma$ -globine.....	17
1.3.4 ϵ -globine	18
1.3.5 La région de contrôle du locus (LCR)	19
1.3.6 L'environnement chromosomique du locus	20
1.3.7 Hémoglobinopathies de la β -globine	21
1.4 Le locus de la β-globine chez les autres espèces	26
1.4.1 La souris.....	29

1.4.2	Les prosimiens	31
1.4.3	Les simiens I: les singes du Nouveau Monde (<i>platyrhiniens</i>).....	33
1.4.4	Les simiens II: les singes de l'Ancien Monde (<i>catarhiniens</i>).....	35
1.4.5	Les ruminants.....	36
1.5	Régulation.....	41
1.5.1	La commutation embryonnaire à fœtale	42
1.5.2	La commutation fœtale à adulte.....	42
1.5.3	Éléments <i>cis</i> -régulateurs	43
1.5.4	Facteurs <i>trans</i> -régulateurs.....	55
1.5.5	Épigénétique	58
1.5.6	Modèles de régulation.....	63
1.6	Systèmes modèles	68
1.6.1	Études en culture cellulaire.....	68
1.6.2	Les locus non humains.....	69
1.6.3	Le locus humain en souris transgéniques	71
1.7	Hypothèse et objectifs de la thèse	80
Chapitre 2 Matériel et méthodes		85
2.1	Clonage moléculaire	85
2.1.1	Préparation des fragments d'ADN à partir de plasmide.....	85
2.1.2	Préparation du fragment PBGD à partir de l'ADN génomique	87
2.1.3	Ligation.....	88
2.1.4	Transformation.....	88
2.1.5	Criblage de clones positifs.....	88
2.1.6	Préparation de plasmide.....	90
2.1.7	Caractérisation des clones.....	91
2.1.8	Mutagenèse dirigée	92
2.2	Génération des plasmides d'intégration dans la levure.....	92
2.2.1	Le YIp pRS-PBGD ^{Aγ}	93

2.2.2	Les YIp de délétions	95
2.3	Recombinaison homologue dans la levure.....	99
2.3.1	Culture et manipulation des levures.....	100
2.3.2	Transformation de levures: Insertion du YIp (<i>Pop-in</i>).....	102
2.3.3	Excision du YIp (<i>Pop-out</i>).....	103
2.3.4	Caractérisation des YAC mutants.....	103
2.4	Génération des YAC mutants.....	104
2.4.1	PBGD ^A γ -YAC	105
2.4.2	Δ -842/-384 ^A γ -YAC.....	106
2.4.3	Δ -384/-202 ^A γ -YAC.....	106
2.4.4	Δ -204/-135 ^A γ -YAC.....	107
2.5	Souris transgéniques.....	107
2.5.1	Préparation du YAC pour la microinjection.....	107
2.5.2	Microinjection du YAC	108
2.5.3	Caractérisation des fondatrices	108
2.5.4	Établissement des lignées transgéniques	110
2.5.5	Acquisition de fœtus transgéniques	111
2.5.6	Prélèvement d'organes	112
2.5.7	Les souris PBGD ^A γ -YAC	113
2.5.8	Les souris Δ -842/-384 ^A γ -YAC	113
2.5.9	Les souris Δ -384/-202 ^A γ -YAC	114
2.5.10	Les souris Δ -204/-135 ^A γ -YAC	114
2.5.11	La souris contrôle β -YAC.....	115
2.5.12	La souris EKLF-nulle	115
2.5.13	La souris Ikaros-nulle	115
2.6	Analyses d'expression	116
2.6.1	Quantification des ARNm de globine.....	116
2.6.2	Quantification des chaînes de globine	117

2.6.3	Distribution cellulaire de l'expression	118
2.7	Analyse épigénétique	119
2.7.1	Immunoprécipitation de chromatine (ChIP)	119
2.7.2	Analyse de méthylation de l'ADN	122
2.8	Analyse mécanistique	123
2.8.1	Capture de conformation chromosomique (3C)	123
2.9	Méthodes statistiques.....	128
Chapitre 3 Résultats.....		129
3.1	Génération des YAC mutants.....	129
3.1.1	PBGD ^A γ -YAC	129
3.1.2	Δ -842/-384 ^A γ -YAC	132
3.1.3	Δ -384/-202 ^A γ -YAC	133
3.1.4	Δ -204/-135 ^A γ -YAC	134
3.2	Génération des souris transgéniques.....	136
3.2.1	Les souris PBGD ^A γ -YAC	136
3.2.2	Les souris Δ -842/-384 ^A γ -YAC	137
3.2.3	Les souris Δ -384/-202 ^A γ -YAC	138
3.3	Analyses d'expression dans les souris transgéniques	139
3.3.1	Les souris PBGD ^A γ -YAC	139
3.3.2	Les souris PBGD ^A γ -YAC/EKLF-nulles	149
3.3.3	Les souris PBGD ^A γ -YAC/Ikaros-nulles	150
3.3.4	Les souris Δ -842/-384 ^A γ -YAC	152
3.3.5	Les souris Δ -384/-202 ^A γ -YAC	154
3.4	Études de l'épigénétique chez la souris PBGD^Aγ-YAC	155
3.4.1	Modification des histones dans le locus	155
3.4.2	Méthylation 5mC de l'ADN dans le locus	158
3.5	Études mécanistiques chez la souris PBGD^Aγ-YAC	160

3.5.1	Capture de conformation chromosomique.....	160
Chapitre 4 Discussion		164
4.1	Impact du promoteur PBGD sur la commutation.....	166
4.1.1	Délai de la commutation de l'hémoglobine fœtale à adulte	167
4.1.2	Un modèle.....	185
4.2	Impact des délétions sur la commutation	187
4.2.1	La délétion Δ -842/-384	187
4.2.2	La délétion Δ -384/-202	190
4.3	Perspectives	191
4.3.1	Définition des mécanismes de la coopération bigénique.....	191
4.3.2	Caractérisation de la bicompetition	191
4.3.3	Évaluation du rôle de la région -204/-135	192
Conclusion.....		193
Bibliographie		195
Annexe I Méthodes supplémentaires		I
Listes d'amorces.....		II
Annexe II: Résultats supplémentaires		VII
Caractérisation des YIpΔ.....		VIII
Séquençage de pB Δ -842/-384.....		VIII
Séquençage de pB Δ -384/-202.....		IX
Séquençage de pB Δ -204/-135.....		X
Mutagenèse dirigée $\Delta\gamma$: -117(A \rightarrow G).....		XI
Caractérisation du PBGD$\Delta\gamma$-YAC par PFGE.....		XII
Caractérisation des fondatrice PBGD$\Delta\gamma$-YAC.....		XIII
Quantification postnatale des chaînes de globine		XIV
Les souris PBGD $\Delta\gamma$ -YAC/Ikaros-nulles		XIV
Annexe III Résultats ajoutés en preuve		XV

Expression chez les souris PBGD^A-YAC/βBP1	XVI
Annexe IV Cartes des plasmides	XVIII
pRS406	XIX
pBluescript.....	XX

Liste des tableaux

Chapitre 1

Tableau I: Mutations HPFH non issues de délétions.....	25
Tableau II: Transgéniques avec le locus humain reproduisant des mutations humaines	74
Tableau III: Transgéniques avec le locus humain modifié au niveau des gènes	75
Tableau IV: Transgéniques avec le locus humain modifié au niveau du LCR.....	76
Tableau V: Transgéniques avec le locus humain modifié au niveau de son organisation....	77
Tableau VI: Transgéniques avec le locus humain combinant des modifications	78

Chapitre 2

Tableau VII: Milieux de culture de levure.....	101
--	-----

Chapitre 3

Tableau VIII: Caractérisation des fondatrices Δ -841/-384 ^A γ -YAC.....	138
Tableau IX: Caractérisation des fondatrices Δ -384/-202 ^A γ -YAC	139

Annexe I

Tableau X: Clonage du promoteur érythroïde de PBGD (PCR niché).....	II
Tableau XI: Mutagenèse dirigée du promoteur de ^A γ (-117: A→G) (PCR).....	II
Tableau XII: Ammorces pour les Séquençages	II
Tableau XIII: Criblage et Génotypage (PCR)	III
Tableau XIV: expression des gènes de globine (RT-PCR en temps réel).....	IV
Tableau XV: ChIP (PCR semi-quantitatif).....	V
Tableau XVI: ChIP (PCR en temps réel).....	V
Tableau XVII: Capture de conformation chromosomique (amplification par PCR).....	VI

Annexe II

Tableau XVIII: Caractérisation des fondatrices PBGD ^A γ -YAC	XIII
--	------

Liste des figures

Chapitre 1

Figure 1: Structure de l'hémoglobine.....	3
Figure 2: Structure de l'hémoglobine tel que présentée par Perutz	8
Figure 3: Le locus humain de la β -globine	14
Figure 4: Commutation de l'hémoglobine chez l'humain	15
Figure 5: Région chromosomique entourant le locus de la β -globine	20
Figure 6: Délétions dans le locus conduisant à une HPFH.....	24
Figure 7: Phylogénie, génétique et physiologie du locus de la β -globine	28
Figure 8: Le locus de la β -globine de la souris <i>Mus musculus</i>	29
Figure 9: Commutation de l'hémoglobine chez la souris.....	30
Figure 10: Le locus de la β -globine chez les prosimiens.....	31
Figure 11: Commutation de l'hémoglobine chez les prosimiens	32
Figure 12: Le locus de la β -globine chez les singes du Nouveau Monde	33
Figure 13: Commutation de l'hémoglobine chez les singes du Nouveau Monde.....	34
Figure 14: Le locus de la β -globine chez les singes de l'Ancien Monde.....	35
Figure 15: Commutation de l'hémoglobine chez les singes de l'Ancien Monde	36
Figure 16: Le locus de la β -globine chez ruminants.....	37
Figure 17: Commutation de l'hémoglobine chez les ruminants.....	39
Figure 18: Promoteurs des gènes ϵ -, γ - et β -globine.....	44
Figure 19: Modèle du <i>Looping</i>	65
Figure 20: Modèles du <i>Tracking</i> et du <i>Tracking</i> facilité	67
Figure 21: Les promoteurs γ -globine, PBGD et l'hybride PBGD γ -globine.....	82
Figure 22: Régions excisées du promoteur de la γ -globine	84

Chapitre 2

Figure 23: Plasmide d'intégration dans la levure pRS-PBGD γ	94
--	----

Figure 24: Les YIp de délétion	97
Figure 25: Recombinaison homologue dans la levure.....	99
Figure 26: Capture de Conformation Chromosomique	123

Chapitre 3

Figure 27: Caractérisation de l'intégrité du mutant PBGD ^Δ γ-YAC.....	131
Figure 28: Caractérisation de l'intégrité des mutants Δ ^Δ γ-YAC	135
Figure 29: Caractérisation des fondatrices PBGD ^Δ γ-YAC	137
Figure 30: Expression des gènes du locus PBGD ^Δ γ-YAC (qPCR).....	143
Figure 31: Analyse des chaînes de globine dans les souris PBGD ^Δ γ-YAC	146
Figure 32: Distribution cellulaire de l'expression de γ et β dans les érythrocytes matures adultes	148
Figure 33: Expression des gènes de globine chez les fœtus PBGD ^Δ γ-YAC/EKLF-nuls ...	150
Figure 34: Expression des chaînes de globine chez les souris PBGD ^Δ γ-YAC/Ikaros-nulles	152
Figure 35: Expression des gènes du locus Δ-842/-384 ^Δ γ-YAC.....	153
Figure 36: Expression des gènes du locus Δ-384/-202 ^Δ γ-YAC.....	154
Figure 37: Modifications des histones dans les souris PBGD ^Δ γ-YAC.....	157
Figure 38: Patron de méthylation de l'ADN dans les souris PBGD ^Δ γ-YAC.....	159
Figure 39: Proximité relative du LCR et des différentes portions du locus à e12,5	162
Figure 40: Proximité relative du LCR et des différentes portions du locus chez l'adulte ..	163

Chapitre 4

Figure 41: Modèle de la régulation des gènes de globine dans le locus PBGD ^Δ γ-YAC....	186
Figure 42: modèle expliquant la régulation correcte du gène ^Δ γ chez la souris Δ-842/-384 ^Δ γ- YAC.....	190

Annexe II

Figure 43: Comparaison de séquence entre la construction pBΔ-842/-384 et la séquence sauvage du promoteur de γ -globine	VIII
Figure 44: Comparaison de séquence entre la construction pBΔ-384/-202 et la séquence sauvage du promoteur de γ -globine	IX
Figure 45: Comparaison de séquence entre la construction pBΔ-204/-135 et la séquence sauvage du promoteur de γ -globine	X
Figure 46: Séquençage des constructions mutées à la position -117(A→G)	XI
Figure 47: Caractérisation du mutant PBGD γ -YAC par <i>PFGE</i>	XII
Figure 48: Expression des chaînes de globine chez les souris PBGD γ -YAC/Ikaros-nulles	XIV

Annexe III

Figure 49: Expression des gènes de globine chez les fœtus PBGD γ -YAC/ β BP1-43....	XVII
--	------

Annexe IV

Figure 50: Plasmide d'intégration dans la levure pRS406	XIX
Figure 51: Plasmide pBluescript II	XX

Liste des équations

Chapitre 2

Équation 1:	$\% \text{ méthylation} = \left(\frac{\text{bande(NC)}}{\text{bande(NC)} + \text{bande(C)}} \right) \times 100\%$ 123
Équation 2:	$p(r) = \left(\frac{rA(3C)/rA(\text{contrôle})}{\psi\beta(3C)/\psi\beta(\text{contrôle})} \right)$ 127

Liste des abréviations et des symboles

α	<i>alpha</i> , α -globine
α -	Anti-
β	<i>bêta</i> , β -globine
δ	<i>delta</i> , δ -globine
η	<i>êta</i> , η -globine
γ	<i>gamma</i> , γ -globine
μ	Micro
ω	<i>oméga</i> , ω -globine
λ	Phage <i>lambda</i> , longueur d'onde
φ	<i>phi</i>
ψ	Pseudo
ρ	<i>rhô</i> , ρ -globine
θ	<i>thêta</i> , θ -globine
ζ	<i>zêta</i> , ζ -globine
Δ	Délétion
β DR Γ	<i>β-globin direct repeat factor</i>
γ PE	<i>γ-globin gene promoter and enhancer binding factor</i>
ϵ -SSP	<i>ϵ-globin stage selector protein</i>
2,3-DPG	2,3-diphosphoglycérate
3C	<i>Chromosome conformation capture</i>
4C	<i>Chromosome conformation capture-on-chip</i>
5C	<i>Chromosome conformation capture carbon copy</i>
5-FOA	<i>5-fluoroorotic acid</i>
5mC	5-méthylcytosine
A	Adénine, adulte, ampère
ADA	Adénosine désaminase
ADN	Acide désoxyribonucléique
ADN _{ss}	ADN de sperme de saumon
AHSP	<i>α hemoglobin-stabilizing protein</i>
Ala	Alanine
ANOVA	<i>Analysis of variance</i>
AP-1	<i>Adapter-related protein complex 1</i>
AP-2	<i>Adapter-related protein complex 2</i>
ARN	Acide ribonucléique
ARN _m	ARN messenger
ATP	Adénosine triphosphate
B.D.	Bras droit
B.G.	Bras gauche
BAC	<i>Bacterial artificial chromosome</i>
bHLH	<i>Basic-helix-loop-helix</i>
bp	Paire de bases
BP1	<i>β-protein 1</i>

BP2	<i>β-protein 2</i>
BSA	<i>Bovine Serum Albumine</i>
C	Cytosine, Coupé
C/EBPγ	<i>CCAAT/enhancer binding protein γ</i>
C/EBPδ	<i>CCAAT/enhancer binding protein δ</i>
CAC	Boîte CACCC
CBF1	<i>CRT/DRE-binding factor 1</i>
ChIP	<i>Chromatin immunoprecipitation</i>
Ci	Currie
COUP-TFII	<i>Chicken ovalbumin upstream promoter transcription factor II</i>
CPI	<i>CAAT-box binding protein 1</i>
CP2	<i>CAAT-box binding protein 2</i>
CSBP-2	<i>Cytokinin-specific binding protein 1</i>
CTCF	<i>CCCTC-binding factor</i>
d	Diffus
dATGTP	Désoxyadenosine - guanosine - thymidine triphosphate
DCR	<i>Dominant control region</i>
DEPC	Diéthyle pyrocarbonate
DLX7	<i>Distal-less homeobox gene 7</i>
DMEM	<i>Dulbecco's modified Eagle's medium</i>
DNaseI	Désoxyribonucléase I
dNTP	Désoxynucléotide triphosphate
DRED	<i>Direct repeat erythroid-definitive</i>
DTT	Dithio-thréitol
E, e	Embryonnaire, jour embryonnaire
EDTA	<i>Ethylenediamine tetraacetic acid</i>
EKLF	<i>Erythroid Krüppel-like factor</i>
ERV-9	<i>Human endogenous retrovirus - 9</i>
Est-1	<i>Ever shorter telomeres protein 1</i>
et al	<i>et alia (et autres)</i>
F, f	Fœtal, femto
Facs	<i>Fluorescent-activated cell sorting</i>
Fisher	Test exact de Fisher
FKLF	<i>Fetal Krüppel-like factor</i>
FOG-1	<i>Friend of GATA-1</i>
G, g	Guanine, gramme
GATA-1	<i>GATA binding protein 1</i>
Glu	Glutamate
Gly	Glycine
GTP	Guanosine triphosphate
HAC	<i>Human artificial chromosome</i>
Hb	Hémoglobine
HbA	Hémoglobine A
HbA2	Hémoglobine A2
HbF	Hémoglobine F
HbS	Hémoglobine S

HDAC1	<i>Histone deacetylase 1</i>
HEL	<i>Human erythroleukemia line</i>
hES	<i>Human embryonic stem cells</i>
His	Histidine
HMG	<i>High mobility group chromatin proteins</i>
HOXB2	<i>Homeobox B2</i>
HPFH	<i>Hereditary persistence of fetal hemoglobin</i>
HS	<i>DNaseI Hypersensitive site/Sites d'hypersensibilité à la DNaseI</i>
3'HS	HS situé en 3' du locus
5'HS1-5	Les cinq HS situés en 5' du locus
HUVSM	<i>Human umbilical vein smooth muscle</i>
INI1	<i>Integrase interactor 1</i>
IPP	<i>Inositol pentaphosphate</i>
J	Joule, juvénile
K4/K9	Lysine 4/Lysine 9
kb	Kilo paire de bases
KLF	<i>Krüppel-like factor</i>
KO	<i>Knock-out</i>
L	Litre
LAR	<i>Locus activation region</i>
LB	<i>L-broth</i>
LCR	<i>Locus control region</i>
LMP	<i>Low melting point</i>
LTR	<i>Long terminal repeate</i>
Lys	Lysine
M, m	Mole par litre (molaire), milli ou marqué
Maj	Majeur
MAR	<i>Matrix attachment region</i>
MBD	<i>Methyl-CpG binding domain</i>
mCpG	CpG méthylé sur la cytosine
MEL	<i>Murine erythroleukemia line</i>
Min	Mineur
M-MLV	<i>Moloney murine leukemia virus</i>
n	Nano
NB	<i>Newborn/Nouveau-né</i>
NC	Non coupé
NF-κB	<i>Nuclear factor-kappa B</i>
NF-1	<i>Nuclear factor 1</i>
NF-E2	<i>Nuclear factor erythroid-derived 2</i>
NF-E4	<i>Nuclear factor erythroid-derived 4</i>
NFY/NF-Y	<i>Nuclear factor binding the Y box</i>
nt	Nucléotide
Oct-1	<i>Octamer factor 1</i>
p	Pico
P	Prénatal
CBP	<i>CREB-binding protein</i>
PAC	<i>P1-derived artificial chromosome</i>

pb	Paire de bases
PBGD	Porphobilinogène désaminase
PBS	<i>Phosphate Buffered Saline</i>
PCR	<i>Polymerase chain reaction</i>
qPCR	<i>Quantitative PCR</i>
RT-PCR	<i>Reverse transcriptase - PCR</i>
PEV	<i>Position-effect variegation</i>
PFGE	<i>Pulse field gel electrophoresis</i>
pH	Potentiel hydrogène
<i>pi</i>	<i>Pop-in</i>
Pkd1	<i>Polycystic kidney disease 1</i>
PMS	<i>Pregnant mare serum</i>
PMSF	<i>Phenylmethylsulfonyl fluoride</i>
PNK	Polynucléotide kinase
<i>po</i>	<i>Pop-out</i>
PY	PBGD^{Δγ}-YAC
PyE	<i>Polyoma virus enhancer</i>
PYR	<i>Pyrimidine-rich region</i>
FISH	<i>Fluorescence in situ hybridization</i>
RNase	Ribonucléase
ROH	Récepteur olfactif humain
<i>s, s</i>	<i>Seconde, single</i>
SATB1	<i>Special AT-rich sequence binding protein 1</i>
SCL	<i>Stem cell leukemia</i>
SDS	<i>Sodium dodecyl sulfate</i>
Ser	Sérine
Sev	Sevrage
Sp1	<i>Specificity protein 1</i>
SSE	<i>Stage selector element</i>
SSP	<i>Stage selector protein</i>
STAT3β	<i>Signal transducer and activator of transcription 3β</i>
SFV	Sérum fœtal de veau
T	Thymine
TAE	Tris-acetate-EDTA
Taq	<i>Thermophilus aquaticus</i>
TBE	Tris-borate-EDTA
TE	Tris-EDTA
TEF-2	<i>Translational elongation factor 2</i>
TEMED	<i>N,N,N',N'-Tetramethylethylenediamine</i>
Thal	Thalassémie
TKCAT	<i>Thymidine kinase chloramphenicol acetyltransferase</i>
Trp	Tryptophane
U	Unité
Ura	Uracile
USF	<i>Upstream stimulatory factors</i>
UT-PAGE	<i>Urea/Triton-polyacrylamide gel electrophoresis</i>
UTR	<i>Untranslated region</i>

UV	Ultraviolet
UW	Unité Weiss
V	Volt
Val	Valine
wt	<i>Wild-type/Type sauvage</i>
X-gal	<i>5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside</i>
×g	Force de relative de centrifugation
YAC	<i>Yeast artificial chromosome</i>
Y	β-YAC
YIp	<i>Yeast integrating plasmid</i>
YY1	Yin Yang 1
Zfp37	<i>Zinc finger protein 37</i>

À ma famille

*"You know, hemoglobin is a wonderful substance.
I like it. It's a red substance that brings color into the cheeks of
girls, and in the course of my hemoglobin investigation I look
about a good bit to appreciate it."*

Linus Pauling

*"We think of truth as something that is invariable, but add a new
circumstance and we have a new truth."*

William J. Mayo

Remerciements

Je remercie Marie Trudel, ma directrice de recherche, pour m'avoir accueilli dans son laboratoire, de m'avoir donné un projet aussi fascinant et de m'avoir enseigné tous les rudiments de la génétique et de la biologie moléculaire. Je la remercie également pour son esprit critique et pour les discussions scientifiques qui ont permis à ce projet de progresser. Sans elle, je ne serais pas devenu celui que je suis.

Je voudrais remercier également les membres du jury, pour avoir bien accepté de lire, corriger et donner leurs impressions sur cette thèse.

Évidemment, je me dois de remercier les stagiaires et étudiants qui, au fil des ans, ont mis la main à la pâte et ont ainsi contribué à l'avancement de ce projet. Dans l'ordre de leur passage dans le labo: Julien Desroches, Line Germain, Mireille Lépine, Yvonne Chow, Klodiana Jani, Fahimé Tavanayanfar et Marthe-Sandrine Eiyimo Mwa Mpollo. Merci également à Wafaa Lemsaddek, Marc Andrews et au pouilleux pour avoir accepté de relire et corriger cette thèse, ce fut très apprécié. J'ai également une pensée particulière pour Mylène Cawthorn, qui a pris soins de mes souris pendant toutes ces années.

Merci également à tous ceux qui furent mes amis durant ces longues années, Richard Guillaume pour son imagination fertile, Martin Couillard avec qui on a su passer à travers toutes les tempêtes, Olivier Côté pour les boules de papiers et les araignées, Mohammed el Mabrouk pour son amitié, Fabienne Vozy pour tous les E-mail inutiles et Marie-José Blouin pour avoir été si patiente. J'ai une pensée aussi pour ma famille reconstituée Monica Pata, Janie Beauregard, Céline Héraud, Mathieu Ferron, Nader Chaloub, Mouhamad Rashed et Michel Arsenault. Bien sûr, merci également à tous les autres qui n'ont pas été nommés ici.

Finalement je me dois d'adresser toute ma gratitude à famille. Ma mère Jacqueline Sicotte pour son soutien pendant toutes ces années et ma sœur Isabelle, son conjoint Philippe Chantigny ainsi que leurs deux bouts d'choux Guillaume et Justine, parce qu'ils ont

toujours été là. Finalement et par-dessus tout, je veux remercier ma conjointe Shufeng Zhou, qui me supporte nuit et jour et sans qui, ces deux dernières années m'auraient parut bien plus longues... 爱你到永远

Avant-propos

Comment devient-on biologiste moléculaire? Bien que la question puisse paraître quelque peu banale, en fait elle n'en est rien. Je me souviens il n'y a pas si longtemps, peut-être treize ou quatorze ans de cela, bien qu'intrigué par cette science appelée génétique, pour moi, quelqu'un avait autant de chance de devenir généticien que de devenir astronaute; la génétique était une science réservée à une certaine élite, de celle que l'on ne rencontre jamais, qui sait et qui est capable de choses étranges comme de modifier le code génétique des animaux. À cette époque, je travaillais dans un entrepôt, emplissant des boîtes de boulons et d'écrous, rêvant d'un futur meilleur. Lorsque j'ai décidé de retourner aux études, je ne savais pas alors vers quel domaine me diriger. Mais au tréfonds de ma conscience, à la frontière du subconscient, il y avait cette idée, un peu folle, la génétique. Cependant, à cette époque, je n'avais aucune idée comment on devenait généticien mais je me suis tout de même inscrit au cégep en science de la nature. Ce fut également cette pensée diaphane qui m'a fait choisir la biochimie lorsque je me suis inscrit à l'université, bien que je ne comprenais pas très bien à ce moment ce qu'était cette discipline, ni même si elle avait quelques chances de mener à la science des gènes.

Durant mon baccalauréat, nécessairement j'ai été mis en contact avec la génétique: la structure des gènes, leurs fonctions, comment ils étaient régulés. Tout ça je l'ai appris, mais seulement en superficie et trop de choses demeuraient obscures. C'est vers la fin du baccalauréat, lorsque je devais me trouver un laboratoire d'accueil pour un stage d'été que j'ai approché le laboratoire de génétique moléculaire et développement du Dr Marie Trudel: là, moi aussi j'allais savoir! C'est ainsi que j'ai débuté mes recherches en biologie moléculaire, au cours desquelles j'ai généré des souris transgéniques et j'ai étudié la régulation des gènes de globine pour tenter d'en extirper leurs secrets. Près de 9 ans plus tard, maintenant que je termine mes études de doctorat, je me rends compte que je ne sais toujours pas! Enfin j'ai bien appris quelques trucs, mais lorsque je regarde devant et que je vois ce qui me reste à savoir...

... le futur m'apparaît bien intéressant.

Chapitre 1

Le locus de la β -globine

1.1 L'hémoglobine

Chez les vertébrés, le transport de l'oxygène dans le sang vers les tissus est effectué par une hémoprotéine hétérotétramérique, l'hémoglobine, qui est exprimée exclusivement dans les érythrocytes et est constituée de deux chaînes de β (*bêta*)-globine et deux chaînes de α (*alpha*)-globine (Figure 1). Chaque chaîne de globine est associée à une molécule d'hème sur laquelle l'oxygène peut se lier de façon réversible. L'hémoglobine est la protéine la plus abondante des globules rouges avec une concentration moyenne de ~ 15 g par 100 mL chez l'humain.¹ Cette abondance couplée à la facilité à purifier l'hémoglobine, en a fait l'une des premières protéines découvertes et l'une des plus étudiées. Bien que préalablement identifiée chez les vertébrés, on sait maintenant que l'hémoglobine sous diverses formes se retrouve dans pratiquement toute la biosphère,^{2,3} incluant les bactéries,⁴ les plantes,⁵ les mycètes⁶ et les animaux, si bien que l'on date l'apparition du gène ancestral de l'hémoglobine à près de 1,8 milliards d'années. Dans le règne animal, l'hémoglobine a évolué en trois grandes familles d'hémoprotéines qui se distinguent entre autre par leur couleur vive caractéristique: l'hémoglobine est le pigment rouge du sang de la plupart des animaux supérieurs, l'hémocyanine est la substance responsable de la couleur bleue du sang de plusieurs arthropodes et mollusques alors que l'hémérithrine, de couleur bourgogne, est responsable de la couleur caractéristique des globules roses de quelques invertébrés vermiformes inférieurs.^{7,8}

Si elle est présente sous forme tétramérique ($\alpha_2\beta_2$) chez les vertébrés, l'hémoglobine n'est pas restreinte à cette composition. En effet, on peut la retrouver sous forme monomérique, dimérique, tétramérique et même multimérique chez certains animaux primitifs tel que les vers vestimentifères que l'on retrouve près des cheminées hydrothermales dans les profondeurs océaniques et qui possèdent des hémoglobines étant formées de 24 et 144 chaînes de globine.⁹⁻¹⁸

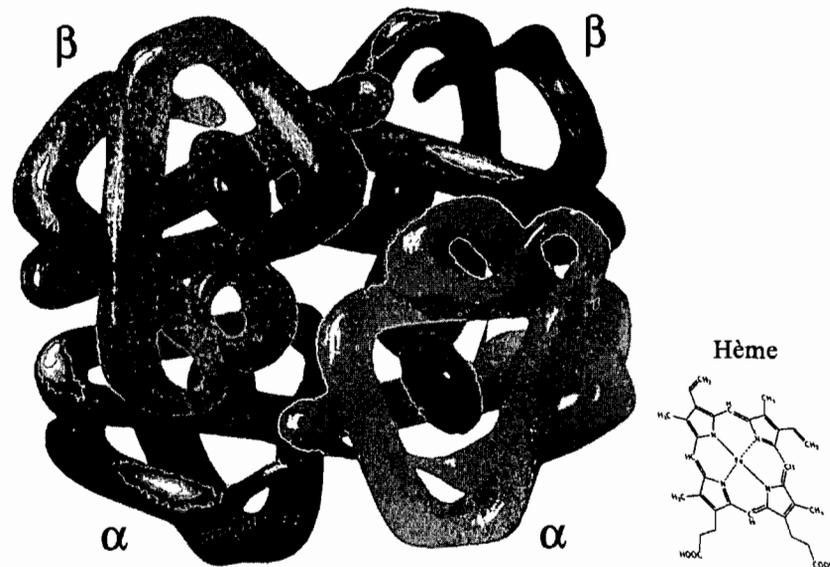


Figure 1: Structure de l'hémoglobine.

Qu'elle s'appelle hémoglobine, leghémoglobine, hémocyanine ou hémérithrine, et quelque soit sa structure quaternaire, la fonction générale de l'hémoglobine demeure presque exclusivement associée au transport ou à la séquestration de l'oxygène, ce qui laisse suggérer que le gène ancestral de l'hémoglobine pourrait être apparu dans la biosphère en réponse à l'élévation des niveaux d'oxygène qui a eu lieu durant l'ère paléoprotozoïque il y a quelques 2 milliards d'années.¹⁹ Ainsi, il est probable qu'au départ, la fonction de l'hémoglobine primitive en fut une de protection en permettant la séquestration de l'oxygène, protégeant ainsi les organisme contre l'effet oxydant de se nouveau gaz qui n'existait que sous forme de traces dans l'atmosphère terrestre. Encore aujourd'hui cette fonction est conservée par certaines hémoglobines comme la leghémoglobine que l'on retrouve dans les nodules de réduction de l'azote chez certaines plantes, où un milieu anaérobie est mandataire à cette fonction.²⁰ Puis, les organismes évoluèrent et se mirent à utiliser l'oxygène dans leur métabolisme, principalement comme molécule échangeuse d'électron. Cet élément auparavant poison, devint alors vital et grâce à la proto-

hémoglobine ancestrale, il pouvait ainsi être utilisé dans certains processus; la chaîne respiratoire venait de voir le jour.²¹

Avec la multicellularisation des organismes vivants, le transport de l'oxygène vers les cellules non-exposées à l'air devenait critique. Chez les animaux, le besoin en oxygène des différents organes internes et plus particulièrement les tissus à haute consommation énergétique comme la fibre musculaire, étant accru, le transport de l'oxygène a été pris en charge par des cellules circulantes spécialisées, riches en hémoglobine: les globules rouges ou érythrocytes. Il devenait alors impératif de contrôler l'affinité de l'hémoglobine pour l'oxygène de façon à faciliter l'apport d'oxygène aux tissus qui en avaient le plus besoin. Le niveau de CO₂ ainsi que le pH sont deux effecteurs de l'affinité de l'hémoglobine pour l'oxygène mais ces deux facteurs sont difficilement contrôlables d'un point de vue métabolique.²² Chez les animaux, l'utilisation de phosphates organiques comme l'ATP, le GTP, le 1,3,4,5,6-inositol pentaphosphate (IPP) et le 2,3-diphosphoglycérate (2,3-DPG) pour réduire l'affinité de l'hémoglobine pour l'oxygène a permis de développer un système de contrôle pouvant être régulé beaucoup plus finement dans le temps et même localement.^{23;24} Chez les animaux ce contrôle est particulièrement important puisque la grande majorité des espèces passe par une étape dite "embryonnaire" lors du développement et que souvent cette étape se fait dans des conditions très différentes des conditions vécues par les adultes. Des exemples peuvent être donnés dans toutes les strates de la vie animale: vie larvaire aquatique et vie adulte aérienne chez de nombreux insectes et batraciens, étape embryonnaire *in ovo* chez les oiseaux et les reptiles et intra-utérine chez les mammifères.²⁴ Même au stade adulte, les besoins en oxygène peuvent varier grandement selon les situations comme la vie en altitude et la plongée sous-marine prolongée. L'effort physique intense a également pour effet d'augmenter brusquement les besoins en oxygène au niveau des muscles, ce qui nécessite un apport massif en oxygène au niveau des tissus en état d'hypoxie et peut être obtenu en augmentant les niveaux de 2,3-DPG au niveau des globules rouges.^{25;26}

Chez les espèces de vertébrés supérieurs tel que les reptiles, les oiseaux et les mammifères, une hémoglobine spécialisée est exprimée au stade embryonnaire précoce puis est remplacée par l'hémoglobine dite fœtale/adulte qui elle restera exprimée tout au long de la vie de l'animal. Généralement, l'hémoglobine embryonnaire possède une affinité intrinsèque pour l'oxygène plus grande que l'hémoglobine de la mère ce qui favorise les échanges d'oxygène de la mère vers l'embryon.²⁷ Le cas des mammifères est particulier puisqu'en plus du stade embryonnaire, tout le stade fœtal est également intra-utérin. Le fœtus se retrouve donc en compétition avec sa mère durant toute la période fœtale pour l'obtention de l'oxygène nécessaire à sa survie. Fait à noter, même chez les mammifères, l'hémoglobine embryonnaire disparaît très tôt pour être remplacée par l'hémoglobine fœtale/adulte au début du stade fœtal. Pour palier à ce problème, trois grandes "stratégies" se sont développées au cours des quelques derniers 100 millions d'années.^{28;29}

1) La stratégie employée par le plus grand nombre d'espèces de mammifères est de diminuer les niveaux de 2,3-DPG érythrocytaire au stade fœtal par rapport au stade adulte où le 2,3-DPG est présent à des concentrations molaires comparables aux concentrations de l'hémoglobine. Ceci a pour effet de diminuer l'affinité de l'hémoglobine maternelle à l'oxygène au profit de celle du fœtus.²⁹

2) Chez les ruminants, où les niveaux de 2,3-DPG sont très faibles aussi bien au stade fœtal qu'au stade adulte, une hémoglobine strictement fœtale est produite ayant une affinité intrinsèque pour l'oxygène plus élevée que l'hémoglobine présente chez l'adulte. Chez ces animaux, l'hémoglobine embryonnaire est remplacée par l'hémoglobine fœtale très tôt au cours du développement, qui est à son tour remplacée à la naissance par l'hémoglobine adulte.³⁰

3) Les primates simiens, dont fait partit l'homme, se caractérisent par un temps de gestation assez long permettant le développement d'un cerveau plus complexe. Ce développement prolongé du système nerveux demande un apport en oxygène particulièrement important et critique. Les niveaux de 2,3-DPG ne présentent pourtant qu'une variation

modeste entre les différents stades du développement.³¹ Comme les ruminants, les simiens produisent aussi une hémoglobine strictement fœtale qui possède cependant une affinité intrinsèque pour l'oxygène équivalente à l'hémoglobine adulte. Ce qui la distingue toutefois, c'est sa faible capacité à lier le 2,3-DPG, ce qui la rend "insensible" à cet effecteur. Ceci a comme effet net d'accroître considérablement l'affinité de l'hémoglobine du fœtus pour l'oxygène en comparaison à l'hémoglobine des adultes, qui elle est sensible au 2,3-DPG.³²

Ces différentes hémoglobines sont codées par des gènes paralogues distribués sur 2 locus distincts. Le locus de la α -globine contient tous les gènes de globine de type α (ζ (*zêta*), α_2 , α_1 et θ (*thêta*) chez l'humain) alors que le locus de la β -globine contient tous les gènes des globine de type β (ϵ (*epsilon*), γ (*gamma*), δ (*delta*) et β chez l'humain). L'hémoglobine embryonnaire chez l'homme est en fait un mélange de trois hémoglobines; les hémoglobines Gower-1 ($\zeta_2\epsilon_2$), Gower-2 ($\alpha_2\epsilon_2$) et Portland-1 ($\zeta_2\gamma_2$).³³⁻³⁷ L'hémoglobine fœtale (HbF) est constituée de deux chaînes α et deux chaînes γ ($\alpha_2\gamma_2$) alors que l'hémoglobine adulte est constituée de deux hémoglobines distinctes: la première, la HbA ($\alpha_2\beta_2$), est l'hémoglobine prédominante chez l'adulte, représentant plus de 95% de l'hémoglobine totale alors que la HbA2 ($\alpha_2\delta_2$) ne dépasse à peu près jamais 3% de l'hémoglobine totale adulte. Un ratio équimolaire des chaînes de type α et β dans les globules rouges est absolument nécessaire puisque tout déséquilibre de ces chaînes produit un phénotype délétère pour les globules rouges appelé thalassémie. Ces maladies génétiques dont la gravité des symptômes varie selon la nature de la thalassémie sont très répandues dans le monde et particulièrement au niveau de la région méditerranéenne. Généralement, une β -thalassémie qui correspond à une anomalie dans la production du gène β et par conséquent, présente un excès de chaînes α , donne des symptômes plus sévères que la α -thalassémie qui elle, présente un excès de chaîne β .³⁸

1.2 La β -globine: l'histoire de... cinq gènes

1.2.1 Le pigment respiratoire

Dès la première moitié du 19^e siècle, les chercheurs se sont intéressés au pigment alors appelé cruorine, qui était responsable de la couleur rouge intense des corpuscules du sang. En 1853, L'anatomiste polonais Ludwig Karol Teichmann identifiait une substance cristalline rougeâtre contenant du fer qu'il nomma cristal de Teichmann, plus tard nommé hémine et qui pouvait être extraite de la cruorine simplement en la traitant avec de l'acide acétique glaciale ce qui ne laissait qu'un produit incolore.³⁹ En 1863, le physicien et mathématicien irlandais George Gabriel Stokes mesura le spectre optique de la cruorine lors de la respiration et conclut alors que l'oxygène se liait faiblement au composé inorganique du pigment.⁴⁰ Quelques mois plus tard, le physiologiste allemand Felix Hoppe-Seyler parvenait à cristalliser la cruorine et proposa que cette molécule était constituée de macromolécules globulaires de plus de 600 carbones, les "globulines" ou globines, ne contenant pas de Fer et de plus petites molécules contenant du fer, les "hématines" ou hème, ce qui fut confirmé en 1871 par William T. Preyer.^{41;42} Hoppe-Seyler proposa une nouvelle appellation pour la cruorine qu'il nomma "hémoglobine", nom qui fut dès lors adopté par la communauté scientifique.⁴¹ Ce ne sera toutefois qu'en 1928 que le rôle physiologique de l'hémoglobine dans la fonction respiratoire sera proposé sur des bases physico-chimiques solides dans un ouvrage majeur sur la fonction du sang par le physiologiste américain Lawrence J. Henderson.⁴³ Toutefois, la structure de cette molécule demeura inaccessible aux chercheurs pendant quelques décennies.

En 1949 les études sur l'hémoglobine firent un bond énorme quand deux études indépendantes, l'une par James V. Neel,⁴⁴ l'autre par le récipiendaire de deux prix Nobel, Linus Pauling,⁴⁵ démontrèrent le lien direct entre une maladie du sang qui faisait rage dans la population afro-américaine, la drépanocytose, et une hémoglobine anormale. Ces deux œuvres, et plus particulièrement celle de Pauling jetèrent les bases de la biologie

moléculaire moderne en prouvant qu'une maladie aussi complexe que la drépanocytose pouvait être provoquée par une modification dans une seule protéine et que cette "maladie moléculaire" était transmise de façon récessive. Pauling et ses collaborateurs émirent, en 1957, l'hypothèse que l'hémoglobine devait être formée de quatre chaînes hétérogènes de globine qui furent nommées l'année suivante, chaînes α et β .^{46,47} En 1959, le collaborateur de Pauling, Walter Adolf Schroeder démontra que la drépanocytose était due à une altération de la chaîne β alors que la chaîne α demeurerait inchangée.⁴⁸ Se sera toutefois le chimiste organicien de Cambridge, Vernon M. Ingram, qui identifia la nature de cette mutation qui se révéla être une substitution d'un acide glutamique par une valine en position six de la β -globine.⁴⁹ Presque au même moment, le biochimiste britannique qui allait plus tard recevoir un prix Nobel pour ses travaux sur la diffraction des rayons X effectuées sur l'hémoglobine, Max Ferdinand Perutz résolu la structure tridimensionnelle de l'hémoglobine à une résolution de 5,5 Å. On connaissait enfin le visage de cette molécule qui avait fasciné les chercheurs depuis plus de 200 ans.⁵⁰

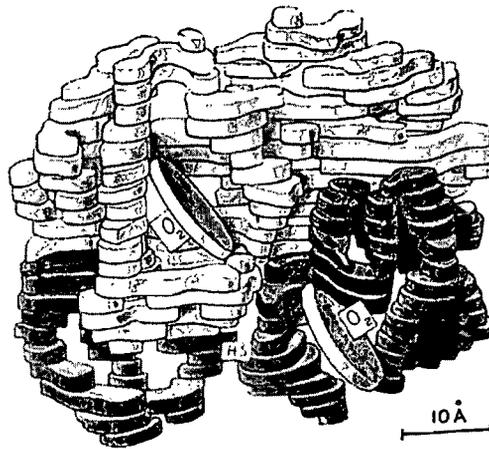


Figure 2: Structure de l'hémoglobine tel que présentée par Perutz à son discours du prix Nobel (1962).⁵¹

1.2.2 La respiration fœtale

Alors que les chercheurs se sont affairés pendant deux siècles à déterminer la composition, la structure et la fonction de la matière rouge du sang, un autre champ d'étude apparemment non-lié aux recherches sur l'hémoglobine, intéressait les scientifiques: la fonction respiratoire chez le fœtus en développement. Dès 1674, le physiologiste anglais John Mayow affirmait que les fœtus, tout comme les adultes, avaient besoin de respirer pour pouvoir se développer. Il proposa même que le placenta agissait comme un poumon utérin, permettant le passage des particules nitro-aériennes vers l'embryon.⁵² Les travaux de Mayow étaient certainement très en avance sur leur temps et il fallut attendre deux siècles avant que le professeur allemand en physiologie animale, Nathan Zuntz, ne démontre en 1877 qu'effectivement l'oxygène que recevait le fœtus lui était fourni par la mère via le cordon ombilical et le placenta. Il démontra également que le fœtus et la mère étaient en compétition pour l'oxygène puisque lorsque une mère lapine était asphyxiée, les fœtus mourraient en même temps qu'elle à moins que les cordons ombilicaux ne soient ligaturés ce qui prolongeait la vie des fœtus de 10 à 20 minutes.⁵³ Ce fut finalement en 1927 que le physiologiste anglais Arthur St George Huggett démontrait chez la chèvre, que cette compétition était possible grâce à une plus grande affinité du sang du fœtus pour l'oxygène que le sang de la mère.⁵⁴ Bien qu'on se doutait déjà que l'hémoglobine était impliquée dans l'affinité du sang pour l'oxygène, La nature moléculaire de la différence d'affinité observée chez le fœtus et chez l'adulte ne demeura toutefois qu'imparfaitement comprise pendant plusieurs années.

Le physiologiste britannique Earl F. McCarthy fut bien près toutefois de découvrir l'explication dès 1933 lorsque ses analyses de dissociation de l'oxygène de l'hémoglobine de chèvre diluée obtenue par la lyse des cellules du sang, démontrèrent que l'hémoglobine du fœtus avait une affinité pour l'oxygène supérieure à l'hémoglobine du sang maternel. McCarthy proposa alors l'existence d'hémoglobines distinctes chez le fœtus et la mère.⁵⁵ L'année suivante, le zoologiste américain Frank Gregory Hall obtenait des résultats simi-

lares avec l'hémoglobine de lapin et de poulet.^{56;57} Toutefois, les résultats que McCarthy obtint de ses travaux sur du sang humain en 1943 compliquèrent considérablement l'interprétation de la théorie d'une hémoglobine fœtale à forte affinité pour l'oxygène permettant les échanges d'oxygène de la mère vers le fœtus. En effet, il mesura que l'affinité pour l'oxygène de l'hémoglobine fœtal humaine diluée obtenue par la lyse des globules rouges n'était pas supérieure, mais plutôt inférieure à celle de l'hémoglobine de la mère obtenue de la même façon.⁵⁸ Ce résultat était opposé à celui obtenu avec l'hémoglobine d'autres espèces, mais aussi aux résultats obtenus sur le sang total et sur des globules rouges humains en suspension.⁵⁹ McCarthy en vint donc à la conclusion que par un mécanisme indéterminé, le globule rouge maternel causait une diminution de l'affinité de l'hémoglobine pour l'oxygène alors que le globule rouge fœtal n'affectait que minimalement cette affinité.⁵⁸ En 1953, le professeur en pédiatrie à l'université de Harvard, Clement A. Smith démontra en analysant l'hémoglobine fœtale humaine purifiée cette fois par dialyse, qu'elle avait une affinité pour l'oxygène ne pouvant être distinguée de celle de l'hémoglobine adulte.⁶⁰ Smith suggéra donc que la différence d'affinité pour l'oxygène entre l'hémoglobine fœtale et adulte résultait d'une différence dans l'environnement soluble où se trouve l'hémoglobine plutôt que dans une différence intrinsèque aux molécules d'hémoglobine elles-mêmes. Toutefois ce ne sera qu'à la fin des années soixante que l'on comprendra ce qui régule les propriétés de liaison de l'oxygène à l'hémoglobine dans les globules rouges adultes.

Pourtant, dès 1925 le chercheur américain Isidor Greenwald décrivait une nouvelle molécule de phosphore organique présent en grande quantité dans les globules rouges et qu'il nomma acide diphospho-*l*-glycérique, qui fut connu plus tard sous l'appellation de 2,3-diphosphoglycérate ou 2,3-DPG.⁶¹ La fonction de cette molécule et son lien avec l'hémoglobine demeurèrent toutefois inconnus jusqu'en 1967 lorsque deux groupes identifièrent le 2,3-DPG comme un modulateur puissant de l'affinité de l'hémoglobine pour l'oxygène.^{62;63} En 1961, la découverte de deux autres hémoglobines présentes seulement chez de très jeunes embryons humain et appelées hémoglobine embryonnaire Gower-1 et Gower-2

venait compliquer les choses,⁶⁴ mais ce ne sera finalement qu'au milieu des années 70 que deux grandes études comparant les hémoglobines embryonnaires, fœtales et adultes, leur réactivité au 2,3-DPG ainsi que leur affinité intrinsèque pour l'oxygène chez plusieurs dizaines d'espèces de mammifères, permirent de mieux comprendre les mécanismes qui contrôlaient la respiration fœtale chez les animaux.^{28;65;66}

1.2.3 L'ère de la génétique

Alors que les recherches essayant de déterminer la structure et la fonction de l'hémoglobine faisaient de grand progrès, la biologie moléculaire connut une importante révolution vers la fin des années 50. Lorsque les chercheurs britanniques Francis H. Crick et James D. Watson présentèrent la structure en double hélice de l'ADN en 1953,⁶⁷ l'ère de la génétique moléculaire venait de débiter. Ce ne sera toutefois qu'au début des années 70 que la génétique des globines allait faire un grand bond suite à la découverte en série des enzymes de restriction, de la transcriptase inverse, et des premiers plasmides pouvant se répliquer de façon autonome en bactérie et permettre le "clonage" de séquences d'ADN.⁶⁸⁻⁷⁴

Grâce à ces technologies, le premier gène fut cloné en 1975 par Bernard Mash de l'université de Genève. Il s'agissait de l'ADN complémentaire de la β -globine du lapin obtenu par la transcription inverse de l'ARN messager de la β -globine.⁷⁵ En 1977, un fragment d'ADN génomique contenant le gène de la β -globine du lapin fut cloné par Richard Flavell, alors chercheur à l'université d'Amsterdam. Flavell démontra que la séquence codante du gène de la β -globine était interrompue par l'insertion d'un large fragment non codant mais transcrit.⁷⁶⁻⁷⁸ Lorsque le gène de la globine de souris fut cloné à son tour en 1978, il ne fallut pas bien longtemps aux chercheurs pour observer que ce gène était également interrompu par une séquence non codante alors appelée "*intervening sequence*" ou séquence "d'entre-temps".⁷⁹ Suite à cette découverte, l'identification de fragments interrompant les séquences codantes de nombreux gènes chez plusieurs espèces animales et même chez la levure mena la communauté scientifique à considérer ces séquences comme étant la norme plutôt que l'exception et ils furent appelés introns.⁸⁰⁻⁸⁵

Toujours en 1978, le clonage du gène foetal humain γ -globine⁸⁶ ainsi que des gènes adultes humains δ - et β -globine⁸⁷ permis de démontrer que ces trois gènes étaient en fait situés à proximité les uns des autres sur le chromosome 11 et formaient un locus de gènes de type β -globine.⁸⁸ Le clonage l'année suivante d'un deuxième gène foetal⁸⁹ et du gène embryonnaire ε -globine⁹⁰ mènera finalement en 1980 à la description, par l'équipe de recherche du professeur Tom Maniatis, de l'organisation générale des gènes du locus de la β -globine 5'- $\psi\beta_2$ - ε - $\overset{G}{\gamma}$ - $\overset{A}{\gamma}$ - $\psi\beta_1$ - δ - β -3' où $\psi\beta_1$ et $\psi\beta_2$ étaient décrits comme des pseudogènes de type β -globine.⁹¹ Deux ans plus tard, le pseudogène $\psi\beta_2$ en amont de ε fut éliminé et le locus fut présenté sous sa forme définitive: 5'- ε - $\overset{G}{\gamma}$ - $\overset{A}{\gamma}$ - $\psi\beta$ - δ - β -3'.⁹² Toutefois, il manquait encore quelque chose.

En effet, dès 1985 les premières études d'expression des gènes individuels de globine humaine en souris transgéniques démontrèrent que les promoteurs proximaux des gènes étaient suffisants pour diriger l'expression spécifique embryonnaire/fœtale et adulte des différents gènes. Cependant, les niveaux d'expression obtenus ne représentaient qu'une infime fraction des niveaux normalement observés, indiquant que quelque chose manquait.⁹³⁻⁹⁶ Il fallut près de deux ans à l'équipe du professeur Mark Groudine, chercheur au centre de recherche sur le cancer Fred Hutchinson à Seattle, pour proposer l'existence d'une région d'activation du locus (LAR) ou région de contrôle dominante (DCR) située en amont du gène ε en se basant sur la découverte faite quelques temps auparavant par Dorothy Tuan *et al* d'une série de sites d'hypersensibilité à la DNaseI (HS) en amont du locus.⁹⁷⁻⁹⁹ Le rôle des HS comme régions d'activation du locus fut démontré ensuite à l'aide d'une série d'expériences où le gène de la β -globine était relié au LAR complet ou seulement à différent HS séparément et qui démontrèrent une expression élevée de β dans les globules rouges.¹⁰⁰⁻¹⁰² Des études subséquentes en souris transgéniques qui utilisèrent des constructions où le LAR était lié au gène de la β -globine ou γ -globine et où l'expression fut étudiée à différents stades du développement, démontrèrent cependant que si les niveaux d'expression étaient effectivement élevés, la spécificité développementale était

perdue puisque les gènes étaient exprimés à tous les stades du développement. Toutefois, des constructions bigéniques contenant à la fois γ et β liés au LAR, permettaient d'obtenir une expression correcte de γ et β au cours du développement, suggérant que la régulation des gènes de globine se faisait par un mécanisme de compétition pour le LAR.¹⁰³ La nomenclature concernant le LAR et les différents sites d'hypersensibilité à la DNaseI posait à cette époque quelques problèmes et manquait de rigueur et en 1990 l'appellation officielle "région de contrôle du locus" ou LCR fut adoptée pour toute séquence capable de conférer une expression indépendante de la position d'intégration à un transgène et une numérotation 5'HS5 à HS1 fut donnée au cinq sites d'hypersensibilité à la DNaseI situés en amont de ϵ .¹⁰⁴

Les études en transgéniques se poursuivirent et de nombreux éléments régulateurs et facteurs de transcription furent identifiés. Cependant, les limitations de cette approche devinrent bientôt flagrantes puisqu'il apparut évident que la structure générale du locus était essentielle pour la régulation des différents gènes de globine au cours du développement. Le problème était que la taille excessive du locus (~70 kb) était beaucoup trop grande pour qu'il puisse être entièrement incorporé dans les vecteurs traditionnels ou même dans les vecteurs à plus grande capacité comme les cosmides qui étaient limités à des inserts d'environ 40 kb;¹⁰⁵ l'obtention de souris transgéniques contenant le locus entier de la β -globine humaine demeurait donc irréalisable. Pourtant, dès 1987, un nouveau type de vecteur à grande capacité avait été développé, il s'agissait des chromosomes artificiels de levure ou YAC, pouvant accommoder des inserts de plus de 1000 kb.¹⁰⁶ De plus, il s'avéra que la mutagenèse dirigée de YAC était assez aisée grâce à la machinerie de recombinaison homologue inhérente à la levure.¹⁰⁷ En 1991, le locus de la β -globine humaine fut cloné dans deux YAC faisant 150 et 248 kb et les premières souris transgéniques faites à partir de YAC furent produites dès 1992 dont certaines, portant le locus entier de la β -globine humaine, montraient une régulation des différents gènes du locus qui suivait celle des gènes endogènes de la souris.¹⁰⁸⁻¹¹² Enfin, la science possédait un modèle "idéal" pour l'étude de la régulation du locus humain de la β -globine.

1.3 Le locus humain de la β -globine

Chez l'humain, le locus de la β -globine est localisé sur le chromosome 11 et est constitué de cinq gènes fonctionnels paralogues s'étalant sur une distance d'environ 40 kb et flanqués de six sites d'hyperméthylabilité à la DNaseI dont les quatre premiers en amont (HS1 à HS4), répartie sur une distance de ~ 15 kb, forment la région de contrôle du locus ou LCR (Figure 3). Les gènes de globine humaine, comme tous les gènes de globine chez les animaux, sont constitués de trois exons et de deux introns dont la position est demeurée très conservée parmi les espèces.

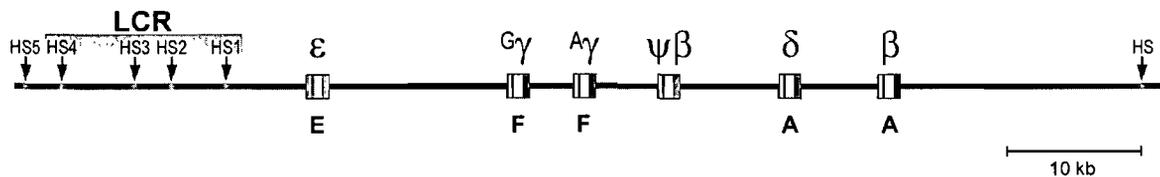


Figure 3: Le locus humain de la β -globine. Les flèches représentent les sites d'hyperméthylabilité à la DNaseI (HS) et le LCR est identifié par un dégradé de gris; Les boîtes représentent les gènes avec leurs exons (couleur) et leurs introns (blanc). La lettre sous chaque gène indique la période du développement où le gène est exprimé (E: embryonnaire, F: foetal, A: adulte). La couleur des gènes représente l'origine évolutive (voir section 1.4) selon le code: (◦): ϵ , (●): γ , (●): η , (◦): δ , (●): β , (●): indique un pseudogène.

1.3.1 La commutation de l'hémoglobine

Les gènes du locus ne sont pas tous exprimés au même moment au cours du développement. Ils sont plutôt exprimés de façon séquentielle dans l'ordre de leur position dans le locus (Figure 4). Le remplacement d'une globine par une autre lors du développement est nommé commutation de l'hémoglobine et chez l'humain, se produisent deux de ces commutations: la première au moment du passage du stade embryonnaire au stade foetal et l'autre, à la naissance. Fait intéressant, ces commutations ne sont pas

strictement génétiques puisqu'elles sont caractérisées par un changement presque simultané au niveau du site de l'érythropoïèse, c'est-à-dire que le site où sont produits les globules rouges change également au même moment (Figure 4). Il est ainsi probable que ces changements d'environnement influencent la régulation des gènes de globine en présentant des ratios différentiels des nombreux facteurs de transcription impliqués dans l'expression de ces gènes au cours du développement.¹¹³

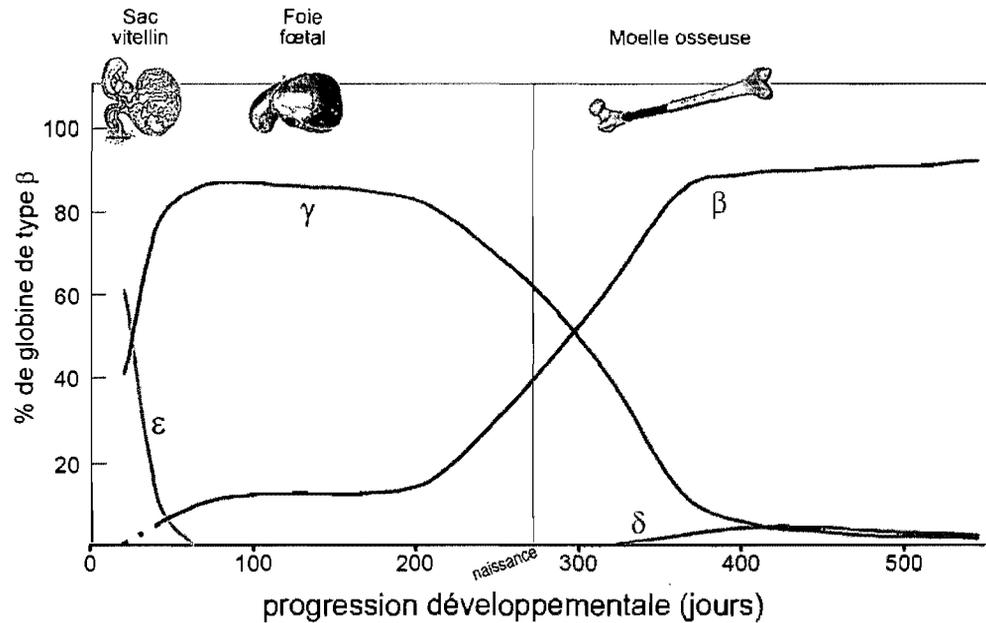


Figure 4: Commutation de l'hémoglobine chez l'humain. Adapté de Kitchen *et al.*²⁸

La commutation de l'hémoglobine embryonnaire à fœtal

À cause de la difficulté à obtenir des échantillons embryonnaires humains, la première commutation qui touche l'hémoglobine a été assez peu étudiée. Toutefois, on sait qu'à l'âge d'environ 5 semaines après la conception, alors que l'hématopoïèse primitive a lieu dans le sac vitellin, les érythrocytes de l'embryon sont nucléés et diffèrent donc

considérablement des érythrocytes de l'adulte.¹¹⁵ Ces érythrocytes appelés, érythrocytes primitifs, en plus d'avoir un noyau, se caractérisent à cette période par l'expression presque exclusive de l'hémoglobine embryonnaire Gower-1. Entre 6 et 7 semaines, la composition en hémoglobine change légèrement et est dès lors constituée d'un mélange d'hémoglobines Gower-1, Gower-2, Portland-1 et HbF. La diversité des hémoglobines dans ces érythrocytes est due à la présence simultanée des chaînes de globine ϵ , γ , ζ et α . Graduellement, ces érythrocytes primitifs sont remplacés par les érythrocytes définitifs lorsque l'hématopoïèse définitive prend place dans le foie fœtal. Ces érythrocytes définitifs appelés cellules F, sont énucléés et leur contenu en hémoglobine est constitué au début surtout de HbF et d'un peu de HbA et de Gower-2 puisque la chaîne ζ est réprimée et que la chaîne β n'est exprimée qu'à très faibles niveaux. Puis, la chaîne ϵ disparaît à son tour et seule la HbF et un peu de HbA sont présentes. Entre deux et trois mois, les érythrocytes primitifs sont complètement remplacés par les cellules F et on ne retrouve les hémoglobines embryonnaires que sous forme de traces; la commutation de l'hémoglobine embryonnaire à fœtale est complétée (Figure 4).^{64;114;115}

La commutation de l'hémoglobine fœtale à adulte

À la naissance, le sang de l'enfant est toujours constitué essentiellement de cellules F. Au cours des premiers mois de la vie, la composition en hémoglobine des érythrocytes change graduellement et la chaîne β devient rapidement prédominante sur les γ . Une deuxième hémoglobine adulte, la HbA₂, apparaît alors mais ses niveaux ne dépassent jamais 2-3 % de l'hémoglobine totale.¹¹⁶ Avant l'âge de six mois, l'hémoglobine HbF ne représente plus que 1-2 % de l'hémoglobine totale et ce niveau est conservé tout au long du stade adulte. Également à la naissance, le site de l'hématopoïèse définitive se relocalise dans la moelle osseuse où elle restera active pour toute la vie. Si la commutation de l'hémoglobine fœtale à adulte se caractérise principalement par le remplacement de la HbF par la HbA, on observe également un reversement des ratios d'expression des deux chaînes γ . En effet, si au stade fœtal $^G\gamma$ représente 70% et $^A\gamma$, 30% des chaînes γ , chez l'adulte, $^A\gamma$

devient prédominante, formant plus de 50% des chaînes γ encore exprimées au stade adulte.^{113;117}

1.3.2 β - et δ -globine

Les gènes de la β -globine et de la δ -globine sont les deux gènes exprimés chez l'adulte. β est la chaîne de globine qui avec la chaîne α forme l'hémoglobine HbA ($\alpha_2\beta_2$) qui est majoritaire alors que δ forme l'hémoglobine HbA2 ($\alpha_2\delta_2$), dont les niveaux n'excèdent pas 3% de l'hémoglobine totale et qui ne semble pas avoir de rôle d'importance physiologique. D'un point de vue évolutif, δ et β sont issus de la duplication d'un gène β ancestral qui a eu lieu avant l'émergence des mammifères euthériens il y a 80-100 millions d'années.¹¹⁸ Chez l'humain, le gène δ a acquis des séquences de β par conversion génique et par conséquent, partage avec ce gène une portion de sa séquence codante.¹¹⁹ Une perte d'éléments de régulation dans le promoteur de δ explique les niveaux très bas d'expression de ce gène en comparaison au gène adulte.^{120;121} L'affinité intrinsèque de l'hémoglobine HbA pour l'oxygène est assez élevée mais une grande sensibilité de cette hémoglobine pour le 2,3-DPG fait en sorte que cette affinité est considérablement réduite en condition physiologique. De son côté, l'hémoglobine HbA2 possède une affinité pour l'oxygène moins grande que la HbA mais répond également moins fortement au 2,3-DPG.¹²²

1.3.3 $A\gamma$ - et $G\gamma$ -globine

Les gènes $A\gamma$ et $G\gamma$ sont les deux gènes exprimés au stade fœtal chez l'humain et forment avec la chaîne α , l'hémoglobine HbF. Les deux gènes γ sont issus de la duplication du gène γ ancestral juste avant l'émergence des simiens il y a ~58 million d'années. Ce gène γ ancestral fut lui-même issu d'un événement de duplication du gène embryonnaire ϵ il y a 80-100 millions d'années. D'ailleurs, chez la grande majorité des espèces mammifères, le gène γ possède toujours une expression embryonnaire plutôt que fœtale. D'un point de vue évolutif, il est intéressant de noter que le recrutement fœtal de l'expression des gènes γ chez

les grands singes a eu lieu au même moment que la duplication et il n'est pas exclu que cette duplication ait eu un rôle à jouer dans le changement ontogénique de ces gènes. Il n'est pas connu si les deux γ -globines ont des fonctions physiologiques distinctes et/ou si elles forment deux hémoglobines différentes de façon préférentielle puisque les techniques d'analyse des tétramères d'hémoglobine ne permettent pas de distinguer les formes $\alpha_2^A\gamma_2$, $\alpha_2^G\gamma_2$ et $\alpha_2^A\gamma^G\gamma$. Cependant, puisque l'unique différence entre les chaînes $^A\gamma$ et $^G\gamma$ est la substitution d'un seul acide aminé en position 136 (Gly→Ala) il est probable que ces deux chaînes se comportent de façon parfaitement identique. L'affinité intrinsèque de l'hémoglobine HbF pour l'oxygène est assez élevée et comparable à celle de l'hémoglobine HbA. Cependant, ce qui la distingue de cette dernière, c'est une substitution de l'acide aminé en position 143 de la chaîne γ (Ser) par comparaison à la chaîne β (His) et qui abolit la liaison du 2,3-DPG.¹²³ Cette caractéristique est importante puisqu'elle permet à l'hémoglobine du fœtus d'avoir une affinité pour l'oxygène supérieure à celle de l'hémoglobine de la mère ce qui favorise l'échange d'oxygène de la mère vers le fœtus. Cette capacité accrue à obtenir l'oxygène permis un temps de gestation plus long, une condition critique pour le développement du cerveau plus complexe et performant des primates.

1.3.4 ϵ -globine

ϵ -globine est la composante de type β de l'hémoglobine embryonnaire Gower-1 ($\zeta_2\epsilon_2$) et de l'une des deux hémoglobines semi-embryonnaire, la Gower-2 ($\alpha_2\epsilon_2$). D'un point de vue évolutif, ϵ a été formé par la duplication d'un gène ancestral proto- β il y a ~200 millions d'années, soit bien avant l'apparition des mammifères. Fait intéressant, ce gène ayant une expression strictement embryonnaire chez toutes les espèces animales semble également moins sujet à conversion génique et aux duplications que les autres gènes de globine. L'affinité intrinsèque de l'hémoglobine embryonnaire Gower-1 est plus élevée que l'hémoglobine HbA. Toutefois, celle de la Gower-2 est identique à celle de la HbA et même que la sensibilité de cette hémoglobine au 2,3-DPG est plus élevée que celle de la HbA suggérant que la forte affinité de l'hémoglobine embryonnaire serait plutôt due à la chaîne ζ

qu'à la chaîne ϵ , ce qui est confirmé par le fait que l'hémoglobine Portland-2 ($\zeta_2\beta_2$) réagit plutôt comme une hémoglobine embryonnaire avec l'oxygène.³⁷

1.3.5 La région de contrôle du locus (LCR)

Le LCR est le plus important élément régulateur distal du locus de la β -globine, situé à 6 kb en amont du gène ϵ . Il s'étend sur une distance d'environ 15 kb et est composé de plusieurs domaines démontrant une sensibilité extrême à la DNaseI, appelés sites d'hypersensibilité (HS). Le plus distal de ces HS, le HS5, a la particularité d'être sous sa forme ouverte de façon ubiquitaire alors que les quatre autres (HS1 à HS4) sont sous leur configuration permissive seulement dans les tissus érythroïdes. Le LCR est essentiel pour l'expression élevée de tous les gènes du locus et ce, à tous les stades du développement, en plus d'être capable de conférer une expression à un transgène, indépendante du site d'intégration.¹⁰⁰ Il est à noter que cette fonction est dépendante de l'orientation du LCR ce qui en fait un *enhancer* atypique. D'un point de vue évolutif, il est notable que la structure générale du LCR ainsi que l'organisation des différents HS qui le constitue sont remarquablement conservées parmi les espèces. Chacun de ces HS est important pour la fonction globale du LCR et ce, même si les HS pris individuellement sont capables de contrôler l'activation des gènes du locus avec plus ou moins d'efficacité. Les HS possèdent un cœur ou *core* qui est défini comme étant la séquence leur conférant leur fonction principale. Il est à noter que les HS2 et HS3 sont les deux sites d'hypersensibilité possédant la fonction activatrice la plus forte, bien que le HS4 ait également cette capacité alors que les HS1 et HS5 ne la possèdent pas.¹²⁴⁻¹²⁹

1.3.6 L'environnement chromosomique du locus

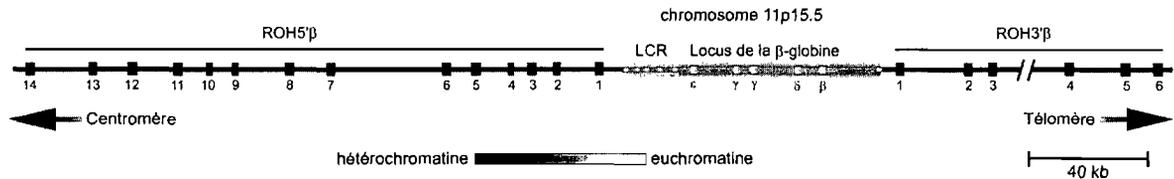


Figure 5: Région chromosomique entourant le locus de la β -globine.

La situation chromosomique du locus de la β -globine est particulière en ce sens que le locus est une région dite AT-riche, c'est-à-dire dont la proportion de A et de T (65%) surpasse fortement celle des G et des C (35%) ce qui est inhabituelle pour une région riche en gène.¹³⁰ Conséquemment, le locus est dépourvu d'îlots CpG qui sont généralement des éléments régulateurs importants des gènes. Une autre caractéristique du locus est son positionnement sur le chromosome 11; en effet, le locus est intégré dans une région contenant une série de gènes codant pour des récepteurs olfactifs (ROH) qui ne sont pas exprimés dans les cellules érythroïdes, mais seulement dans les cellules neuronales de l'épithélium olfactif (Figure 5).^{131;132} De fait, les séquences flanquantes du locus se retrouvent dans un état fortement hétérochromatique dans les tissus érythroïdes alors que le locus de la β -globine est dans un état euchromatique dans ces cellules mais hétérochromatique dans les autres tissus. Comment le locus parvient-il à maintenir un état de la chromatine différent de celui de son environnement chromosomique n'est toujours pas très bien compris mais il est possible que le LCR et le site d'hypersensibilité à la DNaseI en 3' du locus agissent comme *insulators* qui protégeraient le locus contre l'effet de condensation de la chromatine provoquée par la région environnante.^{133;134} Il est à noter que toutes les études cherchant à identifier des éléments régulateurs potentiels de l'expression des gènes de globine et situés à l'extérieur du locus ont échoués chez l'humain ce qui suggère que le locus formerait une sorte d'entité autonome et donc, indépendante de son contexte

chromosomique.^{132;135;136} La seule exception serait l'insertion d'une longue répétition terminale (LTR) du rétrovirus résident humain ERV-9 à la frontière 5' du locus juste en amont du HS5 et qui pourrait avoir une fonction sur la régulation des gènes du locus.^{137;138}

1.3.7 Hémoglobinopathies de la β -globine

Depuis la découverte de Linus Pauling en 1949 démontrant que la drépanocytose était due à une hémoglobine anormale, plus d'un millier d'allèles mutants touchant différents gènes de globine ont été identifiés et caractérisés au niveau moléculaire ce qui en fait le groupe de maladies monogéniques ayant le spectre de mutations le plus répandu jusqu'à présent chez l'homme. Les mutations causant les différentes hémoglobinopathies peuvent toutefois être regroupées en trois grandes familles: le premier groupe comprend des maladies appelées thalassémies qui résultent de mutations provoquant une synthèse réduites d'une ou plusieurs chaînes de globines; le second groupe englobe les maladies qui résultent de la production de globines structurellement et/ou fonctionnellement anormales, la plus connue étant l'hémoglobine S, responsable de la drépanocytose; le troisième groupe inclut les maladies dites Thalassémique-fonctionnelles qui combinent ces deux défauts génétiques, c'est-à-dire l'expression faible d'une globine anormale comme l'hémoglobine Lepore ($\alpha_2[\delta\beta]_2$) issue de la fusion des gènes δ et β .¹³⁹⁻¹⁴² Une autre anomalie génétique quelquefois répertoriée comme hémoglobinopathie bien que ne présentant pas de symptômes pathologiques, est l'expression inappropriée du gène fœtal chez l'adulte.¹⁴³

L'anémie falciforme

Si plus de 700 variants de l'hémoglobine sont connus jusqu'à présent, le premier identifié et de loin le plus étudié est l'hémoglobine S (HbS), responsable de la drépanocytose, aussi connu sous le nom d'anémie falciforme à cause de la forme en faucille que prennent les globules rouges des patients. Cette hémoglobine a la caractéristique de pouvoir polymériser avec d'autres molécules d'hémoglobine S lorsque les niveaux d'oxygène sont faibles et de former des fibres rigides qui déforment l'érythrocyte et lui donnent sa forme

caractéristique. La mutation responsable de cette caractéristique est la substitution Glu₆→Val₆ sur la chaîne de la β-globine, ce qui produit une "zone" hydrophobe favorisant la polymérisation avec une autre HbS. Si cette mutation, qui provoque une maladie infantile grave et souvent létale sous sa forme homozygote, est si présente dans les populations originaires des pays tropicaux, c'est que sous sa forme hétérozygote qui est asymptomatique et appelée "trait drépanocytaire", elle confère une résistance naturelle au paludisme, une maladie endémique des régions chaudes du globe. Il est intéressant de noter que puisque l'hémoglobine fœtale possède une activité protectrice contre la polymérisation de la HbS, les nouveau-nés sont asymptomatiques jusqu'à l'âge d'environ 6 mois, où suite à la disparition de la HbF, les symptômes drépanocytaires surgissent finalement.¹⁴⁴

Les thalassémies

Le terme thalassémie provient du grecque *Θαλασσα (Thalassa)* signifiant "mer" et par extension, la Méditerranée et *αιμα (aima)* qui signifie "sang" et est utilisé pour désigner une série de maladies du sang très répandues dans les régions méditerranéenne et du sous-continent Indien, et caractérisées par un déséquilibre entre les chaînes de type β et α. Les thalassémies ou anémie de Cooley, du nom de celui qui décrivit pour la première fois cette maladie en 1925,¹⁴⁵ sont des maladies génétiques extrêmement hétérogènes en ce qui a trait à la sévérité des symptômes et sont classées en deux grandes familles selon la nature de la chaîne de globine se retrouvant en déficience. Les α-thalassémies regroupent toutes les maladies causées par l'absence ou une insuffisance de la chaînes α alors que les β-thalassémies regroupent les maladies dues à une carence en chaîne β.¹⁴⁶ Quelques formes de thalassémies plus rares peuvent toucher les autres chaînes de globine de type β sans affecter la β-globine elle-même. On parlera alors de γ-thalassémies ou de δ-thalassémies selon le gène altéré.^{147,148} Il existe également de larges délétions dans le locus qui font en sorte qu'aucune globine de type β ne soit exprimée de l'allèle touchée et qui causent certains des cas les plus sévères de β-thalassémies souvent désignées εγδβ-thalassémie. Généralement, toute délétion qui englobe le LCR tombe dans cette catégorie puisque

de telles délétions résultent en la perte d'expression de tous les gènes du locus et ce, peu importe le stade du développement.¹⁴⁹⁻¹⁵¹

Habituellement, la sévérité des β -thalassémies est plus grande que celle des α -thalassémies et ceci est dû au fait qu'un excès de chaînes β peut former une hémoglobine appelée hémoglobine H (β_4) qui est non-fonctionnelle mais soluble et peu toxique pour la cellule. À l'opposé, un excès de chaînes α ne peut former que des dimères peu solubles, qui précipitent à la membrane et provoquent des réactions oxydatives qui endommagent les érythrocytes.^{38;152;153}

La persistance héréditaire de l'hémoglobine fœtale (HPFH)

Normalement, la proportion d'hémoglobine fœtale HbF chez l'adulte ne dépasse pas 1-2 % et est le fait des quelques cellules F circulantes n'exprimant que la γ -globine alors que les autres érythrocytes ne présentent pas du tout d'HbF. Il arrive toutefois que certains individus présentent des niveaux de HbF anormalement élevés et ce, même au stade adulte. Cette anomalie génétique ne provoquant pas de symptômes pathologiques est appelée persistance héréditaire de l'hémoglobine fœtale ou HPFH (pour *hereditary persistence of fetal hemoglobin*). Chez les personnes qui en sont porteur, la HbF peut représenter de 2-30% de l'hémoglobine totale (Tableau I) et ce niveau peut atteindre jusqu'à 100% lorsque relié à certaines β -thalassémies.^{143;154} Généralement, dans une HPFH, l'augmentation de la HbF n'est pas le fait d'une présence accrue de cellule F, mais plutôt d'une expression pancellulaire de la γ -globine, bien que des cas de HPFH hétérocellulaire aient été répertoriés. Plusieurs mutations situées dans le promoteur des gènes fœtaux $G\gamma$ et $A\gamma$ ont été identifiées et qui provoquent une persistance de l'expression de ces gènes au stade adulte (Tableau I). Lorsque issue de larges délétions dans le locus résultant en β -thalassémies, la HPFH peut devenir un modulateur important de la sévérité de la maladie puisque des niveaux élevés de γ -globine peuvent compenser pour la perte de β -globine.¹⁵⁵ Il est à noter cependant que toute délétion des gènes adultes n'entraîne pas forcément une HPFH. De fait,

seul un certain nombre de ces délétions entraînant une activation des gènes fœtaux ont été identifiées (Figure 6). Dans les régions où la drépanocytose est endémique, les mutations HPFH sont également très présentes puisque même un faible niveau de HbF dans les cellules est suffisant pour inhiber la polymérisation de la HbS, conférant à la HbF un rôle protecteur dans cette maladie.¹⁵⁶

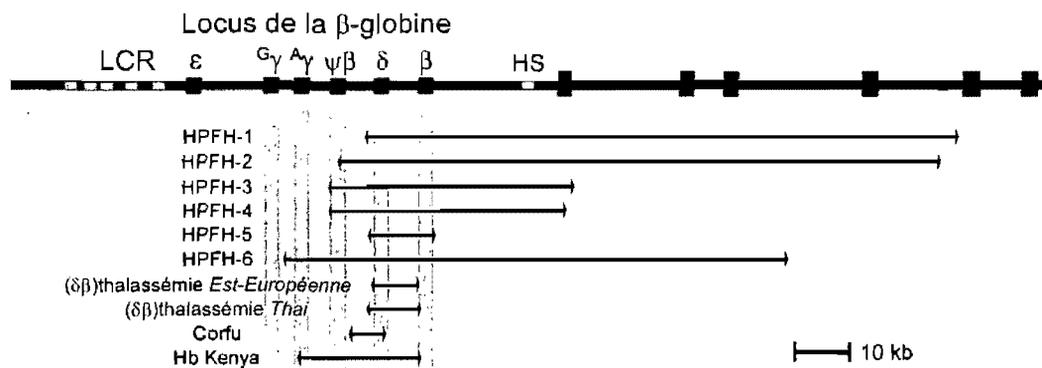


Figure 6: Délétions dans le locus conduisant à une HPFH.^{143;157;158}

Tableau I: Mutations HPFH non issues de délétions.^{143;159-162}

Categories et group racial	Mutation	% de HbF chez l'hétérozygote
G_γ HPFH (pancellulaire)		
Noir	G _γ : -202 (C→G)	15-25
Noir	G _γ : -175 (T→C)	20-30
Sarde	G _γ : -175 (T→C)	17-21
Japonais	G _γ : -114 (C→T)	11-14
Algérien	G _γ : -114 (C→A)	0,6-3,5
Hollandais	G _γ : -37 (A→T)	2,3
Λ_γ HPFH (pancellulaire)		
Sud Italien	Λ _γ : -196 (C→T)	12-16
Chinois	Λ _γ : -196 (C→T)	14-21
Grec	Λ _γ : -117 (G→A)	10-20
Sarde	Λ _γ : -117 (G→A)	12-16
Noir	Λ _γ : -117 (G→A)	11-16
Noir	Λ _γ : -114 à -102 (délétion)	30-32
G_γΛ_γ HPFH (pancellulaire)		
Chinois	Inconnue	20-25
Noir	G _γ : -158(C→T); Λ _γ : -175(T→C)	36-41
G_γ HPFH (hétérocellulaire)		
Noir (Atlanta)	G _γ G _γ : -158 (C→T)	2,3-3,8
Japonais	G _γ : -114 (C→T)	11-14
Λ_γ HPFH (hétérocellulaire)		
Noir	Λ _γ : -202 (C→T)	1,6-3,9
Britannique	Λ _γ : -198 (T→C)	3,5-10
Brésilien	Λ _γ : -195 (C→G)	4,5-7
Noir (Georgie)	Λ _γ : -114 (C→T)	2,6-6
G_γΛ_γ HPFH (hétérocellulaire)		
Suisse	Inconnue	1-4
Noir (Seattle)	Promoteurs γ normaux	3-8

1.4 Le locus de la β -globine chez les autres espèces

Le locus de la β -globine de tous les mammifères euthériens est issu d'un locus ancestral qui comportait déjà 5 gènes (ϵ - γ - η - δ - β) ainsi que le LCR constitué d'une série de sites d'hypersensibilité à la DNaseI très conservés entre les espèces (Figure 7). Les mammifères euthériens se distinguent ainsi des marsupiaux et des monotrèmes qui ne possèdent qu'un locus bigénique ϵ - β associé au LCR, ce qui démontre que la duplication du premier gène β pour donner les proto-gènes ϵ et β a eu lieu il y a de cela fort longtemps, soit avant la divergence évolutive des reptiles, des oiseaux et des mammifères. L'apparition des gènes plus spécialisés γ , η et δ n'a eu lieu que plus récemment, il y a quelques 100 millions d'années. Fait intéressant, l'une des caractéristiques génétiques les plus conservées chez tous les vertébrés supérieurs, est l'expression strictement embryonnaire du gène ϵ -globine. De fait, ce gène n'a subi que très peu de modifications après l'apparition des autres gènes embryonnaires γ et η ce qui tend à démontrer que la fonction de ce gène est primordiale pour la survie embryonnaire, et donc, que la pression sélective pour la conservation de cette fonction était très forte. À l'inverse, les autres gènes du locus ont subi de nombreuses modifications, des conversions géniques et des changements ontogéniques, démontrant que la pression sélective pour ces gènes était moindre, permettant une plus grande liberté d'adaptation aux besoins évolutifs de chaque espèce. Si la plupart des espèces présentent une expression embryonnaire des gènes ϵ , γ et/ou η et une expression des gènes β et δ débutant au stade fœtal, deux grands groupes se distinguent. En effet, les singes simiens ainsi que les ruminants ont des gènes spécialisés pour chacun des stades embryonnaires, fœtaux et adultes, voir même juvéniles chez certains ruminants (Figure 7). Toutefois, ces deux groupes parvinrent à cette spécialisation des gènes de globine par des mécanismes évolutifs différents et indépendants. Cette caractéristique représente un bon exemple d'une évolution convergente étant due à une force sélective commune à deux

groupes phylogénétiquement éloignés, dans ce cas-ci, la nécessité d'une période de gestation prolongée.

Cette section est consacrée à la description du locus de la β -globine de quelques espèces, sélectionnées en fonction de l'intérêt que peut représenter chacun de leurs cas.

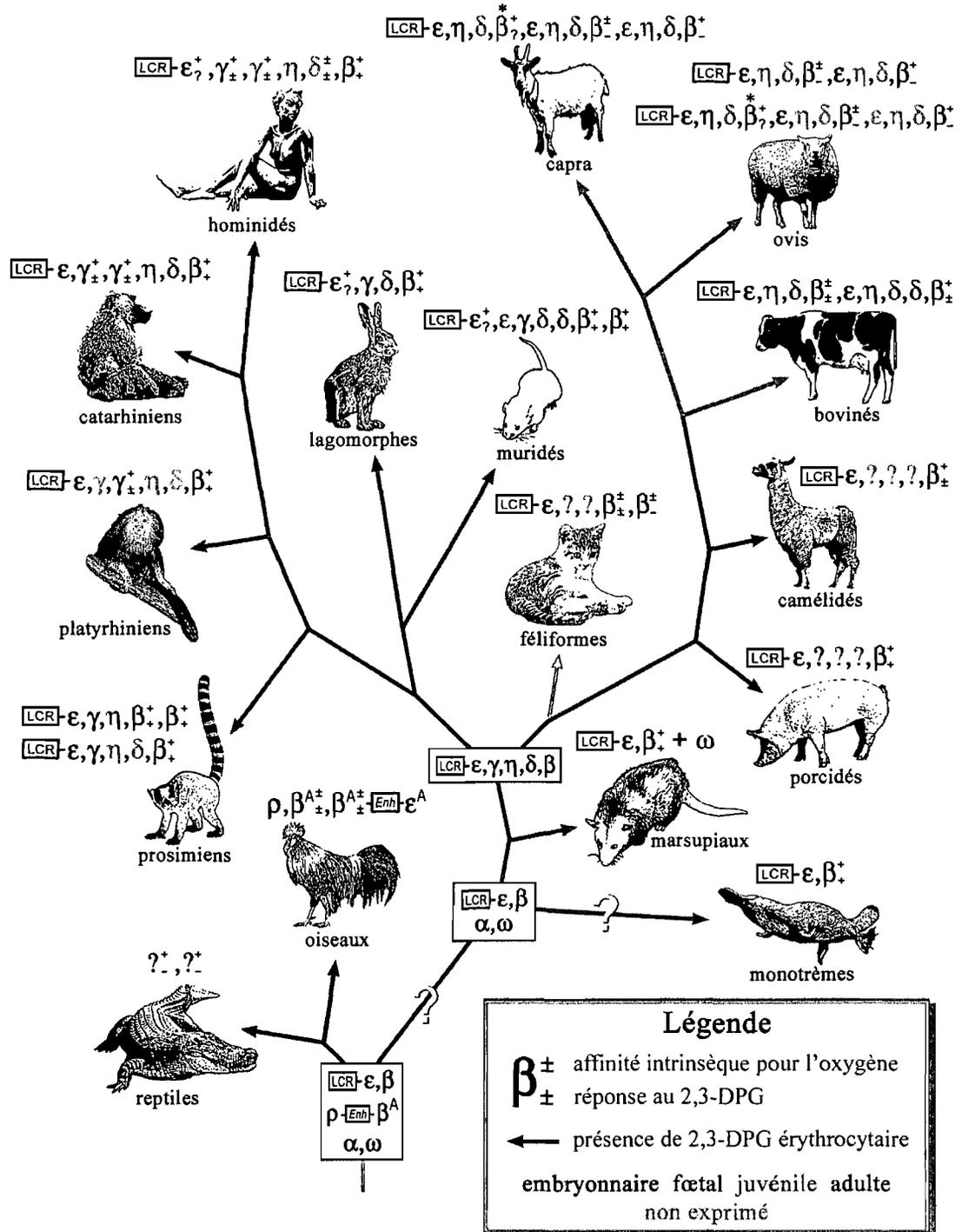


Figure 7: Phylogénie, génétique et physiologie du locus de la β -globine. ^{2;28-31;37;66;122;123;163-201}

1.4.1 La souris

Après le locus humain de la β -globine. Celui de la souris est le locus qui a été le plus étudié par la communauté scientifique. Plusieurs éléments découverts chez la souris ont pu être transposés à l'humain et ce, malgré les différences qui existe entre les deux locus.²⁰²⁻²⁰⁵ Le locus de la β -globine de la souris représente le prototype mammalien puisque sa régulation est semblable à celle observée chez la plupart des autres espèces hormis les singes et les ruminants. Toutefois, ce ne fut pas pour cette raison que les chercheurs se sont tant intéressés à l'hémoglobine de la souris. Ce fut en fait, la possibilité de produire des souris transgéniques comportant les gènes humains de la β -globine qui propulsa ce petit mammifère au firmament des espèces les plus étudiés.^{206,207}

Le locus

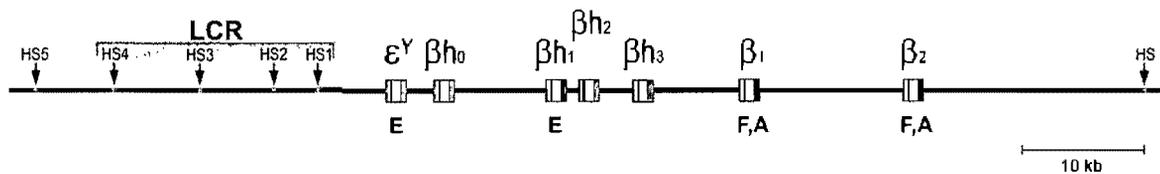


Figure 8: Le locus de la β -globine de la souris *Mus musculus*. La lettre sous chaque gène indique la période du développement où le gène est exprimé (E: embryonnaire, F: fœtal, A: adulte). La couleur des gènes représente l'origine évolutive: (○): ϵ , (●): γ , (◐): δ , (◑): β , (◒): indique un pseudogène.

Le locus de la β -globine de la souris, qui est situé sur le chromosome 7, comporte 4 gènes fonctionnels ainsi que 3 pseudogènes (Figure 8). Les sept gènes sont répartis sur une distance d'environ 50 kb ce qui est semblable au locus humain, alors que le LCR s'étend sur une distance presque deux fois plus grande, soit environ 25 kb. Chez les rongeurs, le pseudogène η n'est pas présent alors que les gènes dupliqués δ (β_{h2} et β_{h3}) sont devenus pseudogènes.²⁰⁸⁻²¹⁰ Le pseudogène β_{h0} quant à lui est issu d'une rare duplication du gène ϵ^Y .² Les gènes β_1 et β_2 sont tous deux orthologues du gène humain β et sont issus de la

duplication du gène ancestral β .²⁰² Chez la souris *Mus musculus* on retrouve deux haplotypes principaux du locus de la β -globine: l'haplotype *diffus* produisant deux globines adultes (β^{majeur} étant codée par le gène β_1 et β^{mineur} encodée par le gène β_2) alors qu'avec l'haplotype *single*, les gènes β_1 et β_2 encodent pour une même protéine, la β^{single} .²¹¹⁻²¹³

La commutation de l'hémoglobine

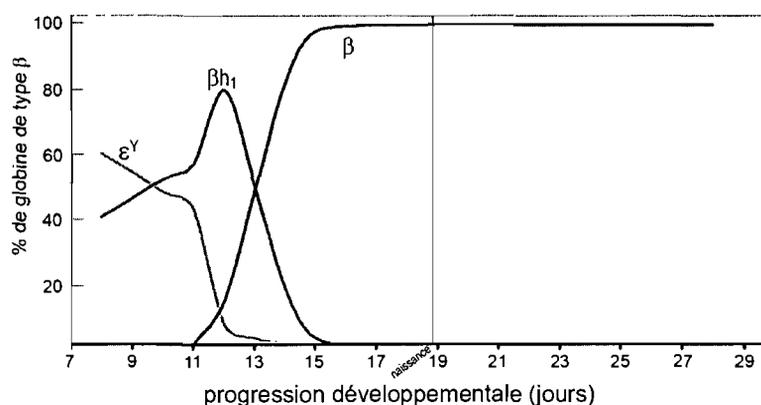


Figure 9: Commutation de l'hémoglobine chez la souris. Adapté de Peterson *et al.*²⁰⁷

Chez la souris il n'y a qu'une seule commutation de l'hémoglobine et elle a lieu entre le stade embryonnaire et le stade fœtal alors que les gènes ϵ^Y et βh_1 sont remplacés par les gènes β_1 et β_2 .²¹⁴ L'ontogénie des gènes embryonnaires n'est toutefois pas complètement synchronisée puisque ϵ^Y est exprimé légèrement plus tôt que βh_1 qui, pour cette raison, est souvent identifié comme gène semi-embryonnaire (Figure 9). Chez la souris, au stade fœtal, les niveaux de 2,3-DPG sont significativement plus bas que les niveaux observés chez l'adulte, ce qui permet d'assurer une affinité pour l'oxygène du sang des fœtus plus grande que celle du sang de la mère (Figure 7, p.28).

1.4.2 Les prosimiens

Les prosimiens regroupent les petits primates les plus distants phylogénétiquement de l'homme et donc également, les plus primitifs. Les prosimiens les plus connus sont les galagos, les lémures et les tarsiers et ce sont également ceux dont le locus de la β -globine a été étudié. D'ailleurs, les prosimiens sont particulièrement intéressants puisque bien qu'appartenant à l'ordre des primates, leur locus de la β -globine ressemble et se comporte comme celui des autres mammifères plutôt que celui des primates supérieurs (Figure 7, P.28). L'étude des différences entre le locus prosimien et le locus simien, évolutivement proches, pourrait s'avérer d'un grand intérêt puisqu'il est proposé que l'expression strictement fœtale des γ -globines chez l'humain, doit ultimement être due à ces différences.^{165;166;215;216}

Le locus de la β -globine

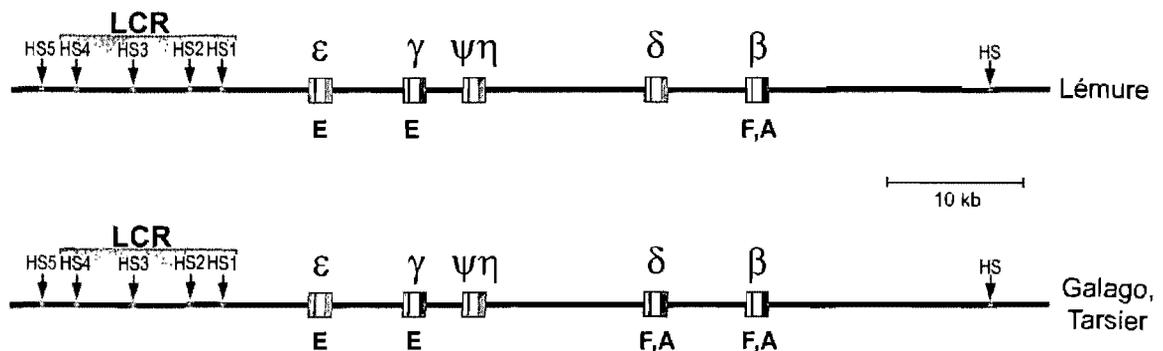


Figure 10: Le locus de la β -globine chez les prosimiens. Les deux locus typiques retrouvés chez les simiens sont présentés, celui du lémure et celui retrouvé chez le galago et le tarsier. La lettre sous chaque gène indique la période du développement où le gène est exprimé (E: embryonnaire, F: fœtal, A: adulte). La couleur des gènes représente l'origine évolutive: (◊): ϵ , (●): γ , (●): η , (◊): δ , (●): β , (◊): indique un pseudogène.

Chez la plupart des singes prosimiens, le locus de la β -globine comporte 4 gènes fonctionnels et un pseudogène répartis sur une distance de ~ 30 kb. Le lémure représente

une exception puisque chez ce dernier, le gène δ a subi une conversion en 5' par le pseudogène $\psi\eta$ et par conséquent, s'est retrouvé inactivé (Figure 10).²¹⁷ Chez les autres prosimiens dont δ est fonctionnel, il est intéressant de noter que celui-ci a également subi une conversion en 5', cette fois par le gène β .¹⁶⁶ Chez tous ces primates primitifs, γ n'est pas dupliqué et la distance entre les gènes ϵ et γ est fortement réduite comparativement à celle observée chez les primates supérieurs.¹⁶⁵

La commutation de l'hémoglobine

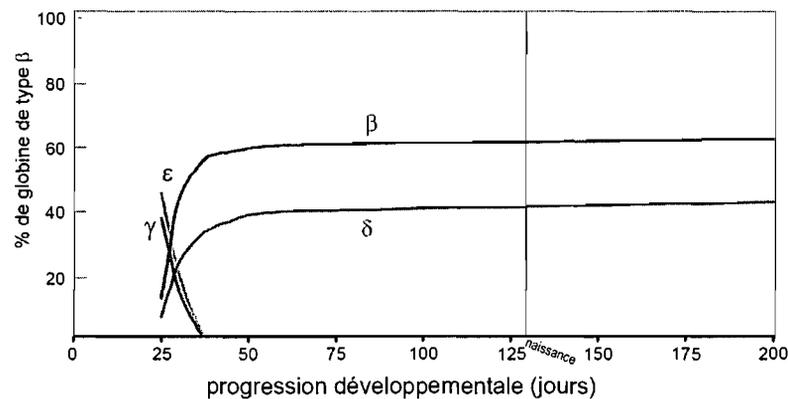


Figure 11: Commutation de l'hémoglobine chez les prosimiens.¹⁶⁵

Chez les prosimiens, ϵ et γ sont exprimés de façon concertée au stade embryonnaire uniquement et sont remplacés lors de la commutation de l'hémoglobine par les globines δ et β . Il est à noter qu'à cause de la conversion de la région 5' de δ par β chez ces primates, δ est exprimé à des niveaux bien supérieurs à ceux observés chez les autres mammifères, représentant près de 40% de l'hémoglobine fœtale/adulte (Figure 11).^{165;166} Comme chez la majorité des mammifères, les niveaux de 2,3-DPG au stade fœtal est plus bas que chez les adultes.

1.4.3 Les simiens I: les singes du Nouveau Monde (*platyrrhiniens*)

Les platyrrhiniens aussi nommés singes du Nouveau Monde puisqu'on ne les retrouve qu'aux Amériques, comprennent plusieurs espèces dont les plus connues sont les ouistitis, les capucins, les singes hurleurs et les atèles ou singes-araignées et se caractérisent par leur queue préhensile et leurs narines écartées. Du point de vue évolutif, ces petits primates représentent une étape de plus vers les grands singes. En effet, les platyrrhiniens sont les singes les plus primitifs à avoir une hémoglobine exclusivement fœtale comme celle retrouvée chez l'humain. Pendant longtemps, les chercheurs ont cru que ces singes n'avaient qu'un seul gène γ ce qui suggérait donc que le recrutement fœtal des γ -globine précédait leur duplication.^{218;219} Toutefois, la vérité s'avéra plus complexe et le débat demeure ouvert sur ce qui précède soit de la duplication du gène γ ou du recrutement fœtal de son expression.

Le locus de la β -globine

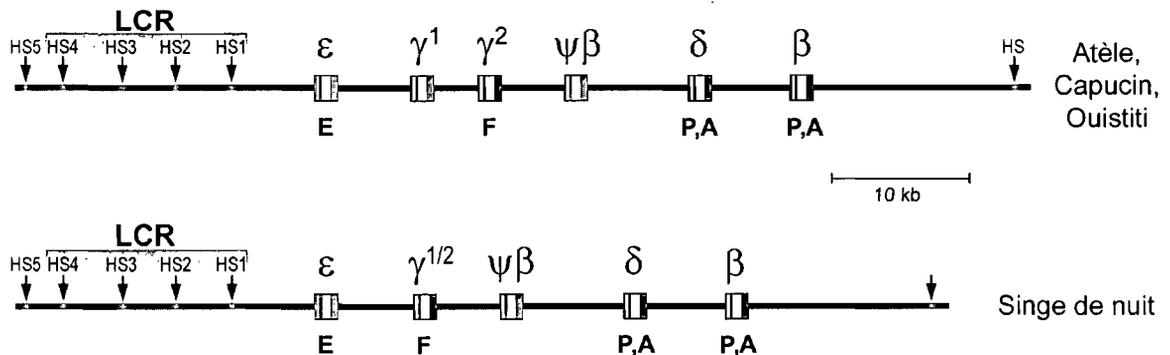


Figure 12: Le locus de la β -globine chez les singes du Nouveau Monde. Le locus retrouvé chez la grande majorité des singes du Nouveau Monde est présenté ainsi que le locus du singe de nuit *A. azaræ* qui se caractérise par la présence d'un seul gène γ issu de la fusion des gènes γ^1 et γ^2 . La lettre sous chaque gène indique la période du développement où le gène est exprimé (E: embryonnaire, F: fœtal, P: prénatal, A: adulte). La couleur des gènes représente l'origine évolutive: (○): ϵ , (●): γ , (◐): η , (◑): δ , (◒): β , (◓): indique un pseudogène.

Le locus de la β -globine des singes du Nouveau Monde est structurellement très semblable à celui des humains à l'exception du pseudogène γ^1 , orthologue de $G\gamma$ (Figure 12). Les 4 gènes fonctionnels et les deux pseudogènes sont répartis sur une distance d'environ 37 kb dû à une réduction de la distance entre ϵ et γ par rapport à celle observée chez les catarhiniens.¹⁶⁸ Le singe de nuit (*Aotus azarae*) représente une exception intéressante puisque les gènes γ ont fusionnés et forment ainsi un seul gène fonctionnel (Figure 12).¹⁶⁸

La commutation de l'hémoglobine

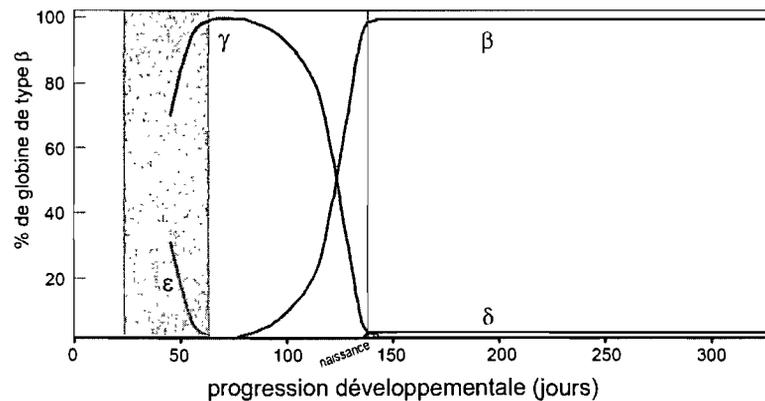


Figure 13: Commutation de l'hémoglobine chez les singes du Nouveau Monde.¹⁶⁸

Les platyrhiniens sont les primates les plus primitifs à présenter deux commutations de l'hémoglobine. Si on connaît peu de chose de la commutation de l'hémoglobine embryonnaire à fœtale qui n'a pas été étudiée, on sait que la commutation fœtale à adulte se produit quelques temps avant la naissance (Figure 13) contrairement aux singes de l'Ancien Monde et aux humains chez qui cette commutation survient après la naissance (Figure 4. p.15 et Figure 15, p.36).^{220;221} Ces singes représentent ainsi un état intermédiaire entre les prosimiens et les grands singes. Chez *A. azarae*, qui ne possède qu'un seul gène γ , il est intéressant de noter qu'au stade fœtal, γ ne représente qu'un très faible pourcentage de

l'hémoglobine, suggérant une expression plus embryonnaire, bien que ceci resterait à être confirmé.¹⁶⁸

1.4.4 Les simiens II: les singes de l'Ancien Monde (*catarhiniens*)

Les catarhiniens sont des primates plus évolués que l'on retrouve en Afrique et en Asie. Ils se caractérisent par leur nez aux narines ouvertes vers le bas et leur queue non préhensile ou absente. Les espèces les plus connues de catarhiniens comprennent les babouins, les macaques et les gibbons. Bien que les chimpanzés et les gorilles fassent aussi parti des catarhiniens, ces derniers ne sont généralement pas appelés singes de l'Ancien Monde, mais plutôt hominidés ou grands singes et sont regroupés avec l'homme.²²²⁻²²⁴ Vu la grande proximité phylogénétique entre les singes de l'Ancien Monde et les hommes, les catarhiniens ont été utilisés comme modèle d'étude pour la régulation de l'hémoglobine au cours du développement.^{169,225}

Le locus de la β -globine

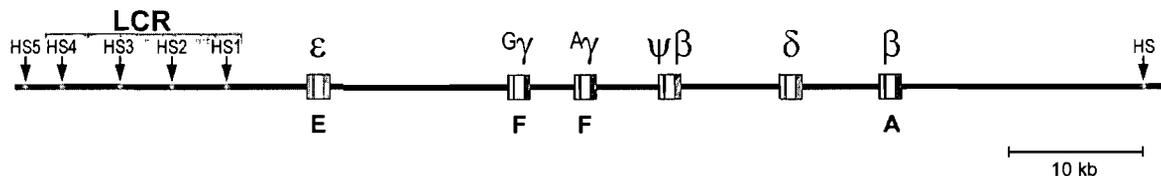


Figure 14: Le locus de la β -globine chez les singes de l'Ancien Monde. La lettre sous chaque gène indique la période du développement où le gène est exprimé (E: embryonnaire, F: fœtal, A: adulte). La couleur des gènes représente l'origine évolutive: (○): ϵ , (●): γ , (●): η , (○): δ , (●): β , (○): indique un pseudogène.

Le locus de la β -globine des catarhiniens n'a subi que peu de changements par rapport au locus humain. La seule différence notable est l'inactivation du gène δ devenu un pseudogène dans cette lignée (Figure 14) ce qui pose la question à savoir si chez les grands

singes, l'expression de δ est due à une réactivation de ce gène, ou si l'inactivation de δ chez les singes de l'Ancien Monde s'est produite après la séparation de ces deux groupes.^{167;169}

La commutation de l'hémoglobine

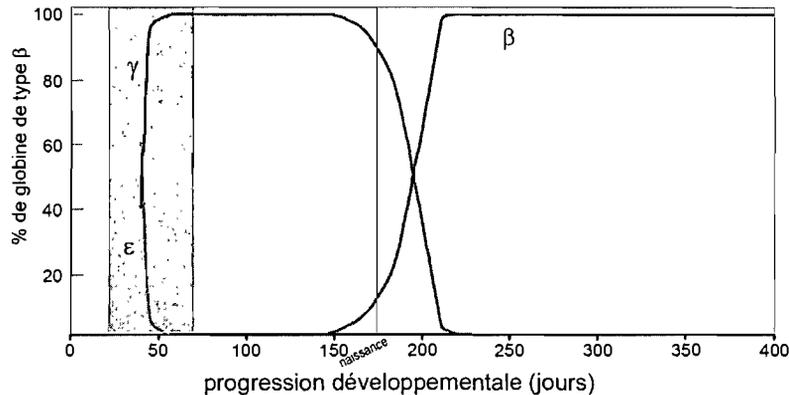


Figure 15: Commutation de l'hémoglobine chez les singes de l'Ancien Monde. Adapté de Kitchen *et al.*²⁸

La commutation de l'hémoglobine chez ces primates suit un profil qui ressemble assez à celui observé chez l'homme avec la différence que la commutation fœtale à adulte se produit sur une période plus courte que celle observée chez l'humain (Figure 15 et Figure 4, P.15). Également, si chez le fœtus humain, β représente toujours environ 8-15% de l'hémoglobine totale (Figure 4), elle n'est pas détectable chez les fœtus de ces singes.²⁸

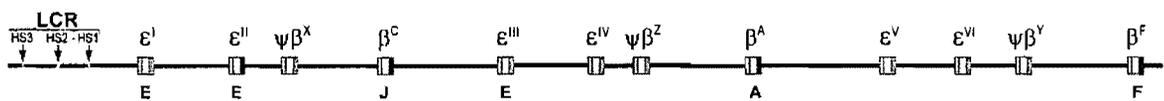
1.4.5 Les ruminants

Les ruminants représentent un groupe intéressant et unique. En effet, ces animaux qui sont phylogénétiquement très éloignés de l'homme ont développé indépendamment un système d'hémoglobines spécialisées. Mis à part les primates supérieurs, les ruminants sont les seuls animaux à présenter une hémoglobine strictement fœtale.²²⁶ Comme chez les primates, cette spécialisation permis un temps de gestation plus long permettant ainsi la naissance "à terme" des nouveau-nés, améliorant ainsi leur chance de survie face aux

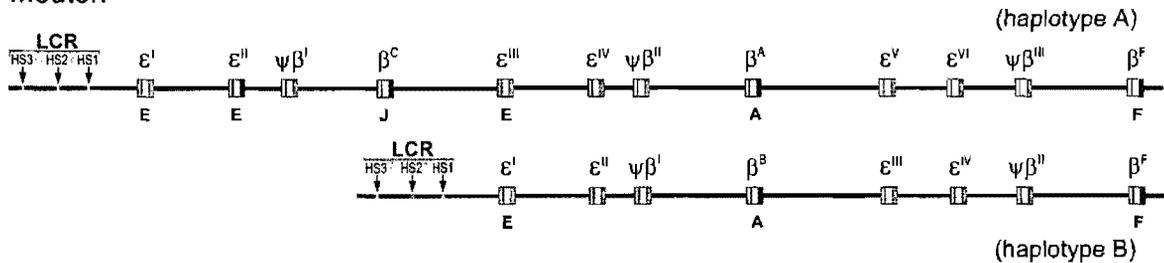
prédateurs. Deux caractéristiques uniques aux ruminants sont la présence chez certaines espèces d'une hémoglobine dite "juvénile" pouvant être réactivée lors de stress hématopoïétiques,^{182;227} et l'absence de 2,3-DPG érythrocytaire à tous les stades du développement, à l'exception d'un pic observable à la naissance (Figure 7, p.28).²²⁸ De fait, chez ces animaux, toutes les hémoglobines ne présentent qu'une très faible, voir même aucune réactivité au 2,3-DPG.²²⁹ Les trois espèces de ruminants les plus étudiées sont la chèvre, le mouton et la vache. Fait à noter, le gène γ est absent chez tous les ruminants.

Le locus de la β -globine

Chèvre



Mouton



Vache

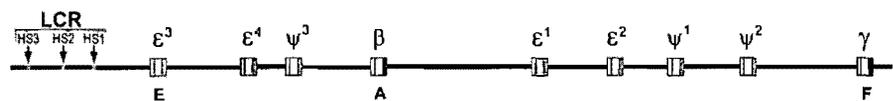


Figure 16: Le locus de la β -globine chez ruminants. Les locus de trois espèces de ruminant sont présentés: celui de la chèvre, les deux haplotypes du mouton et celui de la vache. La lettre sous chaque gène indique la période du développement où le gène est exprimé (E: embryonnaire, F: fœtal, J: juvénile, A: adulte). La couleur des gènes représente l'origine évolutive: (◐): ϵ , (◑): η , (◒): δ , (◓): β , (◔): indique un pseudogène.

La chèvre

Le locus de la β -globine de la chèvre est issu d'une triplification des quatre gènes ϵ - η - δ - β qui a eu lieu après la perte du gène γ . Le locus qui comprend 6 gènes fonctionnels et 6 pseudogènes s'étend sur une distance de ~125 kb (Figure 16).²³⁰ Précédant le locus, le LCR se distingue par la présence de seulement 3 sites d'hypersensibilité puisque les sites HS4 et HS5 n'ont pu être identifiés chez ces animaux.^{231;232} Chez la chèvre, les gènes embryonnaires sont d'origine ϵ et η ce qui est assez unique,²⁰⁸ alors que les gènes fœtaux, juvéniles et adultes sont tous paralogues à β .²³³ Les gènes δ sont quant à eux tous réduits à l'état de pseudogènes.^{233;234}

Le mouton

Chez le mouton, on retrouve deux haplotypes pour le locus de la β -globine. L'haplotype A est identique au locus de la chèvre alors que l'haplotype B est un locus délesté du premier triplicata.²³⁵ Ce locus possède 3 gènes fonctionnels et 5 pseudogènes répartis sur 75 kb et donc, ne possède pas de β -globine juvénile (Figure 16).^{179;185;236;237}

La vache

Le locus de la vache est similaire au locus de l'haplotype β du mouton à l'exception d'une duplication du pseudogène δ produisant un sixième pseudogène (Figure 16).¹⁷⁶ L'hémoglobine juvénile est inexistante chez la vache.

La commutation de l'hémoglobine

Contrairement à tous les autres groupes d'animaux, chez les ruminants l'expression des différents gènes du locus au cours du développement ne suit pas l'ordre de leur positionnement dans le locus. En effet, chez ces animaux, le gène fœtal est situé en aval du gène adulte. De plus, dans le cas de la chèvre et du mouton, un gène exprimé strictement au

stade embryonnaire est situé au centre du locus, précédé par les deux autres gènes embryonnaires et par le gène juvénile (Figure 16).

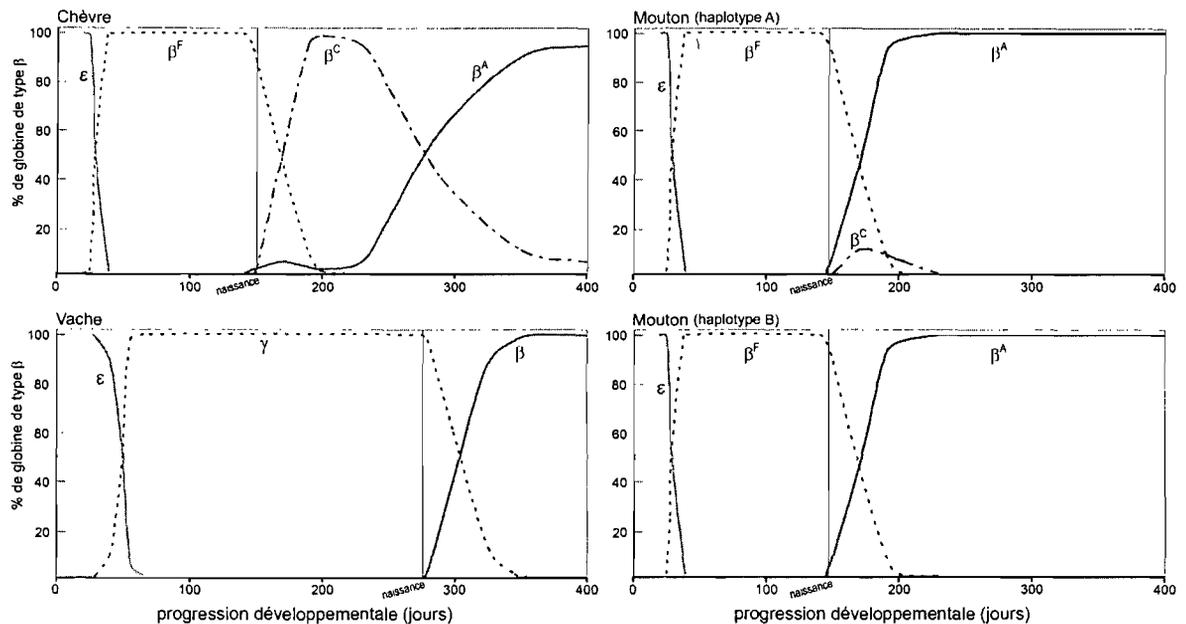


Figure 17: Commutation de l'hémoglobine chez les ruminants. Adapté de Hammerberg *et al.*¹⁸³ et de Huisman *et al.*¹⁸²

La chèvre

La chèvre est unique en ce que son hémoglobine subit trois commutations complètes lors du développement (Figure 17). La commutation de l'hémoglobine embryonnaire à fœtale a lieu très tôt comme chez les autres animaux. Par la suite, comme chez l'humain, la chèvre produit lors du stade fœtal une hémoglobine strictement exprimée à ce stade et qui disparaît à la naissance. Toutefois, cette hémoglobine n'est pas remplacée par l'hémoglobine adulte mais par une hémoglobine transitoire appelée hémoglobine juvénile. Cette hémoglobine est ensuite graduellement remplacée par l'hémoglobine adulte

qui devient l'hémoglobine majoritaire environ quatre mois après la naissance.¹⁸² L'hémoglobine juvénile se distingue également par sa capacité à être activée en réponse à un stress hématopoïétique et ce, au détriment des gènes fœtaux et adultes.²²⁷ De fait, cette hémoglobine est la seule hémoglobine connue dont l'expression peut être activée par l'érythropoïétine. La fonction de cette hémoglobine à faible affinité pour l'oxygène n'est toutefois pas comprise tout comme la signification de l'expression transitoire du 2,3-DPG qui survient après la naissance, puisque aucune hémoglobine ne répond à cet effecteur chez la chèvre.^{66,238} Il est toutefois proposé dans ce cas, que l'élévation soudaine du 2,3-DPG pourrait provoquer un changement de pH qui affecterait l'affinité de l'hémoglobine pour l'oxygène.²⁹

Le mouton

La commutation chez le mouton d'haplotype A se fait également en trois temps comme chez la chèvre mais dans ce cas, l'expression de l'hémoglobine juvénile ne dépasse jamais 10% de l'hémoglobine totale et de ce fait, l'hémoglobine adulte succède presque directement l'hémoglobine fœtale (Figure 17).¹⁸³ Par contre, en stress hématopoïétique, l'hémoglobine juvénile est fortement activée au détriment des hémoglobines fœtales et adultes ce qui laisse suggérer que la fonction de cette hémoglobine chez le mouton en serait une de réponse à des situations où la demande en oxygène est déterminante.²²⁷ Les moutons de l'haplotype B ne présentent que deux commutations dont la commutation fœtale à adulte qui survient après la naissance et ne répondent pas au stress hématopoïétique en produisant une hémoglobine spécialisée (Figure 17).

La vache

Comme les moutons d'haplotype B, la vache ne présente que deux commutations de l'hémoglobine, l'une très tôt et l'autre après la naissance (Figure 17).^{239,240}

1.5 Régulation

Les gènes de globine de type β sont exprimés de façon séquentielle au cours du développement, dans l'ordre de leur agencement dans le locus. Comme je l'ai mentionné précédemment, chez l'humain il y a deux commutations de l'hémoglobine, impliquant de ce fait, la présence de deux périodes de transition au cours desquelles l'expression d'une ou de deux globines est diminuée et que celle de deux autres globines est augmentée. Malgré cela, le niveau global des chaînes de type β dans la cellule demeure toujours constant et proportionnel aux chaînes de type α .²⁸ Cette caractéristique suppose une régulation très serrée et coordonnée, aussi bien des différents gènes du locus de la β -globine que des gènes du locus de la α -globine qui sont situés sur un chromosome différent. Ce qui coordonne l'expression des gènes β avec les gènes α demeure toujours inconnu mais il y a quelques années, une étude a démontré que cette corégulation ne passait pas par la communication croisée entre les loci via un hypothétique "*trans*-balanceur".²⁴¹ Cependant, une étude récente a permis de suggérer un rôle de la chaperonne AHSP et un mécanisme de régulation post-transcriptionnelle.²⁴²

Toutefois, ce qui coordonne l'expression des différents gènes du locus de la β -globine entre eux a été largement étudié et de nombreux éléments impliqués dans cette régulation ont été identifiés à la grandeur du locus. Certes il a été démontré que les promoteurs des gènes du locus ainsi que les différents sites HS du LCR étaient importants pour la régulation des gènes, mais plusieurs éléments extragéniques ont également été identifiés.²⁴³ De plus, l'importance d'une régulation épigénétique a clairement été établie, démontrant de ce fait que la régulation des gènes du locus ne se résume pas qu'à des ratios différentiels de facteurs de transcription mais fait également intervenir une fonction de l'organisation supérieure de la chromatine.²⁴⁴ Ces différents modes de régulation font du locus de la β -globine un système très complexe dont l'ordre des événements assurant une régulation adéquate au cours du développement ne demeure toujours qu'imparfaitement élucidé.²⁴⁵⁻²⁴⁸

1.5.1 La commutation embryonnaire à fœtale

Des deux commutations, la commutation de l'hémoglobine embryonnaire à fœtale est celle qui est la plus mal comprise et ce, pour deux raisons principales. La première est que chez l'humain, cette commutation intervient très tôt lors du développement et demeure donc très difficile à étudier dans son contexte naturel même si maintenant, certains modèles cellulaires existent permettant de récapituler cette commutation.^{115;249} La deuxième étant que chez les souris transgéniques arborant le locus humain, la commutation ϵ à γ ne se fait pas comme chez l'humain puisque ces gènes sont exprimés de façon concomitante au stade embryonnaire et que γ est toujours exprimé plus fortement que ϵ .¹¹⁰⁻¹¹² Si on sait maintenant que l'inactivation de ϵ dans les cellules érythroïdes définitives se fait de façon autonome et donc que ce gène n'a pas besoin de l'activation de γ ou de β pour être inactivé,^{250;251} ce qui contrôle l'activation de γ chez l'humain demeure toutefois toujours inconnu, bien que l'on sait que cette activation dépend de l'environnement "fœtal" cellulaire.^{252;253} Il est intéressant de remarquer que les gènes ϵ et γ ne sont pas régulés par le LCR de la même façon puisque l'activation de ϵ est dépendante de sa distance avec le LCR et de la transcription active de la région précédant son promoteur alors que l'activation des gènes γ en est indépendante.²⁵⁴

1.5.2 La commutation fœtale à adulte

La seconde commutation de l'hémoglobine à survenir chez l'humain se caractérise par la disparition des chaînes γ et leur remplacement par la chaîne β ainsi que de façon plus limitée, par la chaîne δ . Parce qu'elle survient après la naissance chez l'homme et qu'elle est proprement régulée chez la souris transgénique, la commutation de l'hémoglobine fœtale à adulte a été beaucoup plus étudiée et en est de fait, mieux comprise. Toutefois, les mécanismes qui la régissent ne demeurent toujours que partiellement élucidés. Différentes études ont permis de démontrer que l'inactivation des gènes γ était obtenue par un processus essentiellement autonome, semblable à celui régissant l'inactivation de ϵ . Ce processus

implique des séquences situées jusqu'à 730 kb en amont du site d'initiation de la transcription puisque des promoteurs plus courts échouent à être inactivés au stade adulte.²⁵⁵⁻²⁵⁹ L'activation de β quant à elle est surtout contrôlée par un mécanisme de compétition entre les promoteurs pour le LCR et est donc dépendante de l'inactivation de γ .²⁶⁰⁻²⁶² Il est à noter cependant que l'affinité pour le LCR du promoteur de β au stade embryonnaire est plus faible que chez l'adulte mais demeure suffisante pour une activation embryonnaire de β en absence de compétition ce qui démontre que, bien que l'activation de β dépende de l'inactivation de γ , une certaine potentialisation de β a lieu de façon autonome.²⁶³⁻²⁶⁵

1.5.3 Éléments *cis*-régulateurs

De nombreux éléments de séquences ayant un rôle dans la régulation des différents gènes de globine ont été identifiés dans le locus et les facteurs ou complexes qui lient ces éléments ont également pour la plus grande majorité été identifiés. Ces éléments sont surtout concentrés au niveau des séquences promotrices des différents gènes du locus et des sites HS du LCR mais ne sont pas exclusifs à ces régions puisqu'on en retrouve presque partout dans le locus (*e.g.* 367 sites putatifs de liaison du facteur érythroïde GATA-1 ont été identifiés dans le locus, mais seulement quelques uns ont été démontrés comme étant fonctionnels lorsque étudiés dans les cellules érythroleucémiques K562²⁶⁶⁻²⁶⁹). Cependant, la fonction exacte de la plupart des éléments de liaison à des facteurs qui ont été identifiés dans le locus demeure toujours incomprise à part certains dont le rôle a été démontré de façon non équivoque, comme la boîte proximale CACCC du promoteur β qui est liée par le facteur EKLF et qui est essentielle à l'activation du gène.^{261;270}

Dans cette section seront présentés les nombreux éléments situés dans les différentes régions du locus et identifiés comme étant importants dans la régulation des gènes, ainsi que leur rôle probable.

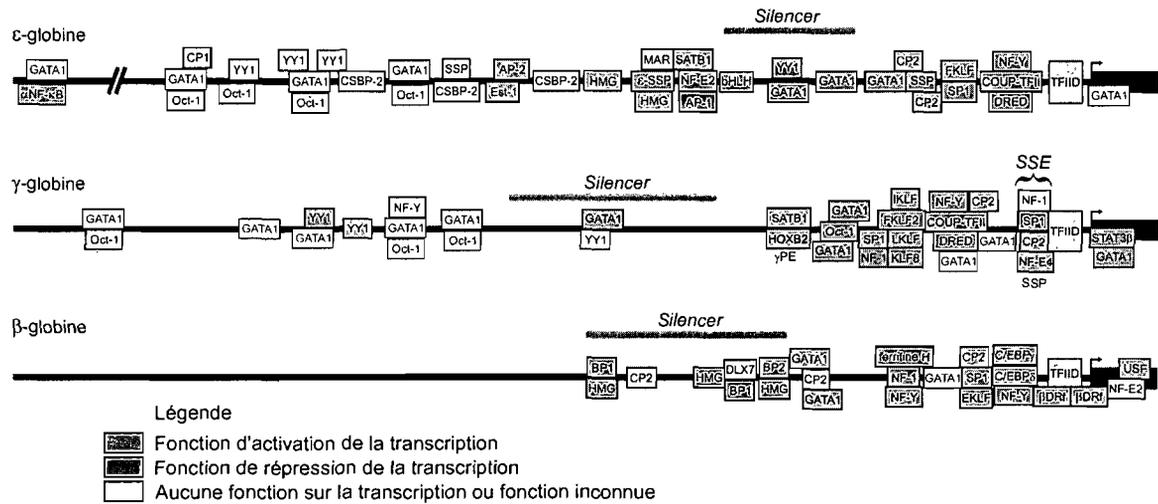


Figure 18: Promoteurs des gènes ϵ -, γ - et β -globine.

Le promoteur de ϵ -globine

Puisque le gène ϵ démontre une grande autonomie dans sa régulation, caractérisée par une inactivation complète lors de l'érythropoïèse définitive malgré la proximité de ce gène au LCR, des recherches intensives d'éléments régulateurs dans la région environnante immédiate du gène ont été entreprises et ont permis d'identifier plusieurs de ces éléments.

Globalement, le promoteur de ϵ peut être découpé en quatre régions fonctionnelles (Figure 18). La partie proximale du promoteur, de -165 au site d'initiation de la transcription, forme un domaine activateur de la transcription. En effet, à l'exception de deux sites en tandem situés dans la région la plus proximale du promoteur et qui sont liés par le facteur DRED qui possède une fonction répressive sur la transcription de ϵ , ainsi que par le facteur COUP-TFII qui lui, possède une fonction bivalente répressive et activatrice selon le contexte,²⁷¹⁻²⁷³ tous les facteurs qui lient cette région sont importants dans l'activation de ce gène. Parmi ces facteurs il y a GATA-1,²⁷⁴ SSP,^{275;276} CP2,²⁷⁷ FKLf,²⁷⁸ Sp1,^{274;275} et NF-Y^{279;280} et tous ont été démontrés comme étant importants individuellement pour l'activité transcriptionnelle de ϵ au stade embryonnaire. Toutefois, la fonction des éléments

régulateurs liant les répresseurs DRED et COUP-TFII est probablement de réprimer l'expression de ϵ dans l'hématopoïèse définitive puisque une expression forcée de DRED provoque une répression prématurée de la globine embryonnaire.²⁷³

Cependant, la fonction principale de répression autonome de ϵ est gérée par la région qui précède la région activatrice et qui couvre une distance de 140 pb, de -392 à -251.^{281;282} Cette région *Silencer* recrute plusieurs facteurs dont le rôle dans la répression de ϵ a été démontré. Cette région possède ainsi deux sites de liaison à GATA-1,^{274;283} ainsi que des sites de liaison pour les facteurs YY1,²⁸³ NF-E2 et AP-1^{284;285} ainsi qu'un site de liaison à un facteur de la famille des bHLH.²⁸⁴

Précédant la région de contrôle négative et couvrant une distance d'environ 320 pb de -770 à -453 du site d'initiation de la transcription de ϵ , se trouve une deuxième région activatrice importante comprenant des sites de liaisons au facteurs HMG,^{286;287} ϵ -SSP,^{284;288} SATB1,²⁸⁹ AP-2 et Est-1.²⁸⁴

De façon intéressante, un élément de liaison à la matrice nucléaire (MAR) est situé à la frontière entre la région activatrice et la région répressive et l'intégrité de cet élément est important pour préserver la fonction répressive de la région négative.^{287;290} Une autre région à fonction répressive a été identifiée loin en amont des autres régions importantes du promoteur de ϵ à environ 2 kb du site d'initiation de la transcription. En effet, il a été démontré qu'un site de liaison à NF- κ B était important pour la répression de ϵ au stade fœtal.^{291;292} Cependant, entre ce dernier élément répresseur et la région activatrice située entre -770 et -453, se retrouve une série d'éléments de liaison à des facteurs de transcription dont le rôle dans la régulation du gène ϵ demeure inconnu.²⁷⁶

Le promoteur de γ -globine

Une vue générale du promoteur des gènes $A\gamma$ et $G\gamma$ dévoile une organisation des différentes régions beaucoup moins définie que les régions du promoteur de ϵ (Figure 18). En effet, le promoteur de γ montre un plus grand dynamisme dans sa façon d'être régulé

que ϵ et les différents éléments *cis*-régulateurs peuvent le plus souvent avoir une fonction répressive ou activatrice selon le contexte cellulaire et les facteurs qui s'y lient. Toutefois, une certaine spécialisation de différentes sections peut être remarquée.

La première région importante de régulation de γ est un élément situé à -50 du site d'initiation de la transcription, donc très près de la boîte TATA et appelé élément sélecteur de stade ou *SSE* (*Stage Selector Element*). Cette courte région a pour fonction d'activer sélectivement γ au stade fœtal précoce et embryonnaire et de lui donner un avantage sélectif dans sa compétition avec β pour le LCR.^{293;294} Plusieurs facteurs peuvent lier le *SSE* dont le complexe SSP (*stage selector protein*) composé de deux sous-unités, NF-E4 et CP2 qui possèdent une fonction activatrice de γ en permettant le recrutement du LCR.^{277;295;296} Une troisième sous-unité du complexe SSP a également été identifiée, la protéine NF-1 mais dont le rôle dans le contexte du complexe *SSE* demeure toujours inconnu.²⁹⁷ Si le *SSE* possède généralement une activité positive sur la transcription de γ , lorsque les cytosines de cet élément sont méthylées, l'affinité de ce dernier pour le facteur de transcription ubiquitaire Sp1 est fortement augmentée. Sp1 qui le plus souvent agit comme activateur transcriptionnel, vient dans ce cas en compétition avec le complexe SSP et le déplace ce qui résulte en une répression du gène γ qui perd alors son avantage sélectif contre β .^{298;299}

À une soixantaine de paires de bases en amont du *SSE*, se trouve une autre petite région comportant des boîtes CCAAT dupliquées et pouvant également être liée par plusieurs facteurs antagonistes. Cette région partage des caractéristiques retrouvées dans la séquence proximale de ϵ et est liée par les mêmes facteurs répresseurs DRED et COUP-TFII^{300;301} dont le deuxième, contrairement à ce qui est observé avec ϵ , ne possède qu'une fonction répressive sur γ .²⁷² Cet élément peut aussi recruter le facteur activateur NF-Y qui entre alors en compétition directe avec COUP-TFII et favorise l'expression de γ en ouvrant la chromatine au niveau du promoteur.^{268;300;302} Cette région possède également des sites de liaison pour les facteurs CP2, un activateur, et GATA-1 dont la fonction dans ce contexte est inconnue.^{277;303} Il est à noter que la mutation HPFH grecque γ : -117 (G→A)

(Tableau I, page 25) affecte le site de liaison de DRED et agit donc en éliminant l'effet répresseur de ce facteur.³⁰¹

Le principal élément activateur des gènes γ se trouve à environ 150 pb en amont du site d'initiation de la transcription et comprend une boîte CACCC pouvant être liée par les activateurs spécifiques à γ , FKLf-2, IKLF et LKLF.^{304;305} Cette région possède également un site de liaison au facteur Sp1 qui dans ce contexte possède une activité activatrice de la transcription de γ .^{306;307} Deux facteurs ayant des activités antagonistes aux facteurs précédents peuvent cependant lier cette région. Ce sont les facteurs NF-1 et KLF8 mais leur importance physiologique demeure encore inconnue.^{305;306}

En amont de cet élément activateur, se trouve une région présentant une fonction répressive assez faible, s'étalant sur une distance d'environ 25 pb de -200 à -175 et comportant quelques sites de liaison à des facteurs de transcriptions. Deux sites de liaisons à GATA-1 et un site de liaison à Oct-1 démontrent une activité de répression²⁶² et plusieurs mutations HPFH localisées dans cette région affectent la capacité de liaison de ces facteurs.³⁰⁸ La région comporte également un site de liaison au facteur CP2 mais dont la fonction est inconnue.²⁷⁷

La région située entre -300 et -200 comporte un important élément activateur pouvant lier le complexe γ PE constitué des deux sous-unités SATB1 et HOXB2 et étant important pour obtenir une activation maximale du gène fœtal.^{309;310} Toutefois la région qui la précède, située entre -382 et -730 a depuis longtemps été associée à la fonction de *Silencer* principal des gènes γ et donc, responsable de l'inactivation autonome de ces gènes chez l'adulte.²⁵⁷ Toutefois, cette région est pauvre en sites de liaison aux facteurs de transcription, à l'exception d'un site de liaison à YY1 et d'un site de liaison à GATA-1 qui lui, est directement impliqué dans cette répression.³¹¹⁻³¹² Finalement, en amont de la région *Silencer* se retrouvent plusieurs sites de liaison aux facteurs Gata-1, Oct-1 et NF-Y dont les fonctions ne sont pas connues ainsi que deux sites de liaison à YY1.³¹³ Le premier, situé à environ 1 kb en amont du site d'initiation de la transcription possède une fonction

activatrice de la transcription du gène γ au stade fœtal³¹⁴ alors que le deuxième, situé environ 200 paires de bases plus en amont, possède une fonction répressive.³¹³ Il est intéressant de noter que lorsque YY1 est associé à GATA-1 comme on peut le voir dans les promoteurs de ϵ et de γ (Figure 18) son activité est négative alors que lorsqu'il est seul, il possède une activité positive.

Il est à noter que le gène γ possède également un élément répresseur situé dans sa séquence non traduite 5' (5'UTR) et pouvant lier les facteurs STAT3 β et GATA-1, indiquant que le 5'UTR de ce gène participe également dans sa régulation.^{315;316}

Le promoteur de β -globine

Contrairement à ϵ et γ , le gène β ne démontre pas de régulation autonome lorsque lié au LCR et cela se reflète dans la plus grande simplicité du promoteur adulte (Figure 18). Globalement, le promoteur de β peut être divisé en deux régions. Contrairement aux promoteurs fœtaux et embryonnaires, toute la partie proximale du promoteur de β s'étalant sur une distance de 220 pb de -200 en amont du site d'initiation de la transcription à +20 est presque exclusivement consacrée à l'activation. Ce promoteur ne possède pas en fait de site de liaison aux répresseurs COUP-TFII et DRED qui sont d'importants répresseurs des gènes fœtaux et embryonnaires. Cette région possède toutefois de nombreux éléments de liaison à des activateurs transcriptionnels dont plusieurs sont spécifiques à β . On y retrouve ainsi des sites de liaison pour EKLF,³¹⁷⁻³¹⁹ GATA-1,^{320;321} Sp1,^{322;323} CP2,²⁷⁷ NF-Y,^{265;320;324-326} NF-1,^{265;327} C/EBP δ ,³²⁸ β DRf^{329;330} ainsi qu'un site situé dans le 5'UTR pour le facteur USF.^{331;332} Cette région possède également un site de liaison au facteur C/EBP γ dont la fonction peut être répressive ou activatrice selon le contexte^{326;328} ainsi qu'un site reconnu par la ferritine H, le seul véritable répresseur pouvant lier cette région.^{333;334}

C'est en amont de cette région que se trouve la région à activité répressive responsable d'inactiver β au stade fœtal ou à tout le moins de réduire son affinité pour le LCR. Cette région contient des sites de liaison à des facteurs qui ont été démontrés comme

étant capables de diminuer l'expression de β en culture cellulaire. Parmi ces facteurs, BP1 est celui qui a été le mieux caractérisé^{321;335-337} mais les facteurs BP2³²¹ et HMG^{265;336} seraient également impliqués dans la répression de β . Les facteurs DLX7³³⁸ et CP2²⁷⁷ peuvent également lier cette région mais n'ont soit pas de fonction régulatrice sur le gène de β -globine (DLX7) soit la fonction est inconnue (CP2). Toutefois, l'effet global de cette région sur le gène β est plutôt faible, étant insuffisante pour obtenir une répression complète du gène durant l'érythropoïèse primitive en absence de compétition.

Le LCR

Le plus important et le plus imposant élément *cis*-régulateur du locus est sans contredit le LCR situé en amont du gène ϵ . Si le LCR a été défini au départ comme étant un élément capable de conférer une expression position-indépendante élevée de façon érythroïde-spécifique à un transgène, sa fonction naturelle dans le locus de la β -globine est plus complexe. Ce qui distingue le LCR des autres éléments *cis*-régulateurs du locus, c'est sa non spécificité d'action. C'est-à-dire que le LCR n'est pas un élément spécifique à un ou quelques gènes du locus, mais est important pour l'expression de tous les gènes et possède donc une fonction globale dans le locus. Si la principale fonction du LCR est l'activation transcriptionnelle des gènes, d'autres fonctions lui ont été également attribuées. De fait, l'ouverture de la chromatine dans le locus, l'isolement du locus contre les effets négatifs de la chromatine environnante ainsi qu'une implication dans le timing de la réplication du locus en phase S précoce sont toutes des fonctions attribuées au LCR, bien qu'elles soient toutes de plus en plus controversées.^{133;204;339-342} Généralement, le LCR est perçu comme un *Enhancer* atypique dont la fonction est dépendante de son orientation.³⁴³

Le LCR peut être divisé en cinq éléments distincts caractérisés par leur hypersensibilité à la DNaseI. Ces éléments, les HS1 à HS5 possèdent généralement une région centrale appelée cœur ou *core* possédant la fonction principale d'*enhancer* mais pas nécessairement toutes les fonctions des HS dans leur entier. De fait, lorsqu'un LCR reconstitué à partir du *core* des HS est utilisé avec un gène rapporteur, un effet additif des

différents éléments est obtenu alors qu'un effet synergétique est obtenu avec un LCR reconstitué à partir des HS entiers, indiquant de ce fait que les régions flanquantes des *cores* sont importantes pour la cohésion de LCR et pour assurer sa pleine fonction.³⁴⁴ Les HS se caractérisent également par leur capacité à lier de nombreux facteurs de transcription, érythroïde-spécifiques et ubiquitaires, qui sont souvent communs aux promoteurs des gènes de globines et ont été hautement conservés lors de l'évolution.¹²⁴ Si la plupart des HS seuls sont capables d'activer la transcription des gènes de globine, la fonction intégrale du LCR dépend largement de la présence de tous les HS intacts et des interactions entre ceux-ci.³⁴⁵

HS1

L'élément HS1 est le site le plus proximal des gènes de globine étant situé à environ 6 kb en amont de ϵ . Ce site possède un rôle plutôt limité dans la fonction activatrice du LCR mais la délétion de ce site en transgénique rend le locus plus sujet à la variabilité due à l'effet de position en plus d'affecter spécifiquement l'expression du gène ϵ au stade embryonnaire et γ au stade fœtal.³⁴⁶⁻³⁴⁹ Contrairement aux autres HS, le HS1 ne possède pas de séquence centrale ou *core* spécifique et ne possède essentiellement que des sites de liaison à GATA-1 et Sp1 ce qui en fait un élément structurellement simple.^{124;350}

HS2

Le site HS2 qui est situé à environ 11 kb en amont de ϵ est l'un des deux éléments apportant la plus importante contribution à la fonction du LCR. De fait, lorsque utilisé seul en transgénique lié au gène β , le HS2 est capable d'activer l'expression de β à tous les stades du développement alors que lorsqu'un gène γ est inséré entre les deux, HS2 est suffisant pour contrôler la commutation de l'hémoglobine fœtale à adulte. En fait, le HS2 ne montre aucune préférence pour, ni n'exclut aucun des gènes embryonnaires, fœtaux et adultes.^{101;345;347;351-353} Toutefois, le HS2 n'a pas la capacité à lui seul de conférer une expression indépendante de la position à un transgène et contrairement au LCR, l'activité du HS2 est indépendante de son orientation et de sa position.^{354;355}

Le HS2 peut être divisé en deux sections, le *core* et la région flanquante du *core*. Le *core* possède la fonction principale de *enhancer* du HS2 alors que la région flanquante possède une fonction de modulation du *core*.^{356;357} En effet, lorsque le HS2 est excisé dans son entier, l'impact sur l'expression de tous les gènes du locus est minime alors que lorsque seul le *core* est enlevé, la perte d'activation est beaucoup plus importante.^{125;358} Toutefois, cette dichotomie de la fonction du HS2 est une caractéristique qui ne se retrouve que chez l'humain puisque la délétion du HS2 complet ou simplement du *core* chez la souris ne provoquent qu'une diminution négligeable de l'expression des gènes du locus.^{359;360}

Le HS2, dont le *core* est capable de lier la matrice nucléaire de façon érythroïde-spécifique,³⁶¹ comprend des sites de liaison à de nombreux facteurs érythroïdes et ubiquitaires tels que GATA-1, CBF1, USF, AP-1, EKLF, FKLf2, YY1, SCL, SSP, Sp1 et particulièrement NF-E2 qui sont importants pour sa fonction.^{269;311;350;354;362-365}

HS3

Le site HS3, situé à environ 15 kb en amont du gène ϵ , est avec le HS2 l'un des deux plus importants éléments constituant le LCR. En transgénique, le HS3 est capable de conférer une expression élevée et position-dépendante des gènes qui lui sont lié et ce, de façon similaire mais non identique au HS2. En effet, contrairement au HS2, lorsque lié au gène β , le HS3 active ce gène préférentiellement dans le foie fœtal et dans la moelle osseuse mais non dans le sac vitellin au stade embryonnaire.^{346;353;366} Également, à cause d'une préférence générale du HS3 pour le gène γ , la commutation fœtale à adulte survient plus tard en transgénique lorsque le locus est contrôlé par le HS3 seul que par tout autre HS incluant le HS2.³⁴⁷ De plus, le HS3 possède la capacité d'ouvrir la chromatine dans les transgéniques ce qui se reflète dans le fait que l'activation transcriptionnelle due au HS3 est plus important lorsque le transgène est intégré au génome que lorsque non intégré en essais transitoires. Ceci a mené les chercheurs à proposer que la fonction principale du HS3 dans le LCR pourrait en être une de remodelage de la chromatine.³⁶⁷

Le HS3 possède également une région centrale ou *core* et une large région flanquante dont la partie la plus distale en 5' possède en elle-même des caractéristiques d'un petit HS et est quelques fois désignée HS3,5.^{124;368} La délétion du HS3 entier ne provoque une altération importante de l'expression que sur le gène ϵ , alors que les gènes γ et β ne sont pas affectés. Ceci est l'inverse de ce qui se produit lorsque le HS3 du locus de la souris est délété puisque dans ce cas, les gènes adultes sont plus affectés que les gènes embryonnaires.^{358;369} La délétion ciblée du *core* toutefois, a un impact beaucoup plus grand puisque l'expression de tous les gènes du locus est réduite de façon substantielle et ce, à tous les stades du développement. De plus, la mutation d'une courte séquence GGTGGA nommée GT6, située dans le *core* abolie l'expression des gènes ϵ et γ mais n'affecte pas celle de β , démontrant de ce fait que la préférence du HS3 pour les gènes embryonnaires et fœtaux est dépendante d'un petit élément régulateur du *core*.^{127;129;370}

Comme le HS2, le HS3 comporte de nombreux sites de liaison à des facteurs transcriptionnels tel que GATA-1, EKLF, FKLF2, SSP, Sp1, TEF-2, NF-E2, YY1 et Oct-1.^{350;365;371-373}

HS4

Le HS4 qui est situé à environ 18 kb en amont du gène ϵ , est un élément important du LCR. S'il possède une activité modérée d'activation spécifique de la transcription des gènes fœtaux et adultes du locus lors de l'érythropoïèse définitive, il est incapable toutefois d'activer le gène embryonnaire et les gènes fœtaux durant l'érythropoïèse primitive.^{347;374;375}

Bien qu'étant le plus petit des HS du LCR, Le HS4 est néanmoins également constitué d'un *core* et de séquences flanquantes présentant des fonctions antagonistes. La délétion du HS4 entier, comme dans le cas des HS2 et HS3, n'affecte l'expression des gènes du locus que minimalement. En effet, cette délétion ne résulte qu'en une légère diminution de β au stade fœtal, accompagnée d'une augmentation proportionnelle de γ .³⁴⁹ La délétion du *core* affecte de façon beaucoup plus importante l'expression des gènes du locus, surtout durant l'érythropoïèse définitive alors que son impact sur l'expression des gènes de globine

lors de l'érythropoïèse primitive demeure controversé. En effet, sur ce point, deux études qui ont excisé le *core* du HS4, ont donné des résultats contradictoires.^{128;370} En transgénique, le *core* compte pour l'essentiel de l'activité activatrice du HS4 en plus d'être responsable de sa formation en provoquant l'ouverture de la chromatine.^{375;376}

Le HS4 contient également plusieurs sites de liaison à des facteurs de transcription tant érythroïde-spécifiques qu'ubiquitaires. On notera ainsi des éléments de liaison aux facteurs AP-1, Sp1, EKLF, FKL2, NF-E2 et GATA-1, les deux derniers étant essentiels à la formation érythroïde-spécifique du HS4.^{350;365;376;377}

HS5

Si règle générale le LCR est défini par la présence des quatre sites d'hypersensibilité érythroïde-spécifiques HS1 à HS4, environ 22 kb en amont du gène ϵ se trouve un cinquième site d'hypersensibilité à la DNaseI, ubiquitaire cette fois, appelé HS5. Contrairement aux autres éléments HS, le HS5 ne possède aucune fonction intrinsèque d'activation de la transcription mais son inclusion en 5' et en 3' d'un transgène dirigé par un autre HS permet d'obtenir une expression plus stable du transgène en protégeant ce dernier contre l'effet de position.¹³⁴ Cette fonction d'*insulator* a d'ailleurs été démontrée par l'inclusion du HS5 entre le HS3 et le gène β ce qui produit une inhibition de la transcription de ce dernier.¹³³ D'autre part, l'hypothèse que le HS5 formerait une barrière naturelle protégeant le locus de la β -globin contre les effets de la région chromosomique environnante est supportée par la présence dans le HS5 de sites de liaison au facteur CTCF, impliqués dans la fonction de plusieurs *insulators*, ainsi que d'un site de liaison de la topoisomérase II qui elle, est associée à une fonction de liaison à la matrice nucléaire.^{124;378} Toutefois, cette fonction attribuée au HS5 est de plus en plus controversée puisque différentes études ne sont pas parvenues à démontrer la capacité du HS5 à bloquer l'activation de gènes via le LCR, ou même de bloquer la fonction *enhancer* du LTR ERV-9 situé juste en amont du locus, au point où le HS5 est de plus en plus perçu comme un vestige évolutif ayant perdu sa véritable fonction.³⁷⁹⁻³⁸²

Le site d'hypersensibilité 3'HS

Situé à environ 20 kb en aval du gène β se trouve un autre site d'hypersensibilité à la DNaseI appelé 3'HS. La fonction principale du 3'HS en est une d'*insulator* capable de protéger le locus contre les effets de la région chromosomique environnante et ne possède pas de fonction d'activation de la transcription. Tout comme le HS5, le 3'HS possède des sites de liaison à la topoisomérase II et au facteur CTCF ce qui supporte l'hypothèse que ces deux HS forment les frontières du locus.³⁸³ Cependant, l'absence de transition des modifications des histones au niveau du 3'HS soulève des questions quant à la véritable fonction de cet élément.³⁸¹

Le domaine PYR

Situé à environ 960 pb du site d'initiation de la transcription du gène δ , se trouve une région longue de 250 pb nommée le domaine PYR (pour *pyrimidine-rich region*), qui se caractérise par un pourcentage de pyrimidine supérieur à 95% sur l'un de ses brins. Cet élément recrute un complexe de remodelage de la chromatine de type SWI/SNF appelé complexe PYR qui est formé des sous-unités ubiquitaires BAF57, INI1, BAF60a et BAF170 ainsi que du facteur hématopoïétique Ikaros.^{384;385} Cette région est incluse dans la délétion de type Corfu qui conduit à une forme de HPFH et qui enlève une partie du gène δ ainsi que 6 kb de la région en amont du gène (Figure 6, page 24).¹⁵⁷ De fait, lorsque l'élément PYR a été excisé d'un mini locus comprenant la séquence du locus allant de γ à β , attachée à un μ LCR, un décalage dans la commutation de l'hémoglobine a été mesuré, suggérant une implication pour cet élément dans la régulation.³⁸⁶ Toutefois, lorsque une délétion de la région comprenant le domaine PYR a été produite dans un YAC contenant le locus entier, aucun impact n'a été mesuré sur l'expression des gènes fœtaux et adultes et le rôle véritable de cet élément dans la régulation de la commutation de l'hémoglobine fœtale à adulte, demeure donc inconnue.³⁸⁷

Autres éléments régulateurs

Quelques autres éléments *cis*-régulateurs ont été identifiés dans le locus mais si l'un deux possède une fonction bien établie, les autres sont plus controversés. De ces éléments, le mieux caractérisé est certainement l'*enhancer* situé en 3' du gène β dont la présence est essentielle à une expression élevée de β ^{388;389} en plus d'être impliqué dans la régulation de la réplication du locus.³⁹⁰ D'autre part, l'implication dans la régulation des gènes de globine d'un *enhancer* identifié en 3' de Λ_γ ,³⁹¹ d'un site d'initiation de la transcription intergénique situé entre $\psi\beta$ et δ et soupçonner d'agir comme frontière entre les domaines fœtaux et adultes³⁹² ainsi qu'un LTR du rétrovirus résidant humain ERV-9 situé juste en amont du HS5³⁹³ demeurent très contestées au même titre que l'élément PYR.^{387;394}

1.5.4 Facteurs *trans*-régulateurs

Comme nous l'avons vu dans la section précédente, de nombreux facteurs *trans*-régulateurs peuvent lier des éléments du locus pour activer ou réprimer la transcription des gènes. Cependant, il est difficile d'avoir une idée précise de la fonction d'un facteur donné dans la régulation de la commutation de l'hémoglobine simplement en regardant ses sites de liaison et sa fonction à chacun de ces sites. De fait, la fonction d'un facteur à une période précise à un endroit donné dépend largement du contexte dans laquelle cette protéine se trouve. Le but de cette section est de décrire la fonction globale sur la commutation de quelques facteurs importants.

La famille des KLF

Parce que les gènes de globine sont exprimés de façon séquentielle et exclusive les uns des autres à différents stades du développement, l'identification de facteurs stade-spécifiques pour chaque gène, qui seraient ainsi directement impliqués dans le contrôle de cette expression différentielle, a fait l'objet de nombreuses études. L'identification du facteur érythroïde à doigts de zinc de type Krüppel (EKLF pour *erythroid Krüppel-like*

factors) capable de lier spécifiquement la boîte CACCC proximale du promoteur de β et qui est nécessaire à l'expression de ce gène mais dispensable à l'expression de γ et ϵ , représentait une avancé importante dans l'identification des facteurs *trans*-régulateurs responsables de la régulation de la commutation de l'hémoglobine fœtale à adulte.^{270;395} La découverte subséquente du facteur FKLF, qui démontre une spécificité pour le gène ϵ et du facteur FKLF-2 qui lui, est spécifique aux gènes γ , permettait de présenter un modèle relativement simple expliquant le contrôle des gènes de globine au cours du développement. Cependant les choses se sont avérées plus complexes et tout ne pouvait pas être expliqué par ces facteurs. En effet, le fait que EKLF soit exprimé à tous les stades du développement démontre que si ce gène est essentiel à l'activation de β , entre autre en recrutant des complexes de remodelage de la chromatine comme SWI/SNF et p300/CBP au promoteur de β ^{396;397} et en permettant les interactions entre ce promoteur et le LCR,^{398;399} il est insuffisant à l'induction de la commutation.^{261;400} De plus, la capacité de EKLF à se lier au promoteur des gènes fœtaux et embryonnaires spécifiquement lors de l'érythropoïèse primitive, couplée à sa capacité à recruter des co-répresseurs comme Sin3A et l'histone déacétylase 1 (HDAC1) démontre que la fonction de ce facteur ne se réduit pas simplement à activer β au stade adulte.^{365;401} De fait, si les facteurs de la famille des KLF sont importants et ce, de façon non équivoque, ils ne forment en fait qu'une pièce du grand puzzle de la machinerie *trans*-régulatrice de la commutation de l'hémoglobine.

GATA-1

Comme nous l'avons vu à la section précédente, presque tous les éléments *cis*-régulateurs du locus de la β -globine contiennent des sites de liaison au facteur érythroïde à doigts de zinc GATA-1.²⁶⁹ De fait, l'expression de tous les gènes du locus dépend de la présence de GATA-1 ce qui en fait un facteur majeur dans l'activation ou la répression des gènes de globine puisque ce facteur peut avoir une double activité activatrice ou répressive selon le contexte génomique et les partenaires présents.⁴⁰² Ceci n'est pas très surprenant étant donné l'importance de ce facteur pour l'expression de presque tous les gènes

érythroïdes.⁴⁰³ GATA-1 agit en permettant le recrutement de complexe de remodelage de la chromatine et du LCR aux promoteurs de globine et fonctionne en combinaison avec son cofacteur FOG-1 et les facteurs de la famille des KLF.⁴⁰⁴⁻⁴⁰⁶ Étant un facteur impliqué de façon générale dans la régulation de la transcription, la fonction de GATA-1 dans la régulation de la commutation de l'hémoglobine est plutôt limitée puisqu'il lui manque une spécificité essentielle à une régulation génique différentielle. Toutefois, Li *et al*⁴⁰⁷ ont démontré en surexprimant GATA-1 dans des souris transgéniques arborant le locus humain entier que ce facteur peut agir comme répresseur spécifique du gène ϵ alors que Costa *et al*³¹² ont démontré qu'un site de liaison à GATA-1 situé à -570 du site d'initiation de la transcription du gène γ^A était important pour sa répression lors de l'érythropoïèse définitive.

NF-E2

Si le membre de la famille des facteurs de transcription à leucine-zipper basique, NF-E2 ne possède que peu de sites de liaison au niveau des promoteurs de globine, son importance dans la régulation de l'expression de ces gènes n'en est tout de même pas moindre.⁴⁰⁸ De fait, NF-E2 possède beaucoup de sites de liaison dans le LCR et sa capacité à interagir avec les facteurs de la famille des KLF, GATA-1 et certains complexes de remodelage de la chromatine, en fait un élément important impliqué dans la formation des HS, dans l'ouverture de la chromatine et possiblement dans les interactions LCR-promoteur bien qu'il ait été démontré comme n'étant pas essentiel à ces interactions.^{205;409-413} Toutefois, comme pour GATA-1, NF-E2 manque cette spécificité nécessaire à en faire un régulateur majeur de la commutation de l'hémoglobine, ce qui est mis en lumière par le fait que l'ablation génique de ce gène n'a pas d'impact sur la commutation.⁴¹⁴

Ikaros

Ikaros est un facteur à doigts de zinc formant la sous-unité du complexe PYR et responsable de la liaison du complexe à l'ADN. Comme il a été dit à la section 1.5.3, le complexe PYR peut se lier à un élément situé en amont de δ . La fonction du complexe PYR

sur la commutation de l'hémoglobine demeure toutefois controversée puisque des études indépendantes ayant délété son élément de liaison, ont mené à des conclusions opposées quant à son implication dans cette régulation.^{386;387} Toutefois, l'ablation génique d'Ikaros provoque un retard dans l'établissement de la commutation de l'hémoglobine fœtale à adulte, ce qui suggère que ce facteur joue bel et bien un rôle dans le timing de la répression des gènes fœtaux.⁴¹⁵ Puisque d'autres sites capables de lier Ikaros de façon indépendante du complexe PYR ont été identifiés à plusieurs endroits dans le locus, il est possible que cette fonction d'Ikaros, ne passe pas par le complexe PYR lui-même.⁴¹⁶

Autres facteurs de transcription

Quelques autres facteurs sont importants dans la régulation de l'hémoglobine et parmi ceux-ci on retrouve BP1, impliqué dans une répression partielle de β au stade fœtal³²¹; DRED et COUP-TFII, tous deux impliqués dans la répression des gènes fœtaux et embryonnaire^{272;417}; YY1,²⁸³ qui est associé à la répression de ϵ et Sp1 dont la fonction global est plus complexe puisque ce dernier est impliqué aussi bien dans l'activation que dans la répression des gènes du locus.^{306;350;410}

1.5.5 Épigenétique

L'activation et l'inactivation des gènes de globine lors de la commutation de l'hémoglobine sont des processus complexes faisant intervenir comme nous l'avons vu dans les quelques sections précédentes, une étroite collaboration entre de nombreux éléments *trans*-régulateurs et facteurs *cis*-régulateurs, mais également comme nous le verrons dans cette section, d'importantes modifications au niveau de la structure de la chromatine. Si on sait depuis longtemps que les différents domaines géniques du locus de la β -globin subissent des changements épigénétiques lors du développement, les mécanismes qui régulent ces changements ne sont toujours pas très bien compris. Certes, les facteurs de transcription comme EKLF, GATA-1 et NF-E2 sont reconnus comme étant capables de recruter de nombreux complexes de remodelage de la chromatine au niveau des différentes

régions régulatrices de l'expression des gènes de globine, mais l'ordre des événements qui contrôlent ce remodelage demeure largement débattu.^{405;410;418-420} En effet, la question à savoir si c'est la liaison des facteurs de transcription qui permet l'ouverture de la chromatine à un lieu et un temps donnés ou si c'est l'ouverture de la chromatine qui précède et permet la liaison de ces facteurs, reste toujours sans réponse définitive. L'existence d'une transcription stade-spécifique des régions intergéniques et indépendante de la transcription génique,⁴²¹⁻⁴²³ ainsi qu'une potentialisation épigénétique du locus précédant, ou étant indépendante de son activation transcriptionnelle,^{424;425} suggèrent que l'activation des promoteurs et les modifications de la chromatine sont des événements distincts mais capables d'influencer l'autre.⁴²⁶

Domaines fœtaux et adultes

Au cours du développement, le locus subit de profonds changements au niveau de la chromatine, généralement en corrélation avec les changements de l'activité transcriptionnelle des gènes du locus. Aux stades embryonnaire et fœtal, une grande portion du locus comprenant les gènes ϵ et les deux gènes γ , présente une configuration très euchromatique, c'est-à-dire que la chromatine se retrouve dans un état plus permissif à la transcription alors que la région du locus où se trouvent les gènes adultes δ et β , se retrouve dans une conformation plus hétérochromatique, donc moins permissive à l'activité transcriptionnelle. Après la naissance, lorsque la commutation de l'hémoglobine fœtale à adulte prend place, un remaniement important de la chromatine du locus a également lieu, se caractérisant par le déclin de la permissivité de la région contenant les gènes ϵ , γ^G et γ^A , et par l'ouverture de la région contenant les gènes δ et β . Ainsi, la région contenant les gènes ϵ et γ est appelée "domaine fœtal" alors que la région contenant δ et β est appelé "domaine adulte". À ces deux domaines épigénétiques, s'additionne un troisième domaine, celui du LCR, qui possède une conformation ouverte à tous les stades du développement.^{427;428}

Ces domaines de la chromatine se caractérisent par une sensibilité générale à la DNaseI variable, une signature différentielle des modifications covalentes des histones

ainsi que par un enrichissement ou un appauvrissement de la méthylation de l'ADN au niveau des cytosines.

Sensibilité à la DNaseI

Lorsque l'ADN est extrait sous sa forme native, c'est-à-dire sous sa forme de chromatine et qu'il est digéré à la DNaseI, certaines régions sont plus difficiles à digérer que d'autres. On sait depuis longtemps que la sensibilité à la DNaseI est indicative d'une région plus permissive à la transcription que les régions plus résistantes à la DNaseI. De fait, il est généralement observé que la région promotrice des gènes actifs est sous une forme hypersensible à la DNaseI alors que celle des gènes inactifs l'est beaucoup moins.⁴²⁹⁻

⁴³¹ D'un point de vue moléculaire, l'hypersensibilisation de la chromatine indique une plus grande accessibilité de la DNaseI à l'ADN qu'une région résistante. Cette plus grande accessibilité pour l'endonucléase implique de fait, une accessibilité accrue pour les facteurs de transcription et pour la machinerie transcriptionnelle basale due à un relâchement générale de la chromatine.^{432;433} Toutefois, si l'accessibilité d'un promoteur à la machinerie transcriptionnelle est obligatoire, elle n'est pas suffisante à sa transcription et n'explique pas comment ou pourquoi un gène donné est exprimé.^{341;434} Cependant, l'hypersensibilité est une indication du potentiel transcriptionnel d'un gène. Dans le locus de la β -globine, outre le LCR qui se définit par des régions hautement sensibles à la DNaseI, le promoteur des gènes de globine le sont également mais de façon stade-dépendante. En effet, aux stades embryonnaire et fœtal, les promoteurs de tous les gènes de globine se retrouvent dans une conformation hypersensible alors qu'au stade adulte les promoteurs fœtaux deviennent résistants aux nucléases tandis que les promoteurs des gènes adultes demeurent hypersensibles, indiquant un contrôle plus rigoureux de la chromatine du domaine fœtal que celui du domaine adulte.⁴³¹

Histones

Si la sensibilité à la DNaseI est une indication qualitative du niveau général d'ouverture de la chromatine, elle renseigne peu sur la nature moléculaire de cette "ouverture". L'étude des différentes modifications covalentes des histones d'une région chromosomique donne en revanche une image beaucoup plus détaillée de l'état de la chromatine. Il existe plusieurs types de modifications des histones dont les deux plus importantes sont l'acétylation et la méthylation de différents résidus lysine des queues du domaine N-terminal de ces histones. Les différentes combinaisons de ces modifications forment le code épigénétique et déterminent la permmissivité à la transcription d'un domaine génétique.⁴³⁵ Une hyperacétylation de l'histone H4 et une acétylation de la lysine K9 de l'histone H3 sont des marqueurs d'une chromatine permissive, de même que la tri-méthylation du résidu K4 de l'histone H3.⁴³⁶⁻⁴³⁸ À l'inverse, la méthylation de la lysine K9 de l'histone H3 ainsi qu'une mono-méthylation de la lysine K4 sont des indicateurs d'une hétérochromatine, donc d'une région non permissive à la transcription.⁴³⁷ Une di-méthylation en K4 de l'histone H3 représente un état intermédiaire et les régions ainsi marquée peuvent être ou ne pas être transcrites selon le contexte chromosomique et génétique.^{439,440}

Dans le locus de la β -globine une signature épigénétique différentielle au niveau des différents gènes de globine en fonction de leur activité transcriptionnelle est également remarquée.^{425;441;442} En effet, les gènes transcrits à un stade donné présentent par exemple des niveaux de tri-méthylation en K4 de l'histone H3 plus élevés que les gènes non transcrits qui eux présentent des niveaux de mono- et di-méthylation plus importants.³⁸¹ Cette dichotomie entre les gènes embryonnaire/fœtaux et les gènes adultes est également observée par une hyperacétylation des histones H4 et H3 au niveau des gènes transcrits alors que les gènes non transcrits sont plutôt hypoacétylés.⁴⁴³ Il est également intéressant de noter que les sites d'hypersensibilité HS2 et HS3 qui sont les deux plus importants éléments du LCR, se caractérisent par des niveaux de tri-méthylation en K4 de l'histone H3 plus élevés que les autres HS.³⁸¹

L'ouverture de la chromatine dans le domaine approprié lors du développement est certainement un mécanisme important impliqué dans la régulation de la commutation de l'hémoglobine. Cependant, il est difficile de déterminer quel est l'événement qui contrôle ce remodelage épigénétique. En effet, si une chromatine "ouverte" favorise la transcription des gènes, cette activité transcriptionnelle est également impliquée dans l'ouverture de la chromatine puisque la machinerie transcriptionnelle comporte des complexes à fonctions d'acétylation des histones.^{423;444;445}

Méthylation de l'ADN

Le locus de la β -globine se distingue de la plupart des autres loci du génome par le fait qu'il soit AT riche et par conséquent, dépourvu d'îlots CpG qui sont généralement retrouvés en nombre important dans la région promotrice des gènes. La méthylation en position 5 des cytosines (5mC) de ces îlots CpG est associée à une fermeture de la chromatine,⁴⁴⁶ donc à une répression de la transcription en permettant le recrutement de répresseurs tel que les facteurs MBD.⁴⁴⁷ Si les gènes du locus sont dépourvus de ces îlots, la présence de 5mC au niveau des quelques séquences dimériques CpG des promoteurs des gènes du locus peut quand même être mesurée. Toutefois, si cette méthylation que l'on peut retrouver de façon spécifique au niveau des promoteurs inactifs des gènes de globine, est responsable de l'inactivité transcriptionnelle ou si elle en est la conséquence n'est pas connue.⁴⁴⁸ Plusieurs études ont démontré néanmoins qu'une méthylation 5mC dans le promoteur des gènes embryonnaires, fœtaux et adultes pouvait contribuer à l'inactivation de ces gènes.⁴⁴⁹⁻⁴⁵³ De plus, le fait que tout le domaine embryonnaire chez la souris soit largement hypométhylé uniquement durant l'érythropoïèse primitive, supporte un rôle de la méthylation dans la régulation de la commutation de l'hémoglobine.⁴⁵⁴ D'ailleurs, il est intéressant de noter que l'élément SSE du promoteur de γ (Figure 18, page 44) est, lorsque méthylé, lié préférentiellement par le facteur Sp1 qui agit alors comme un répresseur de la transcription en déplaçant le complexe activateur SSP.²⁹⁸

1.5.6 Modèles de régulation

Certes, une analyse exhaustive de chaque élément régulateur peut donner une idée des différents mécanismes qui régissent de façon coordonnée l'expression des gènes de globine au cours du développement. Cependant, à cause de la complexité de ces mécanismes, l'analyse de chaque élément séparé peut difficilement donner une vision globale du mode de fonctionnement du locus qui se comporte plutôt comme une entité aux éléments interdépendants. De nombreuses études ont tenté de comprendre la mécanistique globale derrière la commutation de l'hémoglobine. L'une des difficultés consistait à développer des techniques permettant de détecter les interactions physiques entre les différentes régions du locus de façon à obtenir une image de la structure tridimensionnelle du locus aux différents stades du développement. Les techniques de PIN*POINT,³²³ RNA-FISH-trap,⁴⁵⁵ ainsi que la capture de conformation chromosomique 3C,⁴⁵⁶ 4C⁴⁵⁷ et 5C⁴⁵⁸ ont d'ailleurs été développées dans cette optique. On savait depuis longtemps que la régulation de la commutation foetale à adulte se faisait certainement par un mécanisme de compétition des promoteurs foetaux et adultes pour l'activité du LCR,¹⁹⁹ mais la nature moléculaire et la mécanique de cette compétition demeuraient sujet à débat.

Deux grands modèles aujourd'hui sont proposés pour décrire la mécanique de fonctionnement du locus. Si les études de conformation ont permis d'appuyer la validité de ces modèles, elles n'ont pas permis de déterminer une fois pour toute, lequel de ces modèles est représentatif de la réalité. De fait, il est probable que la régulation du locus passe par différentes voies et que chacun de ses modèles prit individuellement ne peut tout expliquer et ne représente en fait qu'une partie de la réalité.

Le modèle du *Looping*

Le modèle où un élément activateur du locus de la β -globine, situé à grande distance du gène qu'il active, vient directement en contact avec le promoteur grâce à la formation d'une boucle de la chromatine et nommé le modèle du *Looping* a été proposé dès

1989.⁴⁵⁹⁻⁴⁶¹ Dans ce modèle, se sont les facteurs transcriptionnels qui forment un pont entre le promoteur actif et le grand holocomplexe du LCR qui est lui-même formé par des ponts protéiques entre les différents HS (Figure 19). Des analyses de la conformation spatiale du locus ont démontré en effet que le LCR se retrouve à proximité des gènes actifs du locus ce qui supporte fortement le modèle du *Looping*.^{399;462;463} Grâce à ces interactions, le LCR est libre d'ouvrir la chromatine, de recruter la machinerie transcriptionnelle de base ou encore d'activer la transcription. C'est donc l'affinité avec laquelle un promoteur peut attirer le LCR qui détermine le temps que ce dernier passera auprès de lui et l'activera. Si deux promoteurs sont actifs au même moment, ils entrent alors en compétition pour le LCR qui se met à faire la navette entre les deux par un mécanisme appelé *flip-flop* en passant plus de temps auprès du promoteur ayant l'affinité pour le LCR la plus grande.^{112;464} Si les deux promoteurs ont une affinité intrinsèque équivalente, le LCR ira préférentiellement au plus proche.

Cependant, la faiblesse de ce modèle réside dans le fait qu'un mécanisme de *looping* ne peut expliquer pourquoi l'orientation et la position du LCR sont importantes pour l'activation des gènes de globine. En effet, cette dépendance à l'orientation du LCR suppose la présence d'éléments exogènes aux gènes du locus et au LCR, impliqués dans le positionnement approprié de ce dernier au niveau du promoteur et ce, de façon dépendante de la position et de l'orientation.

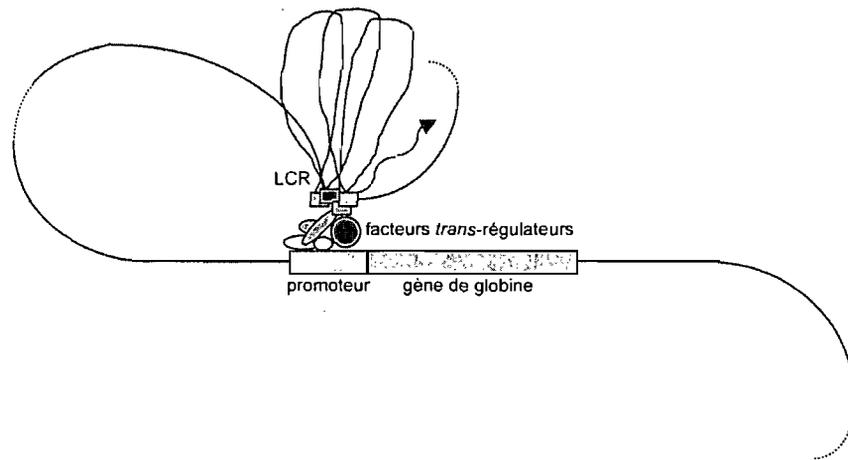


Figure 19: Modèle du *Looping*. Adapté de Harju *et al.*²⁴⁶

Le modèle du *Tracking*

Pour répondre au problème de la dépendance à l'orientation de l'activité du LCR sur les promoteurs de globine soulevé par le modèle du *Looping*, un modèle où la machinerie transcriptionnelle lie en premier le LCR et se déplace ensuite le long de l'ADN jusqu'à ce que le promoteur d'un gène actif soit rencontré a été proposé (Figure 20).^{355;465} Selon le modèle du *Tracking*, c'est l'inactivation des gènes proximaux qui permet une activation des gènes plus distaux puisque le LCR ne se rend qu'au premier gène actif qu'il rencontre. L'observation d'une transcription intergénique LCR-dépendante, orientée dans le même sens que la transcription des gènes du locus et couvrant presque la totalité de ce dernier, vient supporter ce modèle.⁴²³ Toutefois, si le modèle du *Tracking* explique de façon satisfaisante l'importance de la position et de l'orientation du LCR, l'expression différentielle des gènes du locus au cours du développement ainsi que la présence d'une activité transcriptionnelle des régions intergéniques, il ne permet pas d'expliquer le fait que souvent, plus d'un gène sont exprimés dans les mêmes cellules comme c'est le cas des deux gènes fœtaux,^{112;464} ni le fait que la transcription intergénique est spécifique au domaine

actif puisque chez l'adulte, le domaine fœtal n'est que peu transcrit.³⁹² Il est donc en fait improbable que ce mécanisme s'applique aux gènes fœtaux et adultes du locus humain de la β -globine. Toutefois, ϵ dont l'activité transcriptionnelle est dépendante de sa distance au LCR, se prête mieux au modèle du *Tracking* qu'au modèle du *Looping* d'autant plus que l'insertion d'un signal d'arrêt de la transcription entre le LCR et ϵ inhibe l'expression de celui-ci, ce qui suggère que ce modèle pourrait s'appliquer au gène embryonnaire.^{254;443}

Pour pallier aux problèmes soulevés par le modèle du *Tracking* sur l'expression des gènes de γ et de β et en tenant compte des problèmes du modèle du *Looping*, une version alternative combinant ces deux modèles a également été proposé. Le modèle du *Tracking* facilité propose que l'holocomplexe du LCR, comportant la machinerie de transcription de base vient en contact avec une région située en amont des gènes actifs grâce à la formation d'une boucle de la chromatine, puis se met à se déplacer le long de l'ADN jusqu'à ce qu'un gène actif soit rencontré (Figure 20) ce qui permet de combiner les avantages des deux précédents modèles en plus d'expliquer la transcription intergénique spécifique aux domaines épigénétiques actifs.⁴⁶⁶

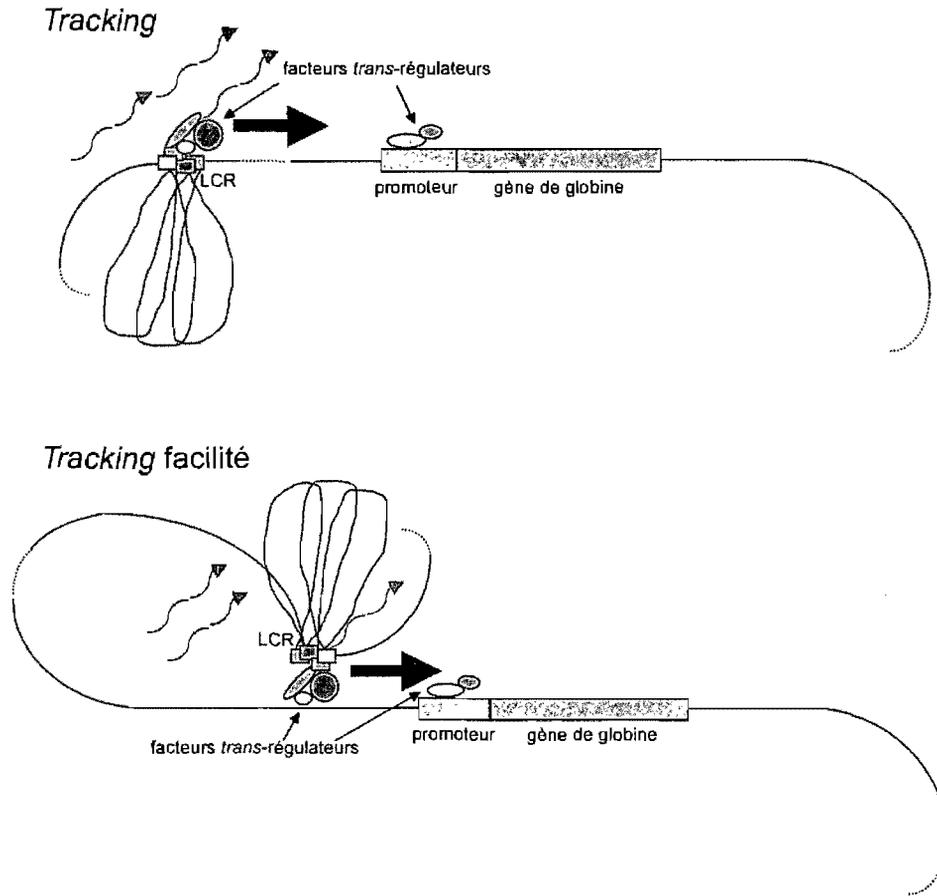


Figure 20: Modèles du *Tracking* et du *Tracking facilitation*. Adapté de Harju *et al.*²⁴⁶

1.6 Systèmes modèles

Comme nous l'avons vu dans les sections précédentes, le locus humain de la β -globine possède des caractéristiques qui sont assez uniques chez les mammifères. L'expression strictement fœtale des gènes γ et strictement adulte des gènes β et δ ne s'observe que chez les singes supérieurs. La seule exception se retrouve chez les ruminants, qui par un mécanisme complètement différent, sont arrivés à un patron d'expression similaire et représentent un exemple d'évolution convergente. Nous avons vu également à la section 1.3.7 (p.21) que plusieurs hémoglobinopathies de la β -globine chez l'humain sont causées par des tares génétiques et pourraient être compensées si la commutation de l'hémoglobine pouvait être contrôlée. En effet, il a été démontré que, par exemple dans le cas de la drépanocytose, l'hémoglobine fœtale pouvait avoir un effet thérapeutique par sa capacité à inhiber la polymérisation de la HbS,^{144;156} ou encore que l'hémoglobine Gower-2 ($\alpha_2\epsilon_2$) avait les mêmes propriétés physicochimiques que la HbA,³⁷ ce qui pourrait en faire une cible thérapeutique intéressante pour les cas de β -thalassémie ou même de $\gamma\delta\beta$ -thalassémie. Pour ces raisons, la compréhension des mécanismes qui régulent la commutation de l'hémoglobine chez l'humain est très importante. Cependant, à cause de la difficulté à étudier ces mécanismes aux différents stades du développement directement chez l'homme et aux coûts prohibitifs liés à l'étude sur les grands primates, le développement de systèmes modèles pour l'étude de la génétique de l'hémoglobine est devenu capital.

1.6.1 Études en culture cellulaire

L'analyse en 1975 de cellules provenant d'une patiente de 53 ans souffrant d'une leucémie myéloïde chronique en phase terminale permis d'isoler ce qui allait devenir le premier modèle cellulaire de la régulation fœtale de l'hémoglobine: les cellules érythroleucémiques K562.⁴⁶⁷ Ces cellules, ainsi qu'une autre lignée de cellules érythroleucémiques humaines, les HEL,⁴⁶⁸ se caractérisent par l'expression inductible des gènes de

globine embryonnaires et fœtaux mais non adultes.^{469;470} C'est en comparant les cellules non-induites avec les cellules induites que les chercheurs ont démontré que la chromatine du locus changeait vers une conformation ouverte lorsque les cellules étaient induites et que cette ouverture de la chromatine était spécifique aux gènes transcrits.⁴⁷¹ Des études utilisant des cellules non-humaines, transfectées avec les gènes humains de globine ont été également effectuées avec toutefois des résultats plutôt mitigés, menant même quelques fois à des résultats artéfactuels tels que l'identification d'un supposé *enhancer* en aval du gène γ .^{389;391;472-475} Malgré les limitations des modèles cellulaires pour l'étude des gènes de globine durant le développement, une commutation de l'hémoglobine a pu être obtenue par des études en hétérocaryons. En effet, la fusion de cellules HEL avec des cellules érythroleucémiques de souris MEL qui présentent une expression spécifique des gènes adultes, a mené à l'activation des gènes adultes humains bien que les gènes fœtaux soient demeurés actifs.^{476;477} Cela démontrait que l'activation des gènes adultes chez l'humain était due à des facteurs *trans*-régulateurs adulte-spécifiques, ce qui fut confirmé lorsque les hétérocaryons MEL-fibroblastes et MEL-lymphoblastes fœtaux humain eurent démontré l'activation spécifique des gènes de globine humaine adulte.⁴⁷⁸ Des cellules érythroïdes non leucémiques ainsi que les progéniteurs érythroïdes ont également été utilisées avec certains succès pour l'étude de la régulation des globines^{479;480} mais les limitations inhérentes aux modèles cellulaires ont vite fait de pointer à la nécessité de générer des modèles plus adaptés à l'étude de la régulation développementale des gènes de globine. L'utilisation plus récente des cellules souches embryonnaires humaines (hES) et leur différenciation en corps embryoides permettant de récapituler la commutation de l'érythropoïèse primitive à définitive, représente une avenue intéressante mais pour le moment peu pratiquée à cause de la difficulté à obtenir le matériel biologique.^{249;481-483}

1.6.2 Les locus non humains

Parce que les modèles cellulaires ont vite démontré leur limites, plusieurs chercheurs se sont tournés vers des modèles non humains mais ayant l'avantage de

permettre l'étude du locus de la β -globine aux différents stades du développement et surtout de récapituler parfaitement la commutation de l'érythropoïèse, c'est-à-dire le déplacement du site érythropoïétique au cours du développement. Cette particularité est importante puisqu'elle assure un microenvironnement adéquat pour la régulation des gènes érythroïdes aux différents stades du développement. Il devenait ainsi donc possible d'étudier le comportement du locus dans un environnement légitime permettant de reproduire les différents stades du développement.

Les ruminants

Des années soixante jusqu'au début des années quatre-vingt les ruminants, et plus particulièrement la chèvre, ont été des modèles de choix pour l'étude de la commutation de l'hémoglobine. En effet, hormis l'humain, les ruminants sont les seuls animaux à posséder une hémoglobine strictement fœtale, distincte des l'hémoglobines embryonnaires et adultes.^{226;237;239} Toutefois, outre la difficulté inhérente au maintien de colonies de grands animaux, il devint rapidement évident, à cause de différences structurelles importantes entre le locus de la β -globine des ruminants et celui de l'humain, que les découvertes faites chez ces animaux ne pouvaient être transposées directement à l'homme.²³⁰⁻²³³

Le poulet

Malgré une structure du locus de la β -globine et une régulation des gènes très différentes du locus humain,¹⁹⁸ le poulet a été largement étudié entre autre au niveau de son élément *enhancer*, mais est surtout devenu un modèle important pour l'étude de la chromatine dans le locus.⁴⁴¹

La souris

De tous les modèles non humains, la souris est de loin celui qui a été le plus étudié. Étant plus rapprochée évolutivement de l'humain que les ruminants, une grande majorité des découvertes faites sur le locus de la β -globine de ce petit rongeur ont pu être

transposées à l'homme.^{202;203;205;213;214} Cependant, de nombreuses études ont également fait ressortir des différences avec l'humain^{425;558;559} et quelques unes ont même donné des résultats opposés, entre autre en ce qui a trait au rôle potentiel du LCR dans l'ouverture de la chromatine dans le locus.^{204;349} Ces études, bien que ne remettant pas en question l'importance des découvertes faites sur le locus murin dans notre compréhension des mécanismes de régulation des gènes de globine humaine, illustrent toutefois les limites de ce système modèle.

1.6.3 Le locus humain en souris transgéniques

Le développement de souris transgéniques permis une grande avancée dans la recherche sur la régulation des gènes de globine. En effet, ce système modèle est le seul à permettre l'étude du locus humain dans un contexte physiologique. La présence du locus endogène représente même un atout dans ce modèle puisqu'elle permet une comparaison directe entre l'expression du transgène humain et les gènes de la souris.

Les constructions tronquées

Puisqu'au départ la taille des transgènes étaient limitée par la capacité des différents vecteurs existants, les premières études en transgéniques ont utilisé des constructions tronquées, c'est-à-dire ne comportant pas le locus entier mais seulement quelques portions reliées entre-elles. Ces études permirent d'identifier de nombreux éléments importants pour la régulation des gènes de globine et permirent d'élaborer les bases de ce qui allait devenir le modèle de régulation par compétition des gènes pour le LCR (Section 1.5, page 41). Grâce à la facilité à générer des transgéniques avec des constructions tronquées et puisque beaucoup d'informations pertinentes peuvent être obtenues de ces constructions, même aujourd'hui de nombreuses études sont effectuées avec cette approche.

L'utilisation du locus entier en transgénique

Le développement de vecteurs à grande capacité a permis de générer plusieurs souris transgéniques arborant le locus entier de la β -globine humaine. De ces études, est apparu l'évidence que l'intégrité générale du locus était importante pour obtenir une régulation fine et appropriée des différents gènes du locus. De fait, de plus en plus l'utilisation de ces transgéniques, malgré la plus grande difficulté à les obtenir, devient la norme puisque les résultats obtenus le sont en fonction d'un contexte global plus complet et adéquat qu'avec les constructions tronquées et sont donc plus reproductibles et représentatifs.

Les vecteurs

Le vecteur le plus employé pour l'étude du locus entier en transgénique est sans contredit le chromosome artificiel de levure ou YAC qui permet des inserts de près d'un mégabase.^{108;110;111;253;484;485} De plus, la machinerie de recombinaison homologue inhérente à la levure a permis l'élaboration de protocoles de mutagenèse dirigée des YAC très performants qui ont permis l'introduction d'une variété de mutations dans le locus et d'étudier leur impact sur la commutation de l'hémoglobine fœtale à adulte.^{207;343;370;486}

Bien que les cosmides n'aient pas la capacité à recevoir des inserts suffisamment grands pour couvrir intégralement le locus humain, une lignée de souris a pu être obtenue par la ligation de deux cosmides de façon à reconstituer le locus entier.⁴⁸⁷ Cette souris a le mérite d'avoir été le premier modèle de souris transgénique à posséder le locus entier et elle est encore quelques fois utilisée comme contrôle pour comparaison avec des mutants. Quelques études utilisant cette approche ont également pu produire des souris transgéniques ayant le locus humain entier comportant des mutations et dont certaines sont devenues d'importants modèles pour les études sur la variabilité (PEV).^{348;463;488;489} Cependant, la facilité à manipuler les YAC a permis à ces derniers de supplanter rapidement l'utilisation des cosmides.

Plus récemment, l'utilisation de chromosomes artificiels de bactérie (BAC) plus pratiques et faciles d'utilisation que les YAC a également permis de générer des souris transgéniques.⁴⁹⁰⁻⁴⁹⁴ Bien qu'encore d'utilisation timide, il est probable qu'à l'avenir les BAC deviennent de plus en plus utilisés dans ces études comme ils sont devenus la norme dans les études du locus de la α -globine.⁴⁹⁵⁻⁴⁹⁷ Quelques laboratoires ont également utilisé des chromosomes artificiels du phage P1 (PAC) pour introduire efficacement des mutations dans le locus de la β -globine humaine et générer des souris transgéniques.^{379;498}

Ultimement, l'utilisation de *knock-in* du locus humain dans le locus de la souris permettant d'étudier la régulation de ces gènes dans un contexte chromosomique plus naturel et conservé d'une lignée à l'autre, deviendra la norme ce qui maintenant est devenu possible grâce à de nouvelles techniques de recombinaison homologue en cellules ES assistée de la *Cre* recombinase.⁴⁹⁹ De même, l'utilisation de chromosomes artificiels humains (HAC) en souris transgéniques pourrait permettre d'éliminer tout effet de position puisque ces chromosomes peuvent se répliquer indépendamment dans la souris sans avoir à s'intégrer dans le génome, et ont démontré une expression tissu-spécifique des gènes du locus humain de la β -globine.⁵⁰⁰

La commutation du locus humain en souris transgéniques

Il a été systématiquement observé que chez la souris transgénique, la commutation de l'hémoglobine fœtale à adulte des gènes humains se produit bien avant la naissance, en même temps que la commutation de l'hémoglobine embryonnaire à fœtale endogène.^{111;207;214;484} Puisque chez la souris il n'y a qu'une commutation de l'hémoglobine (Section 1.4.1, page 29) cela indique que les facteurs de transcription qui régulent la commutation de l'hémoglobine embryonnaire à fœtale/adulte chez la souris sont, à tout le moins en partie, les mêmes qui régulent la commutation de l'hémoglobine fœtale à adulte chez l'humain. Ceci se reflète par l'expression quelque peu anormale du gène embryonnaire humain ϵ chez la souris transgénique.^{111;112;207}

L'effet des mutations sur la commutation

L'intérêt majeur des modèles de souris transgéniques était certainement l'identification des éléments régulateurs de la commutation de l'hémoglobine fœtale à adulte, qui pouvait être obtenue par la génération de mutations dans le locus et par l'analyse des effets qu'elles avaient sur l'expression des gènes. De fait, l'insertion de mutations reconnues chez l'humain pour provoquer une dérégulation de la commutation de l'hémoglobine fœtale à adulte, a généralement conduit à reproduire cette dérégulation dans les souris transgéniques (Tableau II). D'autre part, des délétions ou des mutations produites à des sites précis dans le locus a effectivement permis d'identifier des éléments importants dans le locus (Tableau III) et au niveau du LCR (Tableau IV) en plus de démontrer l'importance de l'organisation générale du locus sur la régulation des gènes (Tableau V) et de mettre en évidence la synergie existant entre différents éléments du locus en combinant des mutations (Tableau VI).

Tableau II: Transgéniques avec le locus humain reproduisant des mutations humaines.

Mutations	Vecteur	Expression			Ref
		embryon	Fœtus	Adulte	
HPFH grecque γ : -117 (G→A)	YAC (248 kb)	ϵ :-, γ :-, β :-	ϵ :-, γ :↑, β :↓	ϵ :-, γ :↑, β :↓	207;501
HPFH2 Δ 83,5 kb ($\Delta\delta\beta$)	YAC (213 kb)	γ :↓	γ :↓	γ :-	135
β^0 -thal β : Δ 4 pb (codon 41-42)	BAC (183 kb)		γ :↑	β :↓	502

Légende: Expression relative au locus sauvage: -, aucun changement (100% de la valeur obtenue avec le locus sauvage); ↑, augmentée [↑(>400%) ↑(200-400%) ↑(150-200%) ↑(<150%)]; ↓, diminuée [↓(<25%) ↓(25-50%) ↓(50-75%) ↓(>75%)]; ∅, non exprimé (0% d'expression). Δ : délétion.

Tableau III: Transgéniques avec le locus humain modifié au niveau des gènes.

Mutations	Vecteur	Expression			Ref
		embryon	Fœtus	Adulte	
boîte CAC ajoutée au promoteur de ϵ	YAC (150 kb)	$\epsilon:\uparrow, \gamma:\downarrow$	$\epsilon:\uparrow, \gamma:\downarrow, \beta:-$	$\epsilon:\uparrow, \beta:-$	503
ϵ^{Bepsi}	YAC (150 kb)	$\epsilon:\uparrow, \gamma:\downarrow$	$\epsilon:\uparrow, \gamma:-, \beta:-$	$\epsilon:\uparrow, \beta:-$	503
Δ Silencer en 5' de ϵ (-304 à -179)	YAC (150 kb)	$\epsilon:\emptyset, \gamma:\downarrow, \beta:-$	$\epsilon:-, \gamma:-, \beta:-$	$\epsilon:-, \gamma:-, \beta:-$	486
Δ Silencer en 5' de ϵ (-406 à -182)	YAC (213 kb)	$\epsilon:-, \gamma:-$	$\epsilon:-, \gamma:-, \beta:-$	$\beta:-$	504
λ (2,3 kb) ajouté en 5' de ϵ (-2 kb)	YAC (150 kb)	$\epsilon:\downarrow, \gamma:-$		$\beta:-$	382
$\Delta\gamma$: -116(A \rightarrow T), -110 (A \rightarrow C), -109(G \rightarrow T)	YAC (150 kb)	$\epsilon:-, \gamma:-$	$\gamma:\uparrow, \beta:-$	$\gamma:\uparrow, \beta:-$	301
$\Delta\gamma$: 9 substitutions de -124 à -100	YAC (150kb)	$\epsilon:-, \gamma:-$	$\gamma:-, \beta:-$	$\gamma:-, \beta:-$	301
Δ enhancer en 3' de $\Delta\gamma$ (Δ 750 pb)	YAC (150 kb)	$\epsilon:-, \gamma:-, \beta:-$	$\epsilon:-, \gamma:-, \beta:-$	$\epsilon:-, \gamma:-, \beta:-$	387;505
Δ enhancer/F en 3' de $\Delta\gamma$ (Δ 1,6 kb)	PAC (185 kb)	$\epsilon:\uparrow, \gamma:\uparrow$	$\gamma:-, \beta:-$	$\beta:-$	498
$\Delta \psi\beta$ (Δ 12,5 kb)	YAC (150 kb)	$\epsilon:-, \gamma:-, \beta:-$	$\epsilon:-, \gamma:-, \beta:-$	$\epsilon:-, \gamma:-, \beta:-$	387
Δ 5'PYR (Δ 2,5 kb)	YAC (220 kb)	$\gamma:-, \beta:-$	$\gamma:\downarrow, \beta:\downarrow$	$\beta:\downarrow$	506
Δ PYR (Δ 3,2 kb)	YAC (150 kb)	$\epsilon:-, \gamma:-, \beta:-$	$\epsilon:-, \gamma:-, \beta:-$	$\epsilon:-, \gamma:-, \beta:-$	387
Δ promoteur β	PAC (185 kb)	$\epsilon:-, \gamma:-$	$\gamma:\uparrow, \beta:\emptyset$	$\delta:\uparrow, \beta:\emptyset$	463
$\Delta \beta$	YAC (248 kb)	$\epsilon:\uparrow, \gamma:\uparrow, \delta:\uparrow$	$\epsilon:-, \gamma:\uparrow, \delta:\uparrow$	$\gamma:\uparrow, \delta:\uparrow$	507
Δ enhancer en 3' de β (Δ 250 pb)	YAC (150 kb)	$\epsilon:-, \gamma:-, \beta:-$	$\epsilon:-, \gamma:-, \beta:\downarrow$	$\epsilon:-, \gamma:-, \beta:\downarrow$	486

Légende: Expression relative au locus sauvage: voir la légende du Tableau II. ϵ^{Bepsi} : gène ϵ dont le promoteur possède les boîtes CAC et CAAT de β . Δ : délétion. λ : séquence neutre provenant du phage λ . F: élément F situé en 3' de l'enhancer de $\Delta\gamma$. PYR: région riche en pyrimidines.

Tableau IV: Transgéniques avec le locus humain modifié au niveau du LCR.

Mutations	Vecteur	Expression			Ref
		embryon	Fœtus	Adulte	
Δ HS1 (0,8 kb)	YAC (213 kb)	$\epsilon:\downarrow, \gamma:-$	$\gamma:\downarrow, \beta:-$	$\beta:-$	349
λ (2,3 kb) ajouté après HS1	YAC (150 kb)	$\epsilon:\downarrow, \gamma:-$	$\gamma:-, \beta:-$	$\beta:-$	382
Δ HS2 (1,9 kb)	YAC (248 kb)	$\epsilon:\downarrow, \gamma:\downarrow$	$\gamma:\downarrow, \beta:\downarrow$	$\beta:\downarrow$	358
Δ HS2 (1,2 kb)	PAC (185 kb)	$\epsilon:\downarrow, \gamma:\downarrow$	$\gamma:-, \beta:\downarrow$	$\beta:\downarrow$	463
Δ HS2 (<i>core</i> , 375 pb)	PAC (185 kb)	$\epsilon:\downarrow, \gamma:\downarrow$	$\gamma:\downarrow, \beta:\downarrow$	$\gamma:\downarrow, \beta:\downarrow$	125
Substitution HS2(<i>core</i>) \rightarrow HS3(<i>core</i>)	YAC (150 kb)	$\epsilon:\downarrow, \gamma:\downarrow$	$\gamma:\downarrow, \beta:\downarrow$	$\gamma:\downarrow, \beta:\downarrow$	125
Δ HS3 (2,3 kb)	YAC (248 kb)	$\epsilon:\downarrow, \gamma:\uparrow$	$\epsilon:\downarrow, \gamma:-, \beta:-$	$\beta:-$	358
Δ HS3 (1,5 kb)	PAC (185 kb)	$\epsilon:\downarrow, \gamma:\downarrow$	$\gamma:\downarrow, \beta:\downarrow$	$\beta:\downarrow$	463
Δ HS3(<i>core</i>)	YAC (150 kb)	$\epsilon:\downarrow, \gamma:\downarrow$	$\gamma:\downarrow, \beta:\downarrow$	$\beta:\downarrow$	370
Δ HS3(<i>core</i>)	YAC (248 kb)	$\epsilon:\emptyset, \gamma:\downarrow$	$\gamma:\emptyset, \beta:\downarrow$	$\beta:\downarrow$	508
Δ HS3 (1,5 kb), Δ promoteur β	PAC (185 kb)	$\epsilon:\downarrow, \gamma:\downarrow$	$\gamma:\emptyset, \beta:\emptyset$	$\delta:\emptyset, \beta:\emptyset$	463
HS3 motif GT6 (TGGA \rightarrow ATTC)	YAC (248 kb)	$\epsilon:\downarrow, \gamma:\downarrow$	$\gamma:\downarrow, \beta:\downarrow$	$\beta:-$	127
Substitution HS3(<i>core</i>) \rightarrow HS4(<i>core</i>)	YAC (150 kb)	$\epsilon:\downarrow, \gamma:\uparrow$	$\gamma:\downarrow, \beta:\downarrow$	$\beta:\downarrow$	370
Δ HS4 (2,9 kb)	YAC (213 kb)	$\epsilon:-, \gamma:-$	$\gamma:\uparrow, \beta:\downarrow$	$\beta:-$	349
Δ HS4 (<i>core</i>)	YAC (150 kb)	$\epsilon:\downarrow, \gamma:\downarrow$	$\gamma:\downarrow, \beta:\downarrow$	$\beta:\downarrow$	370
Δ HS4 (Δ 280 pb)	YAC (248 kb)	$\epsilon:\downarrow, \gamma:\uparrow$	$\gamma:\downarrow, \beta:\downarrow$	$\beta:\downarrow$	128
Substitution HS4(<i>core</i>) \rightarrow HS3(<i>core</i>)	YAC (150 kb)	$\epsilon:-, \gamma:-$	$\gamma:-, \beta:-$	$\beta:-$	370
Substitution HS234(<i>cores</i>) \rightarrow PyE	YAC (150 kb)	$\epsilon:\emptyset, \gamma:\emptyset$	$\gamma:\emptyset, \beta:\emptyset$	$\beta:\emptyset$	509
Δ LCR	YAC (248 kb)	$\epsilon:\emptyset, \gamma:\emptyset$	$\gamma:\emptyset, \beta:\emptyset$	$\beta:\emptyset$	126

Légende: Expression relative au locus sauvage: voir la légende du Tableau II. Δ : délétion. λ : séquence neutre provenant du phage λ . *Core*: partie centrale des HS. PyE: *enhancer* du virus du polyome.

Tableau V: Transgéniques avec le locus humain modifié au niveau de son organisation.

Mutations	Vecteur	Expression			Ref
		embryon	Fœtus	Adulte	
Substitution pr.β→pr.γ	YAC (150 kb)	ε:↓, γ:-, β:↓	γ:↑, β:↓	γ:-, β:↓	259
β repositionné devant ε	YAC (258 kb)	ε:↓, γ:↓, β:↑	γ:↓, δ:↓, β:↑	β:-	507
β ^m ajouté devant ε	YAC (213 kb)	ε:∅, γ:↓, β ^m :↑	γ:↓, β ^m :↑	β ^m :-	258
^A γ ^m ajouté devant ε	YAC (213 kb)	ε:∅, γ:↓, γ ^m :↑	γ:↓, γ ^m :-, β:-	γ ^m :- β:↓	258
β ^m ajouté devant ^G γ	YAC (213 kb)	ε:↑, γ:↓, β ^m :↑	γ:↓, β ^m :↓, β:↓	β ^m : ↓, β:↓	258
ε ^m ajouté devant β	YAC (213 kb)	ε:-, γ:-	γ:-, β:-	β:-	258
HS5 (2,6 kb) ajouté après HS1	YAC (150 kb)	ε:↓, γ:-	γ:↓, β:↓	β:↓	378;382
HS5(ΔCTCF; 2,6 kb) ajouté après HS1	YAC (150 kb)	ε:↓, γ:-	γ:-, β:-	β:-	378;382
Locus inversé par rapport au LCR	YAC (150 kb)	ε:∅, γ:↓, β:↑	γ:∅, β:-	β:-	343;503
LCR inversé par rapport au locus	YAC (150 kb)	ε:↓, γ:↓	γ:↓, β:↓	β:↓	343

Légende: Expression relative au locus sauvage: voir la légende du Tableau II. Δ: délétion. ΔCTCF: délétion du site de liaison du facteur CTCF. Pour les gènes marqués β^m et ^Aγ^m, l'expression est comparée aux gènes sauvages β et ^Aγ respectivement.

Tableau VI: Transgéniques avec le locus humain combinant des modifications.

Mutations	Vecteur	Expression			Ref
		embryon	Fœtus	Adulte	
Substitution pr.β→pr.γ, Δ ^G γ ^A γ	YAC (150 kb)	ε:↑, β:↑	β:↓	β:↓	259
λ (2,3 kb) ajouté après HS1, ε ^{Bepsi}	YAC (150 kb)	ε:↑, γ:↓			382
Δ HS2 (1,2 kb), Δ promoteur β	PAC (185 kb)	ε:–, γ:–	γ:↑, β:∅	δ:↑, β:∅	463
Δ HS3(<i>core</i>), -117 ^A γ (G→A)	YAC (248 kb)	ε:∅, γ:–	γ:↓, β:↓	β:↓	126
Δ HS3 (1,5 kb), Δ promoteur β	PAC (185 kb)	ε:↓, γ:↓	γ:∅, β:∅	δ:∅, β:∅	463
Locus inversé, ε ^{Bepsi}	YAC (150 kb)	ε:↓, γ:–, β:↑			382

Légende: Expression relative au locus sauvage: voir la légende du Tableau II. ε^{Bepsi}: gène ε dont le promoteur possède les boîtes CAC et CAAT de β. Δ: délétion. λ: séquence neutre provenant du phage λ. *Core*: partie centrale des HS.

Le promoteur de la ^Aγ-globine

Compte tenu du grand nombre de mutations identifiées dans le promoteur des gènes γ qui provoquent une persistance héréditaire de l'hémoglobine fœtale chez l'adulte (Tableau I, page 25), ces promoteurs sont soupçonnés depuis longtemps de jouer un rôle central dans la régulation de l'hémoglobine fœtale à adulte. De fait, de tous les promoteurs de globine, celui de ^Aγ est certainement celui qui a été le plus étudié. Ceci a permis de confirmer une implication certaine de ce dernier dans la régulation de l'hémoglobine et son importance dans la compétition pour le LCR.^{96;257;262;293;294;452;510-514} Le modèle courant étant que l'activation des gènes fœtaux empêche l'activation des gènes adultes et c'est leur auto-inactivation qui détermine la régulation temporelle de la commutation de l'hémoglobine fœtale. Toutefois, si cette fonction des gènes γ est due à leur activité transcriptionnelle, ou

encore à certains éléments spécifiques de leur promoteur et qui auraient intrinsèquement cette fonction, demeure incertain. Deux études ont tenté de répondre à cette question en remplaçant le promoteur fœtal-spécifique de γ par un promoteur stade-indépendant mais ont cependant obtenu des résultats contradictoires.

Le promoteur de la thymidine kinase

Il avait été démontré déjà qu'une mini construction comprenant le LCR ou simplement le site HS2, lié directement au gène β (LCR- β ou HS2- β) faisait en sorte que ce dernier était exprimé à tous les stades du développement et que lorsqu'un gène γ était inséré entre les deux (LCR- γ - β et HS2- γ - β), la régulation de β semblait être rétablie.^{103;260;515} Pour déterminer si cet effet était dû à l'activité transcriptionnelle du gène γ ou au gène lui-même, Anderson *et al* ont utilisé deux constructions. La première était la mini construction HS2- γ - β dont le promoteur de γ avait été excisé (HS2- $\gamma_{\Delta P}$ - β) alors que dans la seconde, le gène γ et son promoteur ont été remplacés par le gène de la chloramphénicol acétyl transférase et le promoteur de la thymidine kinase (HS2-TKCAT- β). Dans le premier cas, l'expression de β était dérégulée de façon constitutive, c'est-à-dire que le gène β était exprimé à tous les stades du développement, ce qui indiquait que le promoteur de γ ou l'activité transcriptionnelle du gène γ était importante. La construction avec le promoteur de la thymidine kinase permit toutefois de rétablir l'expression adulte-spécifique du gène β et les auteurs en conclurent que la transcription d'un gène situé entre le site HS2 et β "protégeait" d'une façon ou d'une autre ce dernier contre une activation inappropriée au stade embryonnaire. En d'autres mots, l'inactivation de β au stade embryonnaire nécessitait la présence d'un promoteur actif accessoire entre lui et le LCR et que la nature de ce promoteur antagoniste n'était pas importante.⁵¹⁶

Le promoteur de la β -spectrine

En 1998, une étude similaire effectuée par Sabatino *et al* utilisa cette fois un mini locus comprenant un micro-LCR lié à un fragment du locus contenant γ , $\psi\beta$, δ et β ($\mu\text{LCR}-\gamma\psi\beta\delta\beta$) inclus dans un cosmide et présentant une expression temporellement adéquate des différents gènes du locus.⁵¹⁵ Le promoteur de γ a été remplacé dans cette construction par celui de la β -spectrine. Ce promoteur érythroïde-spécifique fort était capable de conférer une expression élevée au gène γ , même en l'absence du LCR. Cette fois, contrairement à ce qui avait été observé par Anderson *et al*, le gène β était exprimé à tous les stades du développement, tout comme l'était γ ce qui indiquait qu'un gène actif entre β était insuffisant à l'inactivation de β au stade embryonnaire et que l'expression simultanée des deux gènes n'était pas antagoniste.⁵¹²

1.7 Hypothèse et objectifs de la thèse

Si plusieurs études ont démontré que les gènes fœtaux γ et γ étaient surtout régulés de façon autonome, il n'est toujours pas clair si leurs promoteurs ont une fonction intrinsèque importante dans la régulation générale de la commutation de l'hémoglobine fœtale à adulte, entre autre en permettant l'inactivation du gène β au stade fœtal. Il est certes possible que cette fonction ne soit due qu'à l'activité transcriptionnelle elle-même des gènes fœtaux. Les modèles actuels ne permettent pas de déterminer avec certitude le rôle des gènes fœtaux dans la régulation de la commutation mais supportent l'idée en fait que les promoteurs fœtaux participeraient via certains de leurs éléments, à cette régulation.

Suivant les résultats contradictoires des deux études précédentes ayant utilisé les promoteurs de la β -spectrine et de la thymidine kinase, l'objectif de cette thèse était de déterminer la fonction du promoteur fœtal γ dans la régulation de la commutation de l'hémoglobine fœtale à adulte et ce, dans le contexte du locus humain entier contenu dans un YAC de 150 kb en souris transgénique. Cette étude se proposait ainsi dans un premier temps de répondre à la question à savoir si le promoteur γ possède une fonction intrinsèque permettant l'inactivation par compétition avec les gènes adultes ou si cette inactivation est

le résultat obligatoire d'une activation LCR-dépendante d'un gène situé entre β et le LCR. Dans un deuxième temps, l'identification de la ou des régions du promoteur foetal, responsables de cette fonction de régulation, a également fait l'objet de cette étude.

De façon à déterminer l'importance du promoteur de γ -globine dans la commutation de l'hémoglobine, j'ai choisi une approche consistant à remplacer spécifiquement ce promoteur par la région équivalente d'un promoteur érythroïde-spécifique mais stade indépendant. Le choix du promoteur de remplacement devait être fait avec rigueur de façon à ce que cette expérience puisse répondre à la question posée. De fait, le promoteur choisi devait rencontrer certaines conditions essentielles :

- La taille de la région de remplacement devait être équivalente à la région remplacée de façon à éviter un impact sur l'expression des gènes dû à une modification de la structure nucléosomique du promoteur foetal.
- le promoteur devait être érythroïde-spécifique pour assurer que son profil d'expression soit comparable à celui des gènes de globine.
- son expression devait être stable à tous les stades du développement, cette caractéristique étant nécessaire pour permettre l'évaluation de l'importance de l'activité transcriptionnelle dans la compétition pour le LCR.
- L'activité basale du promoteur devait être comparable à celle des promoteurs de globine de façon à éviter l'insertion d'éléments activateurs très puissants dans le locus pouvant interférer avec l'activité des gènes environnants par des mécanismes étrangers au locus.
- Le promoteur devait idéalement être inductible par le LCR des gènes de globine puisque c'est cette dépendance au LCR qui est au cœur du débat sur la fonction des gènes foetaux dans la commutation de l'hémoglobine foetale à adulte.

Ces critères très sévères ont rendu le choix du promoteur assez difficile puisqu'en fait, le promoteur idéal aurait été un promoteur de globine stade-indépendant ce qui n'existe pas. Notre choix s'est arrêté sur le promoteur de la porphobilinogène désaminase ou PBGD. Ce gène qui code pour une enzyme impliquée dans la synthèse de l'hème possède deux promoteurs : l'un est ubiquitaire alors que l'autre est érythroïde-spécifique et rencontre tous

les critères établis pour le choix du promoteur de remplacement. En effet, le promoteur érythroïde de PBGD est un promoteur faible dont le niveau basal d'expression est légèrement inférieurs à celui des promoteurs de globine et qui n'est exprimé que dans les cellules érythroïdes.^{517;518} De plus, le promoteur PBGD, lorsque lié au HS2 du LCR du locus de la β -globine montre une activation de 20 fois par rapport à son niveau basal, indiquant que ce promoteur peut être activé de façon LCR-dépendante.⁵¹⁹

Structurellement, le promoteur de PBGD ressemble à celui des gènes de globine bien qu'il soit beaucoup plus simple (Figure 21). Ce promoteur possède de fait des éléments de liaison à des facteurs activateurs communs aux gènes de globine comme GATA1, EKLF, FKL2, Sp1, NF-E2 et AP1^{517;520-523} mais se caractérise toutefois par l'absence d'éléments répresseurs et de la boîte TATA qui est remplacée par un initiateur.⁵²⁴

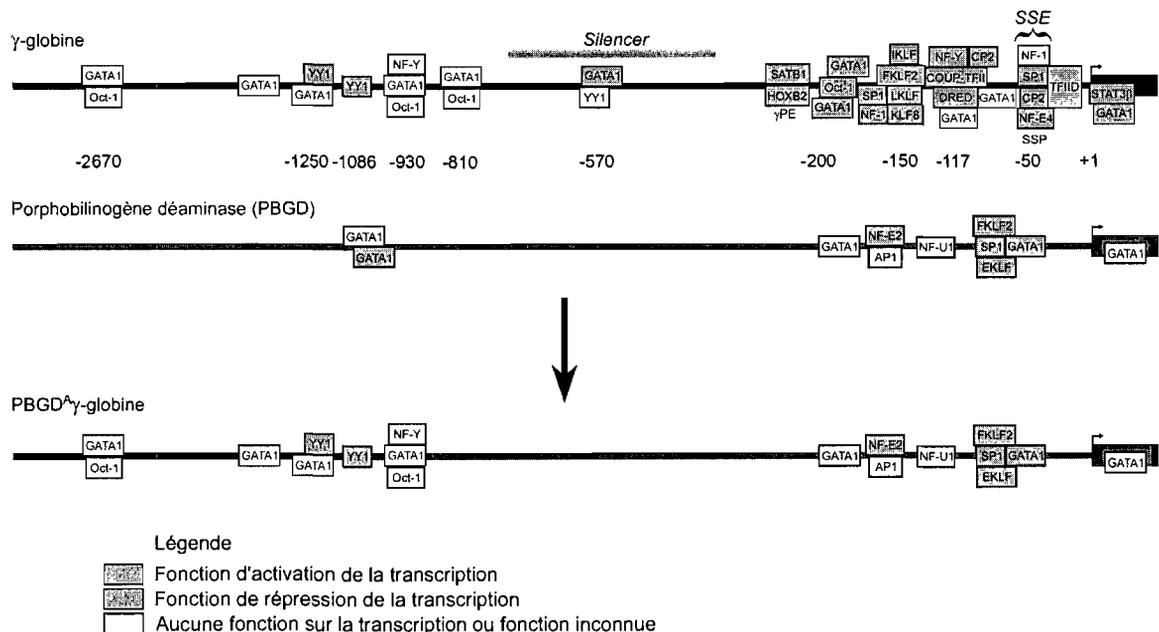


Figure 21: Les promoteurs γ -globine, PBGD et l'hybride PBGD^A γ -globine.

Ainsi, le remplacement du promoteur de γ par PBGD produit un promoteur hybride dont les éléments répresseurs ont été éliminés et remplacés par des éléments à fonction strictement activatrice, entraînant une expression continue de ce gène tout au long du développement et ce, de façon dépendante du LCR (Figure 21). Je me proposais ainsi de mesurer l'expression des différents gènes du locus au cours du développement pour évaluer l'impact de l'activation du gène fœtal à tous les stades du développement.

Pour répondre au deuxième objectif de cette thèse, certaines séquences ont été excisées du promoteur de la γ -globin et ce, dans le contexte du locus humain entier compris dans un YAC en souris transgénique (Figure 22).

Aussi, je me proposais d'exciser la séquence -841 à -384 éliminant de ce fait la région à fonction *Silencer* du gène γ et caractérisée par la présence d'un site GATA-1 important pour cette activité. Une seconde délétion couvrant la région -384 à -202 est également proposée, qui élimine l'élément de liaison au complexe γ PE dont la fonction dans la commutation n'est pas très bien connue. Finalement, une troisième délétion, plus courte cette fois, couvrant la région -204 à -135 permettant d'éliminer une région riche en sites de liaison à des facteurs de transcription tant positifs que négatifs, vient clore cette étude.

Une analyse d'expression des différents gènes du locus chez ces mutants à tous les stades du développement, devait ainsi me permettre d'évaluer l'impact de chacune de ces délétions sur la commutation fœtale à adulte et d'identifier celles qui sont importantes pour la régulation de cette commutation.

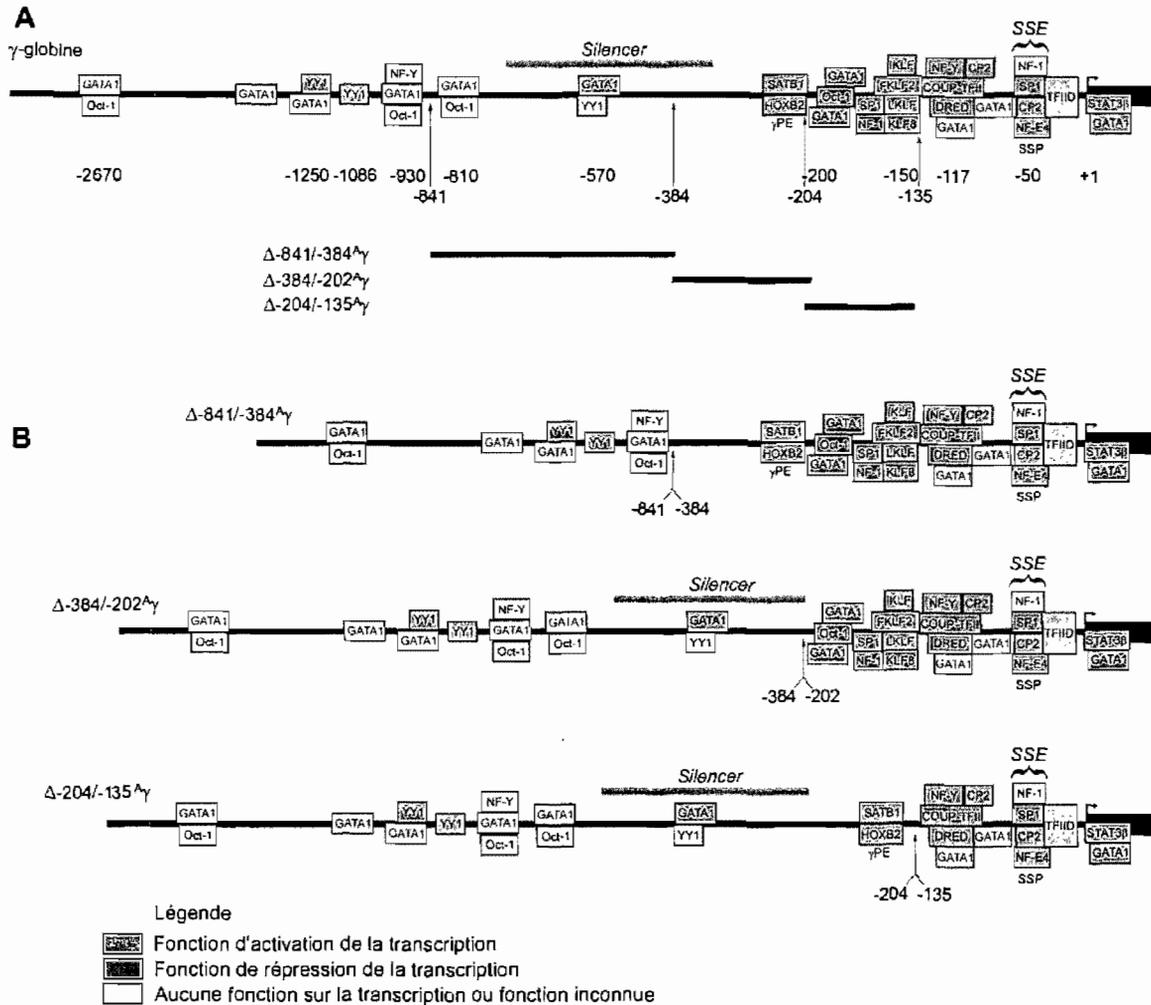


Figure 22: Régions excisées du promoteur de la γ -globine. A) Le promoteur de γ et les trois délétions proposées. B) Structure du promoteur résultant de chacune de ces trois délétions.

Chapitre 2

Matériel et méthodes

2.1 Clonage moléculaire

2.1.1 Préparation des fragments d'ADN à partir de plasmide

Digestion par endonucléases de restriction

L'ADN purifié (10-20 μg) est digéré dans le tampon approprié à l'enzyme utilisée en accord avec les recommandations des fabricants (New England Biolab, Amersham Biosciences, Invitrogen Life Technology). Lorsque plusieurs enzymes sont utilisées, les digestions sont effectuées simultanément ou séquentiellement selon la compatibilité entre les enzymes et leur tampon. Une étape de purification de l'ADN au phénol/chloroforme est effectuée entre chaque digestion séquentielle (tel que décrit ci-dessous).

Extraction de l'ADN de plasmide au phénol/chloroforme

L'extraction de l'ADN de plasmide d'une solution aqueuse se fait dans 1 volume de phénol/chloroforme/alcool isoamyle (25:24:1) équilibré avec du Tris-HCl (100 mM, pH 8,0) suivie d'une précipitation avec 2 volumes d'éthanol (100%) et 0,1 volume d'acétate de sodium (3M, pH 5,2) pendant 1 h à -80°C et d'un lavage avec 1 volume d'éthanol (70%). Le culot d'ADN est séché par centrifugation sous vide avant d'être resuspendu dans le tampon d'intérêt.

Génération de bouts francs

Pour modifier les extrémités cohésives des fragments en bouts francs pour la ligation, deux techniques ont été utilisées en fonction de la nature du brin cohésif et des besoins.

Remplissage d'un brin sortant 5' et déplétion d'un brin sortant 3'

Pour la formation d'une extrémité franche par remplissage d'une extrémité sortante 5' ou par déplétion d'une extrémité sortante 3', l'ADN est incubé avec 1 U de fragment de Klenow de la DNA polymérase I (Invitrogen) par μg d'ADN dans le tampon REact 2 (Invitrogen) en présence des quatre dNTP (33 μM chaque) pendant 30 min à 30°C. La réaction est arrêtée en incubant 10 min à 70°C et l'ADN est extrait au phénol/chloroforme.

Déplétion d'un brin sortant 5'

Pour la formation d'un bout franc par déplétion d'une extrémité sortante 5', l'ADN est resuspendu et incubé avec 1 U d'endonucléase Mung Bean (Amersham Bioscience) par μg d'ADN dans le tampon de digestion [glycérol (5%), acétate de sodium (30 mM), NaCl (100 mM), ZnOAc (1 mM), Triton X-100 (0,001%), Cystéine (1 mM)] pendant 15 min à 37°C. La réaction est arrêtée en ajoutant 25 mM d'EDTA (0,5 M) puis l'ADN est extrait au phénol/chloroforme.

Déphosphorylation des vecteurs

Les vecteurs digérés possédant des extrémités franches ou cohésives compatibles ont été déphosphorylés pour prévenir la recircularisation lors de la ligation. L'ADN est incubé avec 1 U de phosphatase alcaline intestinale de veau (Promega) par μg d'ADN dans le tampon NEB3 (New England Biolab) pendant 30 min à 37°C. Ensuite, 1 U de phosphatase alcaline supplémentaire est ajoutée et l'incubation est poursuivie pendant 30 min. La réaction est arrêtée par une extraction phénol/chloroforme tel que décrit ci-dessus.

Purification des fragments

Les fragments sont isolés sur un gel d'agarose à 0,8-1,2% dans le TAE [(Tris-acétate (40 mM, pH 7,9), EDTA (1 mM))] coloré avec 0,2 mg/L de bromure d'éthidium. Une bande d'agarose contenant le fragment d'intérêt identifié par visualisation sous la lumière UV-B ($\lambda=305$ nm) est excisée du gel et électroéluée dans un sac à dialyse dans du TAE sans bromure d'éthidium pendant 60 min à 100 V. Après une centrifugation pour culotter les

déchets d'agarose, l'ADN du surnageant est récupéré par extraction au phénol/ chloroforme tel que décrit ci-dessus et quantifié par comparaison d'intensité sur un gel d'agarose avec l'ADN de phage λ digéré par HindIII.

2.1.2 Préparation du fragment PBGD à partir de l'ADN génomique

Extraction de l'ADN génomique

L'ADN génomique humain a été obtenu de cellules musculaires lisses de la veine ombilicale en lysant les cellules HUVSM (*Human Umbilical Vein Smooth Muscle*) à confluence dans le tampon de protéolyse [Tris-HCl (50 mM, pH 8,0), EDTA (100 mM), SDS (0,5%), pronase (0,5 mg/mL)] 16 h à 55°C. L'ADN est extrait une première fois avec 1 volume de phénol équilibré avec du Tris-HCl (100 mM, pH 8,0) et une seconde fois avec 1 volume de phénol/chloroforme/alcool isoamyle (25:24:1) équilibré avec du Tris-HCl (100 mM, pH 8,0). L'ADN est précipité avec 1 volume d'éthanol (100%) et 0,1 volume d'acétate de sodium (3 M, pH 6,0) et le culot lavé avec 200 μ L d'éthanol (70%) pour ensuite être séché sous vide quelques secondes. L'ADN est resuspendu dans du tampon TE [Tris (10 mM, pH 8,0), EDTA (1 mM)].

Amplification par réaction de polymérisation en chaîne (PCR)

Le fragment d'ADN contenant le promoteur de la porphobilinogène désaminase (PBGD) a été obtenu d'une amplification par PCR niché au moyen d'un appareil *Gene Amp PCR System 9600* (Perkin Elmer). La méthode correspond à 5 min de dénaturation à 94°C suivit de 28 cycles d'amplification (94°C-50 s, 68°C-50 s, 72°C-50 s) dans une réaction de 20 μ L [Tris-HCl (10 mM, pH 8,0), KCl (50 mM), MgCl₂ (1 mM), NH₄Cl (5 mM), amorces 00-25/00-26 (1 μ M chaque), dNTP (125 μ M), Taq polymérase (0,14 U/ μ L)] voir à Annexe I, Tableau X, page II pour la liste des amorces. Un μ L de ce produit de PCR est amplifié dans un second tour de 27 cycles aux mêmes conditions que le premier tour mais avec les amorces 99-15 (0.6 μ M) et 99-16 (1 μ M). Le nouvel amplicon de cette réaction (941 pb)

possède un site de restriction AvrII introduit à l'aide de l'amorce 99-15. Ce produit de PCR est digéré par des endonucléases de restriction et purifié tel que décrit à la section 2.1.1.

2.1.3 Ligation

La ligation de fragments à extrémités franches est effectuée en utilisant un ratio molaire vecteur:insert de 1:10 dans une réaction de 5-10 μL [Tris-HCl (50mM, pH 7,6), MgCl_2 (10mM), ATP (1 mM), DTT (1mM), Polyéthylène glycole-8000 (5 mM), T4 ADN ligase (Invitrogen, 0,1 U/ μL)]. La réaction est incubée au moins 16 h à 4°C. Les ligations de fragments à extrémités cohésives ont été effectuées avec un ratio molaire vecteur:insert de 1:5 à une température de 16°C pendant 4-16 h.

2.1.4 Transformation

Des cellules compétentes *Escherichia coli* DH5 α ou XL1 *blue* sont transformées avec 100-200 ng de plasmide ou de cosmide. Les cellules (200 μL) sont incubées avec 100 μL de Tris-HCl (100 mM, pH 8,0) et 1-10 μL de la ligation pendant 30 min sur glace. L'ADN est incorporé dans les cellules par la méthode du choc thermique (5 min à 37°C). On ajoute 500 μL de milieu LB [tryptone (10 g/L), extrait de levure (5 g/L), NaCl (10g/L)] et la culture est incubée avec agitation pendant 1 h à 37°C. 3 mL de LB agar mou chauffé à 50°C sont ajoutés à la culture et immédiatement étalés sur un pétri LB-agar/ampicilline (50 $\mu\text{g/ml}$) supplémenté de X-gal (50 $\mu\text{g/ml}$) lorsqu'une sélection à la β -galactosidase est possible. Les pétris sont incubés pendant 16 h à 37°C.

2.1.5 Criblage de clones positifs

Sélection à la β -galactosidase

Les clones obtenus lors de la transformation sur pétri LB-agar/ampicilline/X-gal sont sélectionnés sur la base de la couleur des colonies lorsque le fragment est inséré dans un vecteur contenant le gène de sélection *LacZ* et ne possédant pas d'insert déjà intégré à

son site de polyclonage. Dans ce cas, seuls les clones blancs sont sélectionnés alors que les clones bleus sont éliminés. Les clones sélectionnés sont mis en culture dans 1 mL de milieu LB et incubé pendant 16 h avec agitation à 37°C.

Hybridation de colonies

Une approche de criblage des clones à l'aide de sonde radioactive est utilisée pour identifier les clones ayant intégré l'insert lorsque la sélection à la β -galactosidase n'est pas possible.

Transfert de plages

Les clones sont repiqués et striés en duplicata sur deux pétris LB-agar/Ampicilline et incubés pendant 16 h à 37 °C. Une membrane Hybond-N (Amersham Biosciences) est déposée brièvement sur l'un des duplicata. L'ADN est subséquemment dénaturé 5 min dans une solution de NaOH (0,5 M) et neutralisée 5 min dans une solution de neutralisation [Tris (0,5 M, pH 7,5), NaCl (1,25 M)]. Après deux lavages de 2 min dans la solution de neutralisation, la membrane est séchée et l'ADN fixé à la membrane de façon permanente par exposition au rayonnement UV (1200 μ J) dans un appareil *UV Stratalinker 2400* (Stratagene).

Radiomarquage de la sonde par translation de coupure

L'ADN (100 ng) est incubé dans un tampon de translation [Tris-HCl (50 mM, pH 7,5), MgCl₂ (5 mM), β -mercaptoéthanol (10 mM), BSA (100 μ g/mL), dATGTP (10 μ M de chaque), α^{32} P-dCTP (1,25 μ Ci/ μ L), ADN polymérase I (USB Corporation, 0,5 U/ μ L), DNaseI (100 pg/ μ L)] pendant 45-60 min à 16°C. La réaction est arrêtée avec 5 volumes de la solution d'arrêt [Tris (10 mM, pH 7,5), EDTA (12,5 mM), SDS (0,5%)]. La sonde est dénaturée en la menant à ébullition pendant 5 min puis déposée sur la glace.

Hybridation et révélation autographique

La membrane est incubée dans une solution d'hybridation [Ficoll (2%), polyvinylpyrrolidone (2%), BSA (2%), NaCl (0,75 M), citrate de sodium (75 mM), ADN de sperme de saumon (ADNss)(0,1 mg/mL), SDS 0,1%] pendant 4 h à 65 °C. Une sonde spécifique à l'insert recherché, obtenue à l'aide d'un marquage par déplacement de brèche, est ajoutée et l'incubation à 65°C est continuée 16 h. Les membranes sont lavées successivement dans les solutions de lavage I [SDS (0,1%), NaCl (0,3 M), citrate de sodium (30 mM)] pendant 1 h, II [SDS (0,1%), NaCl (0,15 M), citrate de sodium (15 mM)] pendant 20 min et III [SDS (0,1%), NaCl (15 mM), citrate de sodium (1,5 mM)] pendant 5 min. La membrane est ensuite exposée sur un film d'autoradiographie *X-Omat blue XB-1* (Kodak) à -80°C pendant 4-72 h dépendamment de l'intensité du signal.

2.1.6 Préparation de plasmide

Mini-préparation de plasmide

Les cultures bactériennes sont centrifugées 1 min à vitesse maximale et le culot de cellules est resuspendu dans 100 µL de solution de lyse [Tris-HCl (25 mM, pH 8,0), EDTA (10 mM), glucose (50 mM), lysosyme (2 mg/mL)] puis incubé pour 30 min à 25°C. 200 µL de solution alcaline [NaOH (200 mM), SDS (1%)] sont ajoutés et les tubes sont incubés 5 min sur glace puis la réaction est neutralisée avec 150 µL d'acétate de sodium (3 M, pH 4,8) et laissés sur glace pour 30 min pour permettre la précipitation des fragments cellulaires. Les tubes sont centrifugés 5 min à vitesse maximale et l'ADN du surnageant est précipité avec 2 volumes d'éthanol (100%) et 0,1 volume d'acétate de sodium (3M, pH 4,8) pendant 1 h à -80°C et d'un lavage avec 1 volume d'éthanol (70%) puis l'ADN est resuspendu dans 100 µL de TE.

Grande préparation de plasmide par gradient de densité

Le clone est repiqué et mis en culture avec agitation dans 500 mL de milieu LB pendant 16 h. La culture est centrifugée à 2500×g et le culot resuspendu puis incubé 10 min

dans une solution de sucrose/lysosyme [sucrose (25%), Tris-HCl (50 mM, pH 8,0), lysosyme (1,7 mg/mL)]. La réaction est arrêtée avec 70 mM d'EDTA (pH 8,5) puis la lyse est complétée à l'aide d'une solution de Triton [Tris-HCl (50 mM, pH 8,0), EDTA (62,5 mM), Triton X-100 (0,1%)]. Le lysat est centrifugé 1h à 80 000×g puis le surnageant est déposé sur un coussin de TE saturé en CsCl et centrifugé 16 h à 80 000×g. Le surnageant est ajoutée à du CsCl/bromure d'éthidium puis centrifugé 24 h à 250 000×g. La bande du plasmide superenroulé est ensuite récupérée à l'aide d'une seringue, extrait à plusieurs reprises avec un volume égal d'isopropanol, précipité à l'éthanol/acétate de sodium, puis lavé à l'éthanol 70%. L'ADN est resuspendu dans 100-500 µL de TE.

Grande préparation de cosmide

Le clone est repiqué, mis en culture dans 800 mL de milieu LB et incubé avec agitation pendant 16 h. Le culot est resuspendu dans une solution de lyse [Tris-HCl (8 mM, pH 8,0), EDTA (5 mM), NaOH (0,3 M), SDS (0,7 %)] pendant 3 min sur glace et les débris cellulaires sont précipités par centrifugation à 20 000×g avec 1 M d'acétate de potassium. L'ADN est purifié par extraction avec 1 volume de phénol:chloroforme:alcool isoamyle 25:24:1, précipité avec 0,5 volume d'isopropanol, lavé à l'éthanol (70%) puis resuspendu dans 400 µL de TE.

2.1.7 Caractérisation des clones

Analyse de restriction

Les clones obtenus sont caractérisés à l'aide de leur patron de restriction. L'ADN (100 ng) est digéré à l'aide d'endonucléases tel que décrit à la section 2.1.1 et vérifié par électrophorèse sur gel d'agarose.

Séquençage

Les clones dont le patron de restriction s'avère correct, sont séquencés. L'ADN obtenu d'une mini-préparation est surpurifié en utilisant le kit QIAprep[®] Spin Miniprep

(Qiagen) selon le protocole du fabricant puis resuspendu dans du Tris (10 mM, pH 8,0). L'ADN (375 fmoles) ainsi que l'amorce (16 pmoles) sont fournis au service de séquençage du plateau de biologie moléculaire pour être séquencés. La liste des amorces utilisées pour les séquençages est donnée au Tableau XII (Annexe I, page II). La séquence obtenue est visualisée à l'aide du logiciel Chromas v1.45.

2.1.8 Mutagenèse dirigée

La génération d'une mutation spécifique dans un plasmide est effectuée par PCR à haute fidélité. Le plasmide (300 pg) est amplifié avec la polymérase à ADN *Pfx Platinum*[®] (Invitrogen) dans une réaction de 50 µL [tampon d'amplification *Pfx* (2,5 X), solution *Enhancer* (0,25 X), MgCl₂ (1 mM), dNTP (300 µM de chaque), amorces (0,3 µM de chaque), *Pfx Platinum*[®] polymérase (0,05 U/µL)]. Après une étape de dénaturation de 5 min à 94°C, l'amplification se fait en un tour comprenant 10 premiers cycles (94°C-20 s, 65°C-30 s, 68°C-8 min) suivi de 18 cycles subséquents (94°C-20 s, 65°C-30 s, 68°C-8 min + 20 s par cycle). La mutation est introduite à l'aide d'une des amorces dont l'extrémité 5' correspond à l'extrémité 5' moins 1 nt de l'amorce antisens (voir la liste des amorces utilisées à l'Annexe I, Tableau XI, page II). L'amplicon est purifié à l'aide d'un kit QIAquick[®] PCR purification (Qiagen) et resuspendu dans 50 µL de tris (pH 8,0). L'amplicon (2 µL) est circularisé par une réaction de ligation en présence de la polynucléotide kinase [Tris-HCl (50mM, pH7,6), MgCl₂ (10mM), ATP (1 mM), DTT (1mM), polyéthylène glycole-8000 (5 mM), T4 ADN ligase (0,1 U/µL), T4 PNK (New England Biolab, 0,1 U/µL)] pendant 5-24 h à 37°C. L'ADN est transformé et séquencé tel que décrit aux sections 2.1.4 et 2.1.7 respectivement.

2.2 Génération des plasmides d'intégration dans la levure

La première étape du processus devant mené à la génération de YAC mutants par recombinaison homologue dans la levure consiste à générer les plasmides d'intégration dans la levure ou YIp (pour *yeast integrating plasmid*). Ces YIp sont constitués d'un gène de

prototrophie à l'uracile et de deux bras d'homologie avec le YAC correspondant aux séquences flanquant la région ciblée et servant à la recombinaison homologue. Ces deux bras peuvent être séparés par une séquence de remplacement (comme le promoteur PBGD) ou encore contiguës ce qui produira une délétion dans le YAC, correspondant à la région ciblée. Ce vecteur, dérivé du plasmide pBluescript, comporte également une origine de répllication bactérienne non-fonctionnelle en levure, un gène de résistance à l'ampicilline et le gène de la β -galactosidase (Voir la carte à Annexe IV, page XIX). Les bras d'homologie entourant la séquence de remplacement ou la délétion doivent faire plus de 1 kb pour favoriser les événements de recombinaison homologue.

2.2.1 Le YIp pRS-PBGD^{A γ}

Le YIp pRS-PBGD^{A γ} a été produit par l'insertion séquentielle de trois fragments (Figure 23A). Le premier élément inséré fut le fragment de 2,6 kb MscI-SapI couvrant la région -3423 à -838 en amont du promoteur de ^{A γ} , isolé du plasmide pVZ-ES (donné généreusement par le Dr. Richard Gelinas). Ce fragment a été inséré dans le site NotI rendu franc par remplissage de l'extrémité sortante 5'. Le deuxième élément inséré dans la construction a été un fragment de 887 pb du promoteur de PBGD de -815 à +72. Ce fragment a été obtenu par digestion BamHI-AvrII d'un amplicon de 1,3 kb obtenu par PCR sur de l'ADN génomique humain et dont les extrémitées ont été rendues franches puis insérées dans la construction obtenue précédemment et suite à une double digestion XbaI-BamHI rendue franche, ce qui a créé un site ClaI à la jonction BamHI/BamHI. Finalement, le troisième élément inséré fut un fragment de 6,5 kb BsaHI-EcoRV rendu franc par remplissage de l'extrémité sortante 5' laissée par BsaHI et couvrant la région +49 à +6558 du gène ^{A γ} . Ce fragment a ensuite été inséré dans le site ClaI créé à la jonction BamHI/BamHI.

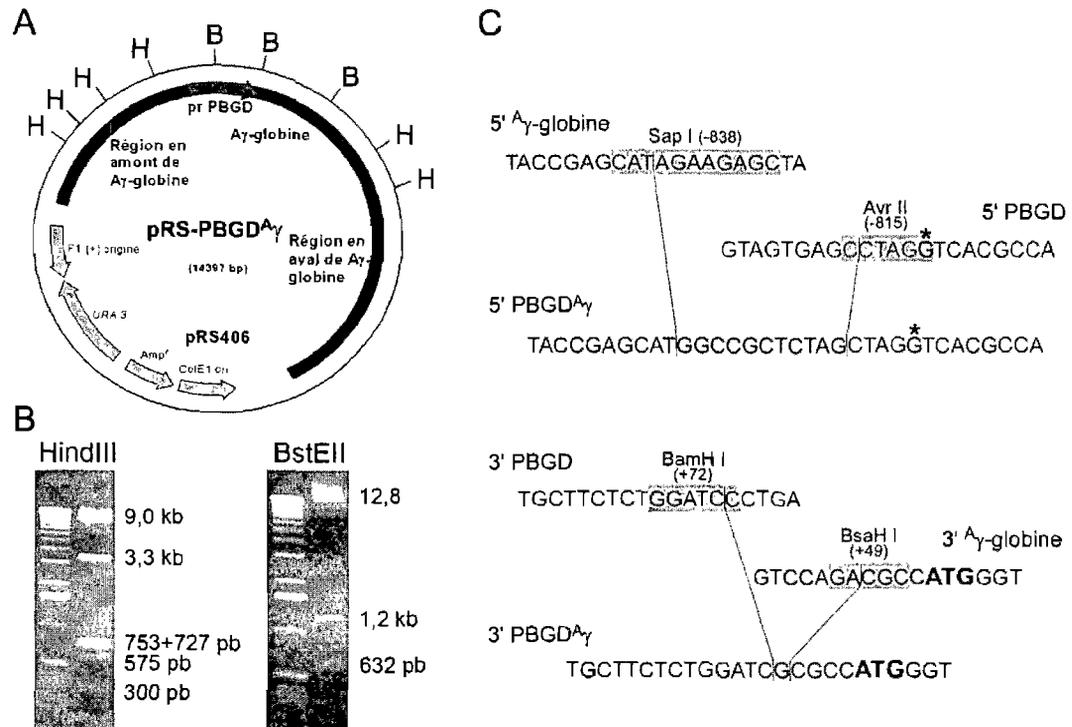


Figure 23: Plasmide d'intégration dans la levure pRS-PBGD $\Delta\gamma$. Génération et caractérisation du YIp utilisé pour la recombinaison homologue avec le β -YAC. (A) Carte du pRS-PBGD $\Delta\gamma$ montrant les sites de restriction HindIII et BstEII utilisés pour la caractérisation. (B) Caractérisation par digestion enzymatique démontrant l'intégrité de la construction. (C) Jonctions A γ /PBGD en 5' du promoteur (haut) et PBGD/ $\Delta\gamma$ en 3' du promoteur (bas). Les sites de restriction utilisés lors du clonage sont indiqués par des boîtes jaunes. Bleu: séquences provenant de A γ (l'ATG du gène γ est indiqué en caractères gras); Rouge: séquences provenant de PBGD; Noir: séquence provenant du site de polyclonage du vecteur pRS406 alors que l'astérisque (*) indique la conversion d'un nucléotide (A \rightarrow G) obtenue par mutagenèse dirigée pour générer le site AvrII.

L'intégrité de la construction pRS-PBGD^{Aγ} a été vérifiée à l'aide de digestions enzymatiques HindIII et BstEII qui ont démontré que tous les éléments de la construction étaient intégrés de façon correcte (Figure 23B). Le plasmide a également été séquencé pour vérifier les jonctions et le promoteur PBGD. Ainsi, une courte séquence provenant du site de polyclonage du vecteur a été intégrée à la jonction 5' du promoteur alors qu'une guanine a été ajoutée dans le 5'UTR à la jonction BamHI/BsaHI, 5 nucléotides devant l'ATG, donc sans altérer la séquence de Kozak (Figure 23C).

2.2.2 Les YIp de délétions

Les YIp de délétion pRSΔ-842/-384^{Aγ}, pRSΔ-384/-202^{Aγ} et pRSΔ-204/-135^{Aγ} ont été obtenus en employant une stratégie différente de celle utilisée pour générer le YIp-PBGD^{Aγ}.

Clonage de ^{Aγ} dans pBluescript II

La première étape de cette stratégie consistait à cloner le fragment SacI/SacI de 3,6 kb contenant le gène et la région promotrice de ^{Aγ} (-2316 à +1307) obtenu du plasmide pVZ-ES, dans le site SacI d'un vecteur pBluescript II dont une partie du site de polyclonage avait été éliminée pour faciliter les étapes suivantes de la génération des délétions (voir Figure 51 à l'Annexe IV, page XX). Ce plasmide pBluescript II modifié avait été obtenu par double digestion ClaI-ApaI suivi de la génération d'extrémités franches par traitement à la Mung Bean nucléase et d'une recircularisation par traitement à la ligase. L'analyse par séquençage du plasmide ainsi obtenu a démontré que le traitement à la Mung Bean nucléase avait enlevé des séquences supplémentaires de chaque côté des sites ApaI et ClaI mais sans affecter le site SacI servant au clonage (Figure 51).

La délétion Δ-842/-384^{Aγ} (génération de pBΔ-842/-384^{Aγ})

La génération de cette délétion a été obtenue par deux digestions successives suivi d'une ligation. La première étape consistait en fait en une digestion partielle SapI (-841)

puisque un site SapI était également présent dans le vecteur pBluescript. Après la génération de bouts francs à l'aide de la Klenow, une digestion StuI (-384) a été effectuée et le plasmide recircularisé. La construction obtenue a été caractérisée à l'aide d'une digestion DraI puisque cette enzyme possède un site de restriction dans la région éliminée du promoteur γ , permettant ainsi de distinguer le mutant de la séquence sauvage (Figure 24). Le séquençage du promoteur a démontré que, suite à la stratégie employée, la délétion couvrait la région -842 à -384 (Figure 43, Annexe II, page VIII).

La délétion Δ -384/-202 γ (génération de pB Δ -384/-202 γ)

Cette délétion a été obtenue par des digestions successives ApaI (-202) et StuI (-384) suivi de la génération d'extrémités franches à l'aide de la Klenow, puis de la recircularisation du plasmide. Le criblage et la caractérisation des clones ont été effectués à l'aide d'une digestion BpmI puisque un site de restriction est présent dans la région couverte par la délétion (Figure 24). Le séquençage du promoteur a confirmé que la délétion a été produite telle que prévue (Figure 44, Annexe II, page IX).

La délétion Δ -204/-135 γ (génération de pB Δ -204/-135 γ)

Cette délétion a été produite en employant une double digestion ApaI (-202) et PflmI (-141) suivi de la génération d'extrémités franches avec la Mung Bean nucléase et de la ligation. Le criblage et la caractérisation a été effectuée avec l'enzyme BanII, étant la seule enzyme coupant dans la région éliminée (Figure 24). Le séquençage a démontré toutefois que la délétion couvrait la région -204 à -135 (Figure 45, Annexe II, page X).

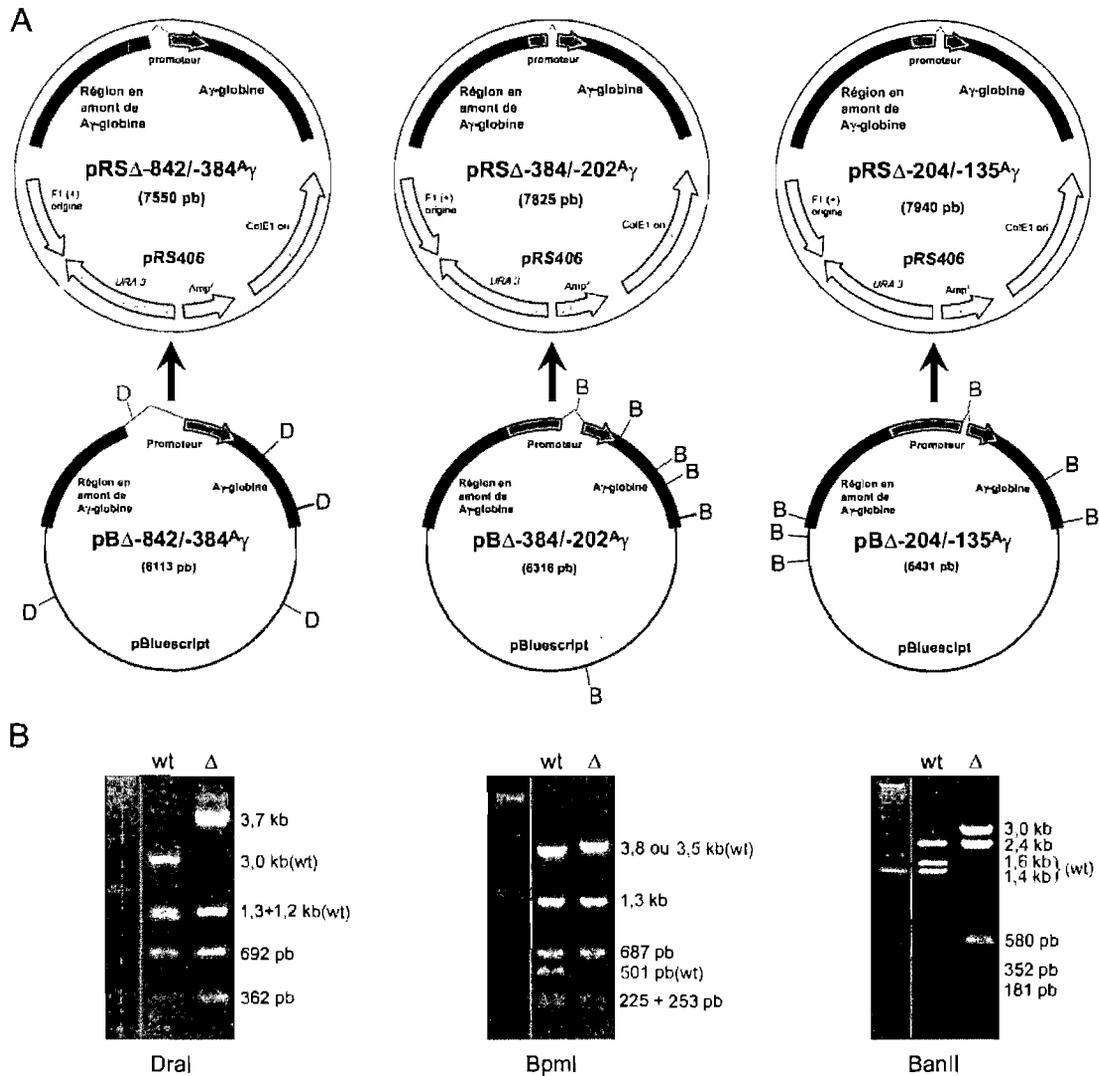


Figure 24: Les YIp de délétion. Génération et caractérisation des plasmides contenant les trois délétions du promoteur de $A\gamma$ et génération des YIp pour la recombinaison homologue. (A) Carte des 3 YIp pRS Δ et des plasmides pB Δ ayant servi à la génération des délétions. (B) Caractérisation par digestion enzymatique démontrant l'intégrité des plasmides pB Δ après les délétions; En gris sont indiquées les bandes uniques aux constructions sauvages (avant les délétions). (wt): type sauvage (*wild type*).

Génération des trois YIp

Les fragments contenant la séquence de $^A\gamma$ et les différentes délétions du promoteur ont été excisés des trois plasmides pB Δ par une digestion PstII (un isoschisomère de SacI et plus efficace que ce dernier) puis clonés dans le site SacI du vecteur pRS406 pour générer les YIp pRS Δ -842/-384 $^A\gamma$, pRS Δ -384/-202 $^A\gamma$ et pRS Δ -204/-135 $^A\gamma$ (Figure 24A). Ces vecteurs ont été caractérisés par séquençage ce qui a permis de démontrer que les trois fragments étaient intégrés correctement dans le vecteur. Cependant, le séquençage a démontré que la position -117 de chacune de ces constructions présentait le A caractéristique de la mutation HPFH grecque -117(G \rightarrow A) (Annexe II, Figure 44 et Figure 45, pages IX-X).

Mutagenèse à la position -117 des trois YIp

Pour éviter que cette mutation -117(G \rightarrow A) qui est reconnue pour activer le promoteur au stade adulte ne fausse les résultats obtenus par la délétion des régions en amont, la séquence sauvage G a été réintroduite par mutagenèse dirigée dans chacun des YIp Δ (Figure 46, Annexe II, page XI). Toute la séquence de l'insert des YIp Δ a également été séquencée ce qui a démontré que l'utilisation d'une polymérase à haute fidélité avait permis d'éviter la création de mutations ailleurs dans les constructions.

2.3 Recombinaison homologue dans la levure

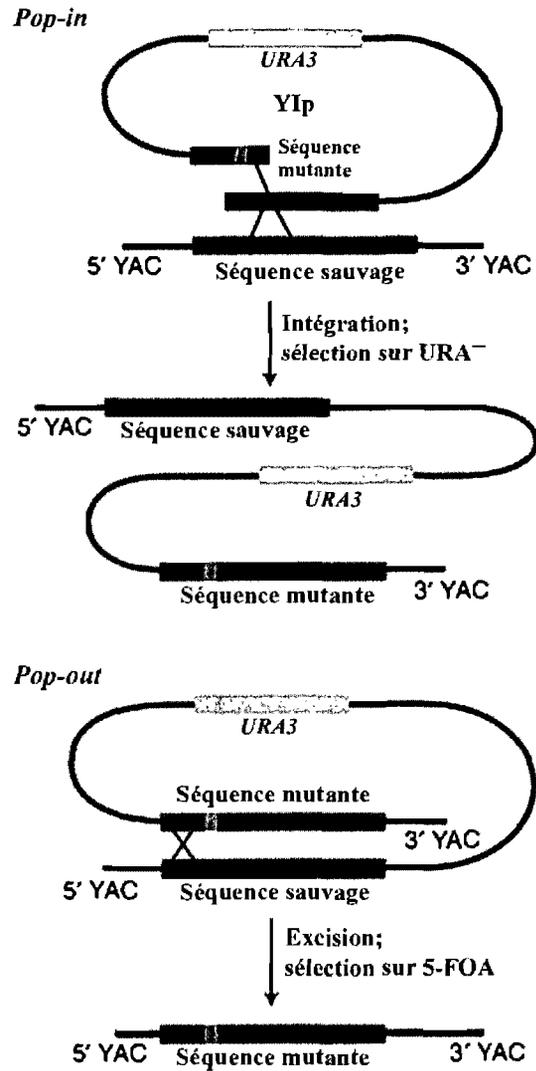


Figure 25: Recombinaison homologue dans la levure. Adapté de Peterson *et al.*⁵²⁵

2.3.1 Culture et manipulation des levures

Souche de levure

La recombinaison homologue est effectuée dans la levure *Saccharomyces cerevisiae* A201F4. Cette levure de souche AB1380 (MATa, *ura3-52*, *Trp1-289*, *Lys2-1*, *His5*, *can100*) possède un chromosome artificiel de levure (YAC) d'environ 150 kb. Celui-ci contient le locus entier de la β -globine humaine, ~40 kb de séquence en amont du LCR, ~20 kb en aval du gène de β -globine¹⁰⁸ ainsi que les gènes de prototrophie TRP1 et LYS2.³⁷⁰

Milieux de culture

Toutes les cultures de levure ont été faites dans les milieux *Drop-out* Trp⁻Lys⁻Ura⁻ ou *Drop-out* Trp⁻Lys⁻. Le milieu *Drop-out* 5-FOA, plus riche en acides aminés, a servi à la sélection à l'acide 5-fluoroorotique. Dans tous les cas, lorsqu'un milieu solide en pétri est nécessaire, 20 g/L de bactoagar est ajouté au milieu liquide. La composition des milieux est donnée dans le Tableau VII, ci-dessous.

Tableau VII: Milieux de culture de levure.

<i>Drop-out Trp⁻Lys⁻Ura⁻</i>	<i>Drop-out Trp-Lys-</i>	<i>Drop-out 5-FOA</i>
dextrose (0,1 M) sulfate d'ammonium (40 mM) base azotée de levure (1,7 g/L) adenine-HCl (0,2 mM) L-arginine-HCl (0,2 mM) L-histidine (0,1 mM) L-isoleucine (0,5 mM) L-leucine (0,5 mM) L-méthionine (0,1 mM) L-phénylalanine (0,3 mM) L-thréonine (1,7 mM) L-tyrosine (0,3 mM)	dextrose (0,1 M) sulfate d'ammonium (40 mM) base azotée de levure (1,7 g/L) adenine-HCl (0,2 mM) L-arginine-HCl (0,2 mM) L-histidine (0,1 mM) L-isoleucine (0,5 mM) L-leucine (0,5 mM) L-méthionine (0,1 mM) L-phénylalanine (0,3 mM) L-thréonine (1,7 mM) L-tyrosine (0,3 mM) uracile (0,1 mM)	dextrose (0,1 M) sulfate d'ammonium (40 mM) base azotée de levure (1,7 g/L) adenine-HCl (0,5 mM) L-arginine-HCl (0,2 mM) L-histidine (0,2 mM) L-isoleucine (0,5 mM) L-leucine (1,0 mM) L-méthionine (0,3 mM) L-phénylalanine (0,7 mM) L-thréonine (3,8 mM) L-tyrosine (0,4 mM) L-sérine (0,9 mM) acide L-aspartique (0,7 mM) acide L-glutamique (1,75 mM) L-lysine (0,5 mM) L-valine (3,4 mM) L-tryptophane (0,5 mM) uracile (0,2 mM) acide 5-fluoroortique (5,7 mM)

Mini-préparation d'ADN génomique de levure

Un clone de levure est repiqué et mis en culture dans un milieu *Drop-out* approprié pendant 36 h à 30°C. Les sphéroblastes sont préparés en digérant la paroi cellulaire avec la lyticase dans un volume de réaction de 600 µL [sorbitol (0,75 M), Tris-HCl (83 mM, pH 8,0), EDTA (83 mM), β-mercaptoéthanol (23 mM), lyticase (Sigma, 0,15 U/µL)] pendant 1h à 37°C. Ensuite, les sphéroblastes sont lysés dans une solution de lyse [tris-HCl (50 mM, pH 8,0), EDTA (20 mM), SDS (1%)] pendant 20 min à 65°C. Les débris cellulaires sont précipités sur glace avec 1,3 M d'acétate de potassium. L'ADN du surnageant est précipité avec 1,5 volumes d'éthanol (100%) puis resuspendu dans du tampon TE.

Finalement, l'ADN est traité avec 50 µg de RNase A (Roche Diagnostic) pour éliminer l'ARN.

Encapsulation des levures en agarose *LMP*

Une préculture est transférée dans 500 mL de milieu *Drop-out* approprié pour une durée de 24 h. Après un lavage dans une solution d'EDTA (50 mM, pH 8,0) le culot de cellule est resuspendu et incubé dans 0,3 volume de solution à sphéroblastes [Sorbitol (1 M), citrate de sodium (0,1 M), EDTA (10 mM, pH 7,5), DTT (33 mM), lyticase (0,38 U/µL)] pendant 2-3 h à 37°C ou jusqu'à ce que la sphéroblistation soit complète. L'agarose à bas point de fusion *Seeplaque-LMP GTG* (2%) dans l'EDTA (100 mM, pH 8,0) est ajouté à un ratio de 1 pour 1 aux sphéroblastes et le mélange est moulé dans des Tygon® où les capsules sont laissées à figer pendant 1 h à 4°C. Les capsules sont mises dans une solution réductrice [EDTA (0,48 M), Tris-HCl (10 mM, pH 8,0), DTT (50 mM)] et incubées pendant 16 h à 37°C. Elles sont ensuite transférées dans la solution de protéolyse [EDTA (0,44 M), Tris-HCl (10 mM, pH 8,0), sarcosyle (1%), protéinase K (0,5 mg/mL)] et incubées 4 h à 50°C. Les capsules sont lavées plusieurs fois dans du tampon SE [NaCl (75 mM), EDTA (25 mM, pH 8,0)] et conservées dans une solution d'EDTA (100 mM, pH 8,0).

2.3.2 Transformation de levures: Insertion du YIp (*Pop-in*)

Une colonie est repiquée et inoculée dans 2,5 mL de milieu *Drop-out* Trp⁻Lys⁻ puis incubée à 30°C jusqu'à ce que la culture se trouve en phase logarithmique ($DO_{600} = 0,2$). Les cellules sont transférées dans 50 mL de milieu *Drop-out* Trp⁻Lys⁻ et l'incubation est poursuivie jusqu'à ce qu'une $DO_{600} = 0,6$ soit obtenue. Après une centrifugation à 2500×g, le culot de cellule est resuspendu dans la solution d'acétate de lithium [Tris-HCl (10 mM, pH 8,0), EDTA (1 mM), acétate de lithium (100 mM)] puis 1 à 10 µg de plasmide d'intégration dans la levure (YIp) sont ajoutés. Les cellules sont incubées pendant 30 min à 30°C puis 7 volumes de solution de PEG [Tris-HCl (10 mM, pH 8,0), EDTA (1 mM), polyéthylène glycol (45% m/v), acétate de lithium (100 mM)] sont ajoutés. L'incubation est

continué pour 45 min. La transformation est effectuée par un choc thermique à 42°C pendant 5 min. Les cellules sont ensuite étalées sur un pétri *Drop-out* Trp⁻Lys⁻Ura⁻ solide, puis incubées pendant 72 h à 30°C.

2.3.3 Excision du YIp (*Pop-out*)

Les clones *Pop-in* sélectionnés sont repiqués et inoculés dans 3 mL de milieu *Drop-out* Trp⁻Lys⁻ pendant 16 h. Ils sont ensuite étalés sur un pétri *Drop-out* 5-FOA solide et incubés pendant 72 h à 30°C. Les colonies sont repiquées et striées sur un pétri *Drop-out* Trp⁻Lys⁻. Les clones sont alors mis en culture pour effectuer des mini-préparations d'ADN génomique de levure et/ou pour encapsulation en agarose *LMP* selon les protocoles décrits à la section 2.3.1.

2.3.4 Caractérisation des YAC mutants

Criblage des clones par PCR

Le criblage des *Pop-out* de délétion a été effectué par PCR sur l'ADN génomique des levures (dilué 1:10) dans les conditions suivantes: 5 min de dénaturation à 94°C, puis 30 cycles d'amplification (94°C-45 s, 62°C-45 s, 72°C-1 min 30 s) dans un volume de réaction de 20 µL [Tris (10 mM, pH 8,5), KCl (50 mM), MgCl₂ (1,25 mM), contenant les amorces FR-φ3/00-52 (1 µM de chaque), les dNTP (125 µM) et la Taq polymérase (0,14 U/µL)]. Les amorces sont donnée au Tableau XIII (Annexe I, , page III).

Digestion par endonucléases de restriction

Lorsque l'ADN génomique de levure est produit par mini-préparation, des digestions par endonucléases de restriction sont effectuées tel que décrit à la section 2.1.1. L'ADN en capsule d'agarose *LMP* est digéré pendant 16 h après une équilibration de 2 h dans le tampon de réaction de l'enzyme à être utilisée. Les digestions peuvent être effectuées simultanément ou séquentiellement quand plus d'un enzyme est utilisé pour

digérer l'ADN. Lorsque les digestions sont effectuées séquentiellement, les capsules sont équilibrées à chaque fois dans le tampon approprié pour l'enzyme.

Buvardage de type *Southern*

L'ADN est digéré à l'aide d'une endonucléase de restriction puis est déposé sur un gel d'agarose (1% dans TAE) après l'ajout d'ADNss fragmenté aux ultrasons (5 µg). L'ADN est fait migré sur gel ~4 h à 95 V puis ce dernier est incubé dans une solution de dénaturation [NaOH (0,5 M), NaCl (1 M)] pendant 30 min avant d'être incubé dans une solution de neutralisation [Tris-HCl (0,5 M, pH 7,4), NaCl (3 M)] pendant 60 min. L'ADN est transféré sur une membrane de Nylon par capillarité pour une période de 16 h puis séché. Finalement, l'ADN est fixé de façon permanente à la membrane par exposition au rayonnement UV (1200 µJ). Le radiomarquage de sondes couvrant tout le locus et l'hybridation sont effectuées tel que décrit à la section 2.1.5.

Électrophorèse à champ pulsé (*PFGE*)

L'ADN chromosomique entier ou digéré en capsule d'agarose mou par des enzymes de restriction donnant de larges fragments, est séparé sur un gel d'agarose (1%) dans un tampon 0,5X TBE [Tris-borate (45 mM), EDTA (1 mM)] sur un appareil *PFGE* (Biorad) à 6 V en utilisant un angle de commutation de 120° et un temps de commutation de 14 s pendant 20 h à 14°C. À la fin de la migration, le gel est trempé dans une solution de dépurination de HCl (250 mM) pendant 10 min, puis transféré sur une membrane par buvardage tel que décrit dans la section 2.3.4.

2.4 Génération des YAC mutants

Les YAC comportant le locus humain entier de la β -globine ont été mutés par recombinaison homologue dans la levure par la technique du *Pop-in/Pop-out* en utilisant les plasmides d'intégration dans la levure (YIp) obtenus à la section 2.2.

2.4.1 PBGD^A γ -YAC

La souche de levure A201F4 qui possède le locus humain de la β -globine contenu dans un YAC de 150 kb, a été transformée par choc thermique avec le YIp PRS-PBGD^A γ , puis les clones *Pop-in* ont été sélectionnés pour l'insertion du YIp sur un milieu sans uracile. Les *Pop-out* ont été obtenus par excision du YIp et de la séquence sauvage par recombinaison homologue sur un milieu de sélection négative contenant du 5-FOA. Le criblage des *Pop-in* et les *Pop-out* a été effectué par buvardage de *Southern* en utilisant une digestion EcoO65 I et une sonde spécifique à la région en 5' du promoteur A γ (BanII-SapI, 1,3 kb) qui permettait de distinguer les YAC sauvages, les *Pop-in* dont le YIp est intégré à l'endroit prévu (3' du gène A γ), ceux dont le YIp est inséré au mauvais endroit (5' de A γ) et les *Pop-out*.

Recombinaison homologue dans la levure

Pour favoriser l'insertion du YIp PRS-PBGD^A γ en 3' du gène A γ , le plasmide a été linéarisé par une digestion BspEI qui coupe à +834 du site d'initiation de la transcription. Cette approche devait permettre une plus grande efficacité d'insertion au niveau du grand bras d'homologie.

Caractérisation du YAC mutant

Les trois mutants ont été caractérisés par buvardage de type *Southern* en utilisant des sondes réparties le long du locus: HS2 (HpaI-HindIII, 1,6 kb), ϵ (ClaI-PvuI, 0,9 kb), ^G γ (StuI-SphI, 2,6 kb) et β (BamHI-SphI, 1,1 kb). Il est à noter que la sonde ^G γ reconnaissait également le gène A γ alors que la sonde β reconnaissait partiellement le gène δ . Chacune de ces quatre sondes a été utilisée sur quatre digestions (EcoRI, PstI, BamHI et HindIII) qui permettaient une couverture extensive du locus. Il est à remarquer que la sonde ϵ utilisée, possédait des séquences de vecteur et détectait également un des bras du YAC ce qui explique la présence d'une bande supplémentaire à celle attendue.

2.4.2 Δ -842/-384^A γ -YAC

Pour le mutant de délétion Δ -842/-384^A γ -YAC, un criblage rapide des *Pop-in* et les *Pop-out* a été effectué par PCR en utilisant des amorces situées de chaque côté de la délétion permettant de détecter la présence de cette dernière. Par la suite, une analyse exhaustive a été effectuée sur certains clones de façon à vérifier l'intégrité du locus et dans le cas des *Pop-in*, l'intégration correcte du YIp.

Recombinaison homologue dans la levure

Le plasmide a été linéarisé par une digestion Bsu36I qui coupe à +661 du site d'initiation de la transcription puis transformé dans la souche de levure A201F4. Après la transformation, 10 clones poussant sur le milieu sans uracile ont été criblés par PCR et tous ont démontré l'intégration du YIp.

Caractérisation du YAC mutant

Les 5 mutants sélectionnés ont été caractérisés par buvardage de type *Southern* en utilisant les mêmes sondes réparties le long du locus que pour le PBGD^A γ -YAC (Section 3.4.1, page 105). Chacune de ces quatre sondes a été utilisée sur deux digestions (HS2, ^G γ et β sur EcoRI et BamHI; ϵ sur PstI et HindIII) qui permettent une couverture extensive du locus.

2.4.3 Δ -384/-202^A γ -YAC

Pour le mutant de délétion Δ -384/-202^A γ -YAC, une stratégie identique à celle employée pour le mutant Δ -842/-384^A γ -YAC a été employée (Section 2.4.2).

Recombinaison homologue dans la levure

Le plasmide linéarisé par Bsu36I a été transformé dans la souche de levure A201F4.

Caractérisation du YAC mutant

Les 5 mutants sélectionnés ont été caractérisés par buvardage de type *Southern* en utilisant les mêmes sondes réparties le long du locus que pour le PBGD^A γ -YAC (Section 2.4.1, page 105). Chacune de ces quatre sondes a été utilisée sur deux digestions (HS2, ^G γ et β sur EcoRI et BamHI; ϵ sur PstI et HindIII) qui permettent une couverture extensive du locus.

2.4.4 Δ -204/-135^A γ -YAC

Pour le mutant de délétion Δ -204/-135^A γ -YAC, une stratégie identique à celle employée pour le mutant Δ -842/-384^A γ -YAC a été employée (Section 2.4.2).

Recombinaison homologue dans la levure

Le plasmide linéarisé par Bsu36I a été transformé dans la souche de levure A201F4.

Caractérisation du YAC mutant

Les 8 mutants sélectionnés ont été caractérisés par buvardage de type *Southern* en utilisant les mêmes sondes réparties le long du locus que pour les YAC précédents. Deux ou trois digestions pour chacune des sondes (EcoRI et BamHI pour HS2 et β ; PstI et BamHI pour ϵ ; EcoRI, HindIII et BanII pour ^G γ) permettant une couverture extensive du locus ont été utilisées.

2.5 Souris transgéniques

2.5.1 Préparation du YAC pour la microinjection

Le YAC en capsule d'agarose *LMP* est isolé par *PFGE* tel que décrit à la section 2.3.4 (page 103), sur un gel d'agarose 1% à bas point de fusion. La bande correspondante au YAC est excisée du gel et encastrée de façon parallèle au champ électrique dans un gel

d'agarose 4% à bas point de fusion. Après une migration de 16 h à 35 V, l'ADN concentré à l'interface des gels 1%-4% est excisé et équilibré dans 50 mL de tampon de microinjection [Tris-HCl (10 mM, pH 7,5), EDTA (0,1 mM), NaCl (100 mM), spermidine (7 μ M), spermine (3 μ M)] pendant un minimum de 16 h à 4°C. Les blocs d'agarose sont asséchés sur un filtre et incubé 10 min à 66°C ou jusqu'à fusion complète de l'agarose. L'ADN est libéré par digestion de l'agarose à la β -agarase (New England Biolab, 3 U/100 μ L) pendant 3 h à 42°C. Un aliquote du YAC isolé est de nouveau séparé par *PFGE* pour vérifier la présence de dégradation de l'ADN.

2.5.2 Microinjection du YAC

Les oocytes fécondés sont récupérés des oviductes de femelles (C57Bl/6J X CBA/J) F1 superovulées par injection de sérum de jument gravide (PMS) et de gonadotrophine chorionique humaine (hCG) puis traités à la hyaluronidase (Sigma, 100 U/mL) pour éliminer les cellules du cumulus oophorus. L'ADN est injecté directement dans le pronucléus des oocytes à l'aide d'un micromanipulateur sous microscope inversé *Diaphot* avec contraste interférentiel de Nomarski (Nikon). Typiquement, 200-350 œufs sont microinjectés en une journée et environ 25-50% sont sélectionnés sur la base de leur viabilité pour être transférés dans les pseudogestantes.

Les œufs sont implantés dans l'infundibulum de l'oviducte de pseudogestantes obtenues par accouplement avec un mâle vasectomisé, après dégagement de l'ovaire à l'aide de pinces fines. L'ovaire est par la suite remis en place dans l'abdomen de la souris et l'incision est refermée à l'aide d'agrafes Autoclips (Clay Adams Co.).

2.5.3 Caractérisation des fondatrices

Génotypage

Les souris générées sont génotypées pour la détection du transgène. Les souris n'ayant pas intégré le transgène sont considérées « négatives » et sont éliminées ou peuvent

être utilisées comme souris contrôles de type sauvage. Les souris ayant intégré le transgène sont identifiées "positives" et sont conservées.

Isolation de l'ADN génomique

L'ADN génomique de souris est isolé d'un bout de queue selon le protocole décrit à la section 2.1.2.

Buvar dage de type Southern

Environ 10 µg d'ADN de queue de souris sont digérés par endonucléase de restriction puis migrés sur un gel d'agarose (1% dans TAE) pendant ~4 h à 95 V. Le gel est incubé dans la solution de dénaturation [NaOH (0,5 M), NaCl (1 M)] pendant 30 min puis incubé dans la solution de neutralisation [Tris-HCl (0,5 M, pH 7,4), NaCl (3 M)] pendant 60 min. L'ADN est transféré sur une membrane de Nylon par capillarité pour une période de 16 h puis asséché. L'ADN est fixé de façon permanente à la membrane par exposition aux UV (1200 µJ). Le radiomarquage de sondes couvrant tout le locus et l'hybridation sont effectués tel que décrit à la section 2.1.5.

PCR

Le génotypage a été également effectué par PCR sur l'ADN génomique dilué 1:10 dans les conditions suivantes: 5 min de dénaturation à 94°C suivit de 30 cycles d'amplification (94°C-45 s, 62°C-45 s, 72°C-1 min 30 s) dans un volume de réaction de 20 µL [Tris (10 mM, pH 8,5), KCl (50 mM), MgCl₂ (1,25 mM), amorces FR-φ3/00-52 (1 µM de chaque), dNTP (125 µM), Taq polymérase (0,14 U/µL)]. La liste des amorces est donnée au Tableau XIII (Annexe I, , page III).

Caractérisation de l'intégrité des transgènes

L'intégrité du transgène dans les souris fondatrices est vérifiée par comparaison avec le locus sauvage par deux approches: 1) un buvardage de *Southern* (section 2.3.4) utilisant des sondes couvrant les différentes régions du locus est utilisé pour vérifier l'intégrité du

locus. 2) Un PCR en duplexe de 30 cycles d'amplification (94°C-45 s, 58°C-45 s, 72°C-1 min 30 s) suivant une étape de dénaturation de 5 min à 94°C dans un volume de réaction de 20 µL [Tris (10 mM, pH 8,5), KCl (50 mM), MgCl₂ (2 mM), amorces 98-07/98-08/98-09/98-10 (1 µM chaque), dNTP (62,5 µM), Taq polymérase (0,14 U/µL)] est utilisé pour vérifier la présence des deux bras du YAC. La liste des amorces est donnée au Tableau XIII (Annexe I, page III).

2.5.4 Établissement des lignées transgéniques

Les fondatrices sont croisées avec des souris sauvages (C57Bl/6J X CBA/J) F1 puis les F2 positives sont croisées avec des F2 négatives sélectionnées pour l'haplotype *diffus* tel que décrit à la section suivante. Les descendants d'une fondatrice forment une lignée transgénique et possèdent le numéro de la fondatrice (ou numéro de la lignée) plus un numéro d'identification propre à chaque souris, séparé du numéro de la lignée par un trait d'union.

Sélection pour l'haplotype *diffus*

Chez les souris transgéniques issues de croisements (C57Bl/6J X CBA/J) F1, il existe 2 haplotypes différents pour la β-globine endogène : l'haplotype *single* ou *s* de la C57Bl/6J (*s/s*) qui produit une seule β-globine appelée β-*single* ou β^S et l'haplotype *diffus* ou *d* de la CBA/J (*d/d*) qui produit deux β-globines appelées β-majeure ou β^{Maj} et β-mineure ou β^{Min}. L'identification de l'haplotype de la β-globine endogène est effectuée par électrophorèse sur plaque d'acétate de cellulose. Les globules rouges culottés par centrifugation (~5 µL) sont lysés dans une solution d'hémolyse [cystamine diHCl (0,33 M), DTT (6,7 mM), NH₄OH (0,5 M)]. Les échantillons sont déposés à l'aide d'un applicateur *Super Z-12* (Helena laboratory) sur une plaque d'acétate de cellulose *TITAN III-H* (Helena Laboratory) humectée avec du TBE 1X pour Helena [Tris (84,2 mM), EDTA (1,6 mM), acide borique (51,7 mM)]. La migration se fait pendant 30 min à 300 V, les membranes sont colorées avec une solution rouge de Ponceau [Ponceau S (5 g/L), acide

trichloroacétique (0,3 M)] puis lavées avec une solution d'acide acétique (5%) à trois reprises. Les membranes sont déshydratées dans l'éthanol (100%), trempées dans une solution méthanol/acide acétique (7:3) puis séchée à 55°C. Une souris contrôle (C57Bl/6J X CBA/J) F1 est utilisée pour identifier les deux haplotypes. Seules les souris transgéniques homozygotes pour l'haplotype *diffus* (*d/d*) sont conservées et utilisées dans les expériences subséquentes.

Nombre de copies

Le nombre de copies du transgène est mesuré dans les souris F2, écartant l'effet d'un possible mosaïsme des fondatrices. Cette quantification est faite par buvardage de type *Southern* tel que décrit à la section 2.3.4. L'intensité du signal obtenue pour le gène β -globine est comparée au gène endogène de l'adénosine désaminase (ADA) et corrigée pour la longueur de la sonde. La quantification de l'autoradiogramme est effectuée par la digitalisation à l'aide d'un système *Alpha Imager 2200* (Alpha Innotech) et la quantification avec le logiciel *ImageQuant v5.0*.

2.5.5 Acquisition de fœtus transgéniques

Pour obtenir des fœtus à différents stades du développement, une femelle de type sauvage sur fond génétique mixte (C57Bl/6J et CBA/J), homozygote pour l'haplotype *d/d* est fécondée par un mâle transgénique. La date de la fécondation est déterminée par la présence du bouchon coïtal et est considéré comme le jour embryonnaire e0,5. Les fœtus sont récupérés aux jours e9,5, e10,5, e12,5, e14,5 ou e16,5 par césarienne après euthanasie de la femelle par dislocation cervicale. Le géotypage se fait par PCR sur l'ADN extrait d'une portion du placenta tel que décrit à la section 2.3.4.

2.5.6 Prélèvement d'organes

La moelle osseuse

Les cellules de la moelle osseuse des deux fémurs sont chassées de la cavité médullaire à l'aide d'une aiguille contenant 2 ml de PBS/héparine (1000 U/mL) fait avec de l'eau traitée au diéthylpyrocarbonate (DEPC). Après centrifugation, le culot est rapidement plongé dans l'azote liquide et les tubes sont conservés à -80°C.

Le sang périphérique

Pour l'extraction d'ARN, ~1,5 mL de sang est récupéré par ponction cardiaque sous anesthésie générale à l'aide d'une seringue préalablement traitée à l'héparine. Le sang est centrifugé et le culot est rapidement plongé dans l'azote liquide. Pour une analyse au niveau protéique chez l'adulte, ~50 µL de sang périphérique sont récupérés dans un capillaire traité à héparine (Fisher) par incision de la queue. Par la suite le sang est lavé trois fois dans 1 mL de saline [NaCl (0,9%)]. Chez les fœtus et les nouveau-nés, le sang périphérique est obtenu en plongeant les fœtus dans la saline après section du cordon ombilical ou décapitation. Dans tous les cas, les culots de globules rouges sont conservés à -80°C.

La rate

La rate est prélevée des nouveau-nés par incision du flanc gauche après euthanasie, puis plongée rapidement dans l'azote liquide et conservée à -80°C.

Le foie fœtal et le sac vitellin

Le foie des fœtus aux jours e12,5 à e16,5 ainsi que le sac vitellin des embryons (e9,5 et e10,5) sont récupérés à l'aide de pinces fines sous microscopie puis plongés rapidement dans l'azote liquide et conservée à -80°C.

2.5.7 Les souris PBGD^A γ -YAC

Génération des fondatrices

Pour générer les fondatrices PBGD^A γ -YAC, une quarantaine de jours de microinjection ont été nécessaires. 473 bébés ont été obtenus et le criblage a été effectué par buvardage de type *Southern* en utilisant la sonde contre γ utilisée à la section précédente ou encore par amplification du promoteur γ par PCR (section 2.5.3). Un total de 17 souris se sont avérées positives pour le gène γ et ont donc été caractérisées pour l'intégrité du transgène.

Caractérisation des fondatrices et établissement des lignées

La caractérisation des fondatrices a été effectuée par buvardage de type *Southern* sur l'ADN génomique digéré EcoRI en utilisant les sondes HS2, ϵ , γ et β qui avaient été utilisées pour la caractérisation des clones *Pop-out*, alors que la présence des deux bras du YAC a été détectée par PCR en multiplex (section 2.5.3). Toutes ses lignées ont été produites sur fond génétique mixte C57Bl/6J et CBA/J et ont par la suite été sélectionnées pour l'allèle endogène de globine de l'haplotype *diffus* par électrophorèse sur acétate de cellulose (section 2.5.4, page 110), puisque pour les analyses protéiques, la β^{single} produite par l'haplotype *single* ne peut être distinguée de la β humaine.

2.5.8 Les souris Δ -842/-384^A γ -YAC

Génération des fondatrices

Pour générer les fondatrices Δ -842/-384^A γ -YAC, neuf jours de microinjection ont été nécessaires et 94 souris ont été obtenues. Le criblage des souriceaux a été effectué par amplification du promoteur γ par PCR. Un total de 3 souris se sont avérées positives pour le gène γ et ont donc été caractérisées pour l'intégrité du transgène.

Caractérisation des fondatrices et établissement des lignées

L'ADN génomique des fondatrices a été caractérisé comme à la section précédente. Comme pour les souris PBGD^Δγ-YAC, toutes ces lignées ont été produites sur fond génétique mixte C57Bl/6J et CBA/J et ont été ensuite sélectionnées pour l'allèle endogène de globine de l'haplotype *diffus*.

2.5.9 Les souris Δ-384/-202^Δγ-YAC

Génération des fondatrices

Pour générer les fondatrices Δ-384/-202^Δγ-YAC, huit jours de microinjection ont été nécessaires et 61 souris ont été obtenues. Le criblage a été effectué par amplification du promoteur γ par PCR. Un total de 3 souris se sont avérées positives pour le gène γ et ont été caractérisées pour l'intégrité du transgène.

Caractérisation des fondatrices

L'ADN génomique des fondatrices a été caractérisé comme à la section précédente. La lignée qui ne présentait pas de réarrangement au niveau du transgène et qui a été produite sur fond génétique mixte C57Bl/6J et CBA/J et d'haplotype *d/d* a ensuite été maintenue homozygote pour l'allèle *diffus* (section 2.5.4, page 110). Les lignées qui présentaient un locus tronqué ou réarrangé n'ont pas été retenues pour la suite de cette étude.

2.5.10 Les souris Δ-204/-135^Δγ-YAC

À cause de la grande difficulté à purifier les YAC pour la microinjection, le mutant Δ-204/-135^Δγ-YAC n'a pas été microinjecté et donc les souris n'ont pas encore été produites. La génération de ces souris représente donc une des étapes futures de cette étude.

2.5.11 La souris contrôle β -YAC

Toutes les expériences effectuées sur les souris transgéniques ont été comparées avec une souris ayant le locus sauvage de la β -globine humaine (la souris β -YAC) générée à partir du même YAC que celui utilisé pour notre étude (La souris β -YAC a été obtenue du Dr. James Douglas Engel de l'université Northwestern en Illinois).³⁷⁰ Le criblage des souris β -YAC est effectué par PCR ou par électrophorèse sur acétate de cellulose. Il est à noter que ces souris ont été croisées sur fond génétique mixte C57Bl/6J et CBA/J et ont été sélectionnées pour l'allèle endogène de globine d'haplotype *diffus* par électrophorèse sur acétate de cellulose. La liste des amorces ayant servi au génotypage est donnée au tableau Tableau XIII (Annexe I, page III).

2.5.12 La souris EKLF-nulle

La souris modifiée par ablation génique du gène endogène EKLF^{400;489} a été obtenue du Dr. Éric Milot du centre de recherche Guy-Bernier de l'hôpital Maisonneuve-Rosemont. Le criblage de ces souris est effectué par PCR. Ces souris ont été croisées sur fond génétique mixte C57Bl/6J et CBA/J puis sélectionnées pour l'allèle endogène de globine d'haplotype *diffus* (section 2.5.4, page 110). La liste des amorces ayant servi au génotypage est donnée au tableau Tableau XIII (Annexe I, page III).

2.5.13 La souris Ikaros-nulle

La souris modifiée par ablation génique du gène endogène Ikaros⁴¹⁵ a été obtenue du Dr. Michel Cayouette de l'IRCM. Le criblage des souris est effectué par PCR. Ces souris ont également été croisées sur fond génétique mixte C57Bl/6J et CBA/J et sélectionnées pour l'allèle endogène de globine d'haplotype *diffus* (section 2.5.4, page 110). La liste des amorces ayant servi au génotypage est donnée au tableau Tableau XIII (Annexe I, page III).

2.6 Analyses d'expression

2.6.1 Quantification des ARNm de globine

L'analyse de l'expression des transgènes se fait sur l'ARN total extrait du tissu érythropoïétique spécifique à chaque âge (sac vitellin à e9,5 et e10,5; foie fœtal de e12,5 à e16,5; rate chez les nouveau-nés; moelle osseuse chez les adultes).

Préparation de L'ADN complémentaire

Extraction de l'ARN

L'ARN total est isolé des tissus congelés par homogénéisation au *Polytron* dans 2 mL de TRIzol (Invitrogen). L'homogénat est laissé à reposer 5 min puis mélangé à 0,4 mL de chloroforme. Après centrifugation à 12 000×g, l'ARN se trouvant dans la phase supérieure est précipité avec 1 mL d'isopropanol pendant 10 min à 25°C. Le culot est lavé à l'éthanol 75%-DEPC puis séché à l'air libre pendant environ une heure. L'ARN est resuspendu dans 25-50 µL d'eau DEPC et quantifié par densité optique à $\lambda=260$ nm. L'intégrité est vérifiée sur gel d'agarose à 1,5% contenant de la formaldéhyde (2,2 M).

Traitement à la DNaseI

Pour éliminer les traces d'ADN contaminant, 1 µg d'ARN est incubé pendant 60 min à 37°C dans 10 µL de réaction à la DNaseI [Tris-HCl (50 mM, pH 8,3), KCl (75 mM), MgCl₂ (3 mM), inhibiteur de RNase RNAGuardTM (Pharmacia, 0,32 U/µL), DNaseI FPLCpureTM (Pharmacia, 1U/µL)]. La réaction est arrêtée avec 1,5 mM d'EDTA et incubée pour 10 min à 75°C. Le volume de réaction est conservé à -20°C.

Transcription inverse

L'ARN traité à la DNaseI (2,5 µL) est incubé 5 min à 65°C en présence d'amorces hexamériques dégénérées pd(N)₆ (Pharmacia, 42 ng/µL) puis mise sur la glace. La réaction

de transcription inverse se fait dans un volume final de 20 μL [Tris-HCl (50 mM, pH 8,3), KCl (75 mM), MgCl_2 (3 mM), $\text{pd}(\text{N})_6$ (25 ng/ μL), dNTP (62,5 μM), RNAGuardTM (0,16 U/mL), dithiothréitol (10 mM), transcriptase inverse M-MLV (Invitrogen, 20 U/ μL)] pendant 45 min à 37°C. La réaction est arrêtée par inactivation de l'enzyme à 65°C pendant 15 min. L'ADN complémentaire est conservé dans le milieu de réaction à -20°C.

PCR quantitatif en temps réel

La quantification des ADNc de globine est effectuée par qPCR dans un volume de réaction de 25 ou 15 μL avec 12,5 ou 7,5 μL respectivement de *Mastermix QuantiTectTM SYBR[®] Green* (Qiagen) en présence de 0,3 mM de chaque amorce (voir la liste des amorces au Tableau XIV, Annexe I, page IV). Les conditions d'amplification pour toutes les paires d'amorces sont: une première étape de dénaturation à 95°C pendant 15 min suivi de 45 cycles d'amplification (94°C-30 s, 60°C-35 s, 72°C-30 s) sur un system *Mx4000 Multiplex Quantitative PCR* (Stratagene) ou sur un system *Mx3005P multiplex Quantitative PCR* (Stratagene). La prise des données est faite à trois temps aux étapes d'hybridation (60°C) et d'élongation (72°C). Les données recueillies sont analysées à l'aide des logiciels *Mx4000 v4.20* ou *MxPro v3.00*.

2.6.2 Quantification des chaînes de globine

Préparation du lysat d'hémoglobine

Les culots d'érythrocytes préparés à la section 2.3.6 sont lysés dans 4 volumes d'eau. La concentration en hémoglobine est évaluée par une version modifiée de la méthode colorimétrique développée par David Drabkin⁵²⁶: L'hémoglobine du lysat est transformée en cyan-méthémoglobine dans 50 volumes de solution de Drabkin [KCN (0,77 mM), $\text{K}_3\text{Fe}(\text{CN})_6$ (0,6 mM), KH_2PO_4 (1,03 mM), Triton X-100 (0,1%)] puis la densité optique est mesuré à $\lambda=540$ nm. Le reste du lysat est transféré dans 5 volumes de tampon

d'échantillon [urée (6,7 M), acide acétique (1,45 M), β -mercaptoéthanol (2,4 M)]. Typiquement, de 30-60 μ g d'hémoglobine par échantillon sont analysés par *UT-PAGE*.

Électrophorèse sur gel de polyacrylamide-urée/Triton (*UT-PAGE*)

Les échantillons sont déposés sur un gel de polyacrylamide possédant une phase de séparation [acrylamide (12%), bis-acrylamide (0,08%), urée (6 M), Triton X-100 (2 %), acide acétique (0,87 M), TEMED (0,6%), persulfate d'ammonium (0,06% p/v)] superposée d'une phase de concentration [acrylamide (6%), bis-acrylamide (0,04%), urée (6 M), Triton X-100 (1%), acide acétique (0,96 M), TEMED (0,5%), persulfate d'ammonium (0,06% p/v)]. La migration est effectuée dans un tampon d'électrophorèse à 5% d'acide acétique pendant 17 h à 10 mA/gel avec refroidissement à l'eau. Les gels sont colorés 1 h dans une solution de bleu de Coomassie [Brilliant Blue R (Sigma, 2,5%), méthanol (30%), acide acétique (7%)] et par la suite, incubés 16-24 h dans la solution de décoloration [méthanol (25%), acide acétique (5%)]. L'image de ces gels sont digitalisés à l'aide d'un système *Alpha Imager 2200* et quantifiés avec le logiciel *ImageQuant v5.0*.

2.6.3 Distribution cellulaire de l'expression

La distribution cellulaire de l'expression des globines au niveau protéique est estimée par une analyse cytofluorométrique en flux sur le sang périphérique de souris adultes.

Analyse cellulaire par fluorescence-activée de cellules (facs)

Le marquage intracellulaire de l'hémoglobine pour l'analyse par cytofluorométrie de flux est effectué sur les cellules du sang périphérique à l'aide du protocole modifié de Wallac/Isolab. Le sang (10 μ l) est récupéré dans 1 mL de PBS supplémenté de 2% de sérum fœtal de veau (SFV), lavé dans 1 mL de PBS puis incubé 1 h à 25°C dans 1 mL de PBS/formaldéhyde (2%). Du glutaraldéhyde à 0,01% final est ajouté et l'incubation est poursuivie 30 s. Les cellules sont lavées avec du PBS, resuspendues dans 250 μ L de lait en

poudre sans gras à 5% dans du PBS et incubées pendant 10 min à 25°C. Après centrifugation, le culot de cellules est resuspendu dans 400 µL d'une solution de Triton-0,1% [Triton X-100 (0,1%), BSA (0,1%), NaCl (0,17 M), KCl (3,3 mM), Na₂HPO₄ (10 mM) KH₂PO₄ (1,8 mM)] à laquelle sont ajoutés 6 µg d'anticorps anti-HbA conjugués à la biotine (Perkin Elmer) puis incubé pendant 30 min à 25°C. Après centrifugation, le culot est lavé dans une solution de Triton-0,01% [Triton X-100 (0,01%), BSA (0,1%), NaCl (0,17 M), KCl (3,3 mM), Na₂HPO₄ (10 mM) KH₂PO₄ (1,8 mM)] et resuspendu dans la solution de Triton-0,01% à laquelle sont ajoutés 10 µg d'anticorps anti-HbF conjugués à la fluorescéine isothiocyanate (Perkin Elmer) et 0,4 µg de streptavidine conjuguée à l'allophycocyanine (PharMingen) puis incubé 30 min à 25°C. Les culots sont lavés dans la solution de Triton-0,01%, resuspendus dans 1 mL de PBS et analysés sur un appareil d'analyse par cytométrie en flux *FACSCalibur* (Becton Dickinson) à l'aide du logiciel *CellQuest Pro v4.0.2*. L'analyse est effectuée à l'aide du logiciel *WinMDI v2.8*.

2.7 Analyse épigénétique

2.7.1 Immunoprécipitation de chromatine (ChIP)

Une mesure des niveaux d'acétylation et de méthylation des histones H3 et H4 a été effectuée sur le foie fœtal de fœtus à e12,5 et sur les cellules érythroïdes de la moelle osseuse chez les adultes.

Préparation des échantillons

Le foie des fœtus au jour e12,5, génotypés par électrophorèse du sang sur cellulose acétate, est récupéré à l'aide de pinces fines sous microscopie. Le foie est passé successivement dans une seringue en utilisant différentes jauges d'aiguille (de jauge 18 à jauge 26) pour mettre les cellules en suspension puis culotté par centrifugation. Les cellules de la moelle osseuse des adultes sont récupérées tel que décrit à la section 2.3.6 mais sans être congelées. Les cellules de deux fémurs sont lavées dans 2 mL de PBS/SFV

(2%), resuspendues dans 1mL de PBS et marquées avec $0,2 \mu\text{g}/10^7$ cellules d'un anticorps anti-Ter119 couplé à la R-phycoerythrine (PharMingen) pendant 30 min à 4°C. Les cellules sont ensuite lavées puis resuspendues dans du PBS/SFV (0,8%). Les cellules Ter119⁺ sont séparées sur une trieuse de cellules à haute performance *MoFlo* (Cytomation Inc). Les culots de cellules de foie fœtal et de cellules Ter119⁺ de la moelle osseuse sont fixés 10 min dans du PBS/formaldéhyde (1%) à 37°C puis lavés avec 1,3 mL de PBS contenant des inhibiteurs de protéase [PMSF (1,15 mM), aprotinine (1,15 $\mu\text{g}/\text{mL}$)].

ChIP

Les ChIP sont effectués à l'aide de Kits d'immunoprécipitation (Upstate) selon les directives du fabricant. Les culots sont resuspendus dans la solution de lyse [Tris (50 mM, pH 8,1), EDTA (10 mM), SDS (1%), PMSF (1 mM), aprotinine (1 $\mu\text{g}/\text{mL}$)] puis incubés 10 min sur la glace. L'ADN des lysats est fragmenté aux ultrasons à l'aide d'un sonicateur *Sonic Dismembrator model 100* (Fisher Scientific) pendant 1 min 30 s à force 3,75. Après centrifugation, le surnageant est transféré dans 2,8 mL de tampon de dilution [Tris-HCl (16,7 mM, pH 8,1), EDTA (1,2 mM), Triton X-100 (1,1%), SDS (0,01%), NaCl (167 mM), PMSF (1 mM), aprotinine (1 $\mu\text{g}/\text{mL}$)] puis 200 μL sont conservés pour le contrôle *input*. Au reste, 65 μL de solution de billes d'agarose/protéine A/ADNss (50% dans TE pH 8,0 et 0,05% de NaN_3) sont ajoutés et le tout est incubé 1 h à 4°C avec agitation. Après centrifugation, 800 μL de surnageant sont transférés dans un nouveau tube contenant 8 μL d'anticorps spécifiques et le tout est incubé 16 h à 4°C avec agitation. 65 μL de solution de billes d'agarose/protéine A/ADNss sont ajoutés et incubés pendant 1 h à 4°C. Après centrifugation, le culot est lavé deux fois avec 1 mL de tampon de lavage à faible teneur en sel [Tris-HCl (20mM, pH 8,1), EDTA (2 mM), Triton X-100 (1%), SDS (0,1%), NaCl (150 mM)] 5 min, une fois avec 1 mL de tampon de lavage à haute teneur en sel [Tris-HCl (20mM, pH 8,1), EDTA (2 mM), Triton X-100 (1%), SDS (0,1%), NaCl (0,5 M)] 5 min, 1 fois avec 1 mL de solution de lavage au chlorure de lithium [Tris (10 mM, pH 8,1), EDTA (1 mM), acide désoxycholique (1%), IGEPAL-CA630 (1%), LiCl (0,25 M)] 5 min et deux

fois avec du tampon TE. Le culot est alors resuspendu dans 250 μ L de tampon d'élution [NaHCO_3 (0,1 M), SDS (1%)] et le surnageant est récupéré après centrifugation. Au surnageant est ajouté du NaCl à une concentration finale de 0,2 M et le mélange est incubé 4 h à 65°C. Sont ensuite ajoutés à la réaction, 10 μ L d'EDTA (0,5 M), 20 μ L de Tris-HCl (1 M, pH 6,5) et 2 μ L de protéinase K (100 mg/mL) puis le tout est incubé à 55°C pendant 1 h 30 min. L'ADN est extrait avec 1 volume de phénol:chloroforme:alcool isoamyle (25:24:1), précipité avec 2-2,5 volume d'éthanol à 100% et 5 μ g d'ARN de levure pendant 16 h à -20°C puis lavé avec 150 μ L d'éthanol à 70%. L'ADN est resuspendu dans 30 μ L de TE pH 8,0 et conservé à -20°C.

PCR

Deux approches d'amplification quantitative des produits de l'immunoprécipitation de chromatine ont été utilisées. La première, semi-quantitative, a été utilisée pour une analyse rapide d'un plus grand nombre de régions alors que la deuxième, quantitative, a été utilisée pour l'analyse de régions sélectionnées lors de la première analyse.

PCR semi-quantitatif radioactif

Une analyse des ChIP par PCR radioactif en duplex de différentes portions des promoteurs des gènes γ et β a été faite dans les conditions suivantes: une étape de dénaturation de 5 min à 94°C suivi de 30 cycles d'amplification (94°C-30 s, 60°C-30 s, 72°C-30 s) dans une réaction de 20 μ L [Tris (10 mM, pH 8,8), KCl (50 mM), MgCl_2 (1,25 mM), DMSO (1%), amorces (0,8 μ M chaque), dNTP (125 μ M), $^{32}\text{P}\alpha$ -dCTP (1 nCi/ μ L), Taq polymérase (0,14 U/ μ L)] voir l'Annexe I, Tableau XV, page V pour la liste des amorces. Chaque réaction de PCR amplifie le gène Zfp37 comme contrôle interne et une région d'intérêt. Les produits de PCR sont déposés sur un gel de polyacrylamide de 8%. Après la migration, le gel est séché et ensuite exposé sur un film d'autoradiographie *X-Omat blue XB-1*. La quantification de l'autoradiogramme est effectuée par digitalisation à l'aide d'un système *Alpha Imager 2200* et quantification avec le logiciel *ImageQuant v5.0*.

PCR quantitatif en temps réel

La quantification précise de quelques sites sélectionnés dans le promoteur des gènes γ et β a été fait par PCR quantitatif en temps réel tel que décrit à la section 2.6.1 (page 116) mais avec une température d'hybridation dépendante des amorces utilisées (voir la liste des amorces au Tableau XVI, Annexe I, page V). Contrairement aux analyses semi-quantitatives radioactives, cette analyse se fait en simplex avec le gène contrôle *Zfp37* amplifié séparément.

2.7.2 Analyse de méthylation de l'ADN

Préparation des échantillons

Les foies fœtaux de fœtus à e12,5 sont préparés tel que décrit à la section 2.5.6.

Digestion par endonucléases de restriction sensibles au 5mC

L'ADN des foies fœtaux est isolé selon le protocole décrit à la section 2.1.2 puis digéré par les endonucléases de restriction sensibles à la présence de 5-méthylcytosine (5mC) dans leur site de restriction (*HpaI*, *HhaII* ou *BsaHI*) ou insensible à cette méthylation (*MspI*) de façon séquentielle avec une digestion par l'endonucléase *EcoRI*, selon le protocole décrit à la section 2.1.1.

Buvarbage de *Southern*

Un buvardage de type *Southern* tel que décrit à la section 2.3.4 avec des sondes spécifiques pour le promoteur de γ et de δ/β (section 2.1.5) est effectué pour l'évaluation du niveau de méthylation de ces régions.

Quantification

L'autoradiogramme est digitalisé à l'aide d'un système *Alpha Imager 2200* et quantifié avec le logiciel *ImageQuant v5.0*. Le niveau de méthylation est calculé à partir du ratio de l'intensité de la bande du fragment non coupé méthylé (NC) sur la somme des intensités des bandes non coupées et coupées non-méthylée (C) corrigées pour la longueur.

$$\text{Équation 1: } \% \text{ méthylation} = \left(\frac{\text{bande(NC)}}{\text{bande(NC)} + \text{bande(C)}} \right) \times 100\%$$

2.8 Analyse mécanistique

2.8.1 Capture de conformation chromosomique (3C)

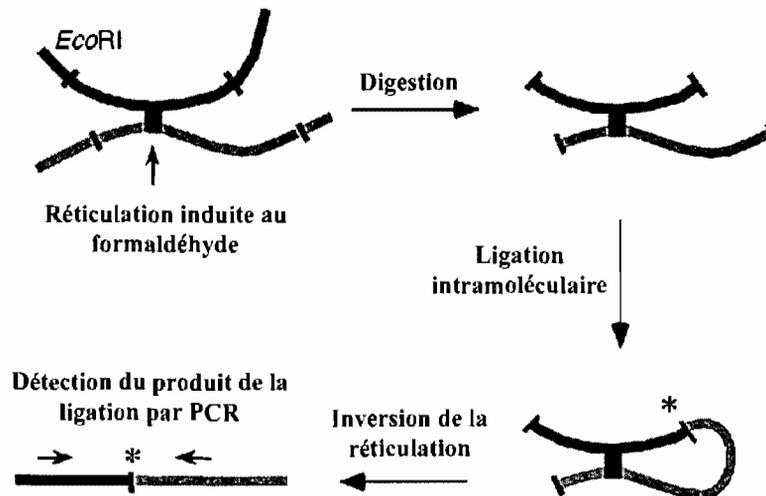


Figure 26: Capture de Conformation Chromosomique. Adapté de Dekker *et al.*⁴⁵⁶

Préparations des échantillons

Foies fœtaux à e12,5

Les foies de fœtus au jour e12,5 qui sont génotypés par électrophorèse du sang sur cellulose acétate tel que décrit à la section 2.5.4, sont récupérés à l'aide de pinces fines sous microscopie et déposés dans 1 mL de milieu essentiel minimum de Eagle modifié par Dulbecco (DMEM) contenant 10% de SFV. Les cellules des foies fœtaux sont séparées mécaniquement par passages répétés dans un embout à l'aide d'une micropipette P-200. Après centrifugation, le culot de cellules est resuspendu dans 3 mL de DMEM/SFV (10%)/formaldéhyde (2%) et incubé pendant 10 min à 25°C. La réaction est arrêtée avec 80 mM de glycine, centrifugée 15 min à 4°C à 400×g puis le culot de cellules est lavé avec 2 mL de PBS. Les échantillons sont regroupés par pools de 3 foies positifs, resuspendus dans 5 mL de tampon de lyse froid [Tris-HCl (10 mM, pH8,9), NaCl (10 mM), Nonidet NP-40 (0,2%), PMSF (0,1 mM), cocktail d'inhibiteurs de protéase (Sigma, 2µL/mL)], incubés pendant 30 min à 4°C, centrifugés 15 min à 500×g pour culotter les noyaux qui sont alors plongés dans l'azote liquide et gardés à -80°C. Un pool de 3 foies est considéré comme $N = 1$ pour les tests statistiques.

Cellules érythroïdes de la moelle osseuse adulte

Les cellules de la moelle osseuse des deux fémurs et des deux humérus de souris adultes sont chassées de la cavité médullaire à l'aide d'une seringue et de 2 ml de DMEM/SFV (10%). Après centrifugation, le culot est rapidement mis sur glace puis lavé avec 11 mL de DMEM/SFV (2%). Le culot est resuspendu dans 500 µL de DMEM/SFV (2%) et les cellules sont comptées à l'hématimètre. Les cellules sont marquées avec 0,2 µg/10⁷ cellules d'anticorps anti-Ter119 couplé à la R-phycoerythrine (PharMingen) et de façon optionnelle avec 0,625 µg/10⁷ cellules d'anticorps anti-CD71 couplé à la fluorescéine isothiocyanate (eBioscience) pendant 30 min à 4°C. Les cellules sont ensuite lavées puis resuspendues dans 2 mL de PBS/SFV (2%). Le culot est resuspendu dans 80 µL de

PBS/SFV (0,5%), dégazé à l'hélium par tranche de 10^7 cellules. Des micro-billes magnétiques *MicroBeads* anti-R-phycoerythrine (Miltenyi Biotec) sont ajoutées ($18 \mu\text{L}/10^7$ cellules) et la suspension est incubée 45 min à $4-8^\circ\text{C}$. Après centrifugation, le culot est resuspendu dans $500\mu\text{L}/10^8$ cellules de PBS/SFV (0,5%)/EDTA (2 mM) et passé à l'*autoMACS* (Miltenyi Biotec) en utilisant le programme de sélection positive *Possel-S*. Les fractions positives ($\text{Ter}119^+$) et négatives ($\text{Ter}119^-$) sont passées sur un appareil d'analyse par cytométrie en flux *FACScan* (Becton Dickinson) pour vérifier l'efficacité de la séparation puis culottées et fixées dans 2 mL de DMEM/SFV (10%)/formaldéhyde (2%) tel que décrit à la section précédente à la différence que chaque culot de noyaux est conservé à -80 séparément. Dans ce cas, une souris donne $N = 1$ pour les tests statistiques.

Génération de la librairie 3C

La génération d'une librairie 3C se fait en deux étapes successives. La première consiste à digérer l'ADN réticulé au formaldéhyde à l'aide d'une endonucléase de restriction et la deuxième étape consiste à la ligation intramoléculaire de l'ADN digéré et à la réversion de la réticulation.

Digestion aux endonucléases de restriction

Les culots de noyaux conservés à -80°C sont resuspendus dans 1 mL de tampon de lyse froid [Tris-HCl (10 mM, pH8,9), NaCl (10 mM), Nonidet NP-40 (0,2%)], culottés par centrifugation puis lysés 1 h à 37°C dans $900 \mu\text{L}$ de tampon REact3 (Invitrogen, 1,1X)/SDS (0,3%). Du Triton X-100 est ajouté à une concentration finale de 1,8% et la solution d'ADN est incubée pendant 1 h à 37°C pour séquestrer le SDS. Des aliquotes de $125 \mu\text{L}$ sont récupérés et à chacun d'eux sont ajoutées 625 U d'enzyme EcoRI à haute concentration (Invitrogen). Les tubes sont incubés à 37°C pendant 16 h puis la réaction est arrêtée en ajoutant du SDS à une concentration finale de 1,6% et en incubant pendant 20 min à 65°C .

Ligation

La solution d'ADN digéré est divisée en aliquotes de 40 μL puis diluée à 800 μL de tampon de ligation [Tris-HCl (30 mM, pH 8,0), MgCl_2 (10 mM), dithiothréitol (10 mM), ATP (1mM)]. Du Triton X-100 est ajouté à une concentration finale de 1,0% et la solution d'ADN est incubée pendant 1 h à 37°C pour séquestrer le SDS. La température est abaissée à 16°C puis 25 Unités Weiss (UW) de ligase à haute concentration (Invitrogen) sont ajoutées. La réaction de ligation est incubée 6-16 h à 16°C puis 250 μg de protéinase K (Roche Diagnostic) sont ajoutés. Les tubes sont incubés 16 h à 65°C. L'ADN est extrait au phénol/chloroforme tel que décrit à la section 2.1.1. L'ADN est resuspendu dans 12,5 μL de TE par tube de ligation et la librairie 3C ainsi générée est conservée à -20°C.

Préparation du contrôle de ligation

Des quantités équimolaires d'ADN génomique de levures contenant le YAC modifié, obtenu par mini-préparation (section 2.2.1.3) et d'ADN de chromosome artificiel de bactérie (BAC) contenant le gène murin *Pkd1*⁵²⁷ sont mises dans un tube, précipitées avec 2 volumes d'éthanol (100%) et 0,1 volume d'acétate de sodium (3M, pH 5,2) pendant 1 h à -80°C. Après un lavage avec 1 volume d'éthanol (70%), l'ADN est digéré avec 200 U d'endonucléase de restriction *EcoRI* à haute concentration pendant 16 h à 37°C dans un volume de 50 μL . La réaction est arrêtée par inactivation de l'enzyme à 65°C pendant 20 min suivi d'une extraction au phénol/chloroforme tel que décrit à la section 2.1.1. Le culot d'ADN est resuspendu dans 50 μL de tampon de ligation, 25 UW de ligase à haute concentration sont ajoutées puis la réaction est incubée pendant 4-5 h à 16°C. L'ADN est de nouveau extrait au phénol/chloroforme puis resuspendu dans 50 μL de TE et conservé à -20°C.

Quantification des librairies 3C

Chaque librairie 3C est quantifiée en triplicata à l'aide d'un spectrophotomètre *Ultrospec 2100 pro* (Biochrom Ltd) par mesure de la densité optique à $\lambda=260$ nm en utilisant le programme interne de quantification de l'ADN.

PCR

La détection des produits du 3C est effectuée par PCR sur une dilution de la librairie 3C (~300 ng/réaction) et du contrôle de ligation (~20 ng) dans les conditions suivantes: Une étape de dénaturation de 5 min à 94°C suivi de 40 cycles d'amplification (94°C-30 s, 62°C-35 s, 72°C-30 s) dans un volume de réaction de 20 µL [Tris (10 mM, pH 8,0), KCl (50 mM), MgCl₂ (1 mM), NH₄Cl (5 mM), DMSO (2,5%), amorces 06-01/amorce complémentaire (0,5 µM chaque), dNTP (87,5 µM), Taq polymérase (0,14 U/µL)] pour toutes les régions du locus sauf pour la région ^Gγ qui est amplifiée dans un tampon moins stringent [Tris (10 mM, pH 8,5), KCl (50 mM), MgCl₂ (1,25 mM), amorces 06-01/06-04 (0,5 µM chaque), dNTP (87,5 µM), Taq polymérase (0,14 U/µL)]. Voir à l'Annexe I,

Tableau XVII, page VI pour la liste des amorces. Les amplicons des librairies 3C et du contrôle de ligation sont déposés sur un même gel d'agarose 2,1% coloré au bromure d'éthidium, révélé aux UV-B (λ=302 nm), digitalisé à l'aide d'un système *Alpha Imager 2200* et quantifié avec le logiciel *ImageQuant v5.0*.

Quantification de la proximité relative

La proximité relative $p(r)$ entre la LCR et une région donnée (région rA) du locus est calculée en divisant le ratio d'intensité de la bande de cette région obtenue de la librairie 3C sur celle obtenue avec le contrôle de ligation par le ratio d'intensité de la bande du pseudogène $\psi\beta$ obtenue de la librairie 3C sur celle obtenue du contrôle de ligation, selon la formule:

$$\text{Équation 2: } p(r) = \left(\frac{rA(3C)/rA(\text{contrôle})}{\psi\beta(3C)/\psi\beta(\text{contrôle})} \right)$$

Cette formule permet de normaliser les résultats obtenus entre les différents échantillons et pour la différence d'efficacité d'amplification entre les différentes régions analysées du locus.

2.9 Méthodes statistiques

La normalité de la distribution des différentes mesures est vérifiée par le test omnibus de normalité de D'Agostino & Pearson lorsque le n est suffisamment élevé. Lorsque le n est insuffisant, la distribution normale est assumée et les valeurs sont présentées en tant que moyennes \pm écarts types à moins de raisons logiques permettant d'assumer une distribution non Gaussienne, auquel cas, les valeurs sont présentées en tant que médianes \pm écarts interquartiles. Lorsque deux groupes sont comparés, le t -test non-pairé et le test de Mann-Whitney sont respectivement utilisés selon que les groupes ont une distribution normale ou non. Lorsque plus de deux groupes sont comparés, un ANOVA à un sens suivi du post-test des comparaisons multiples de Tukey et du post-test de pente ou un test de Kruskal-Wallis suivi du post-test des comparaisons multiples de Dunn sont effectués selon la normalité de la distribution. Les tests pairés de Wilcoxon et de Friedman suivi du post-test des comparaisons multiples de Dunn sont utilisés pour les analyse d'immunoprécipitation de la chromatine, l'ANOVA en mesures répétées suivi du post-test des comparaisons multiples de Tukey dans les analyses d'expression et de 3C, le F -test de pente et le test de non-linéarité de Runs dans les analyses d'expression et le test exact de Fisher dans l'analyse des *Pop-in/Pop-out*. Le test exact de Fisher est également utilisé pour l'évaluation de la transmission mendélienne des gènes KO lors de croisements hétérozygotes alors que pour les croisements homozygotes, le test du χ^2 a été employé. Tous les tests statistiques ont été effectués avec un intervalle de confiance à 95% et un p statistique $< 0,05$ est considéré significatif.

Chapitre 3

Résultats

3.1 Génération des YAC mutants

3.1.1 PBGD^A γ -YAC

Recombinaison homologue dans la levure

Les Pop-in

Après la transformation, 33 clones poussant sur le milieu sans uracile ont été criblés et 23 ont démontré une intégration du YIp telle qu'attendue, 6 clones se sont intégrés via la petite région d'homologie en 5' alors que 4 clones présentaient un patron de digestion aberrant. Il est à noter ainsi que la digestion BspEI n'a pas favorisé l'intégration du YIp en 3' plutôt qu'en 5' puisqu'une proportion de 23 clones sur 29 ayant le patron de digestion attendu n'était pas statistiquement différente de la valeur de 70% attendu d'une intégration aléatoire compte tenu de la longueur des deux bras d'homologie (Fisher: $p = 0,5$)

Deux clones qui démontraient un patron de digestion EcoO65 I correct (Clones *Pi1* et *Pi5*) ont été sélectionnés pour la deuxième étape de recombinaison homologue.

Les Pop-out

Compte tenu de la longueur des deux bras d'homologie du YIp intégré dans le YAC, une efficacité de génération de *Pop-out* mutants de l'ordre de 30% avec 70% de révertants (YAC ayant conservé la séquence sauvage) était attendue. 30 clones *Pop-out* obtenus avec le *Pi1* et 30 autres avec le *Pi5* ont été criblés. Alors que le *Pi5* a donné 8 clones mutants, 20 révertants et 2 clones aberrants ce qui était exactement la proportion attendue (Fisher: $p = 1,0$), le *Pi1* a donné 29 révertants et aucun mutant, démontrant ainsi un fort biais vers l'élimination de la séquence mutante (Fisher: $p < 0,002$). Le clone *Pi1* a donc été abandonné et trois clones *Pop-out* (les clones *Po5-8*, *Po5-11* et *Po5-12*), obtenus du clone *Pi5*, ont été sélectionnés pour être caractérisés en profondeur.

Caractérisation du YAC mutant

Les trois mutants ont été caractérisés par buvardage de type *Southern* en utilisant des sondes réparties le long du locus (Figure 27 et section 2.4.1, page 105). Si les trois clones se sont avérés intègres au niveau du LCR, du gène ϵ et des gènes γ , le clone *Po5-11* a démontré un réarrangement au niveau du gène β et a donc été abandonné.

Une analyse des clones *Po5-8* et *Po5-12* a été effectuée ensuite par *PFGE* permettant de déterminer la présence de larges délétions dans le locus ou ailleurs dans le YAC (Figure 47, Annexe II, page XII). Les deux clones n'ont démontré aucune délétion importante et ont donc été conservés pour la génération des souris transgéniques.

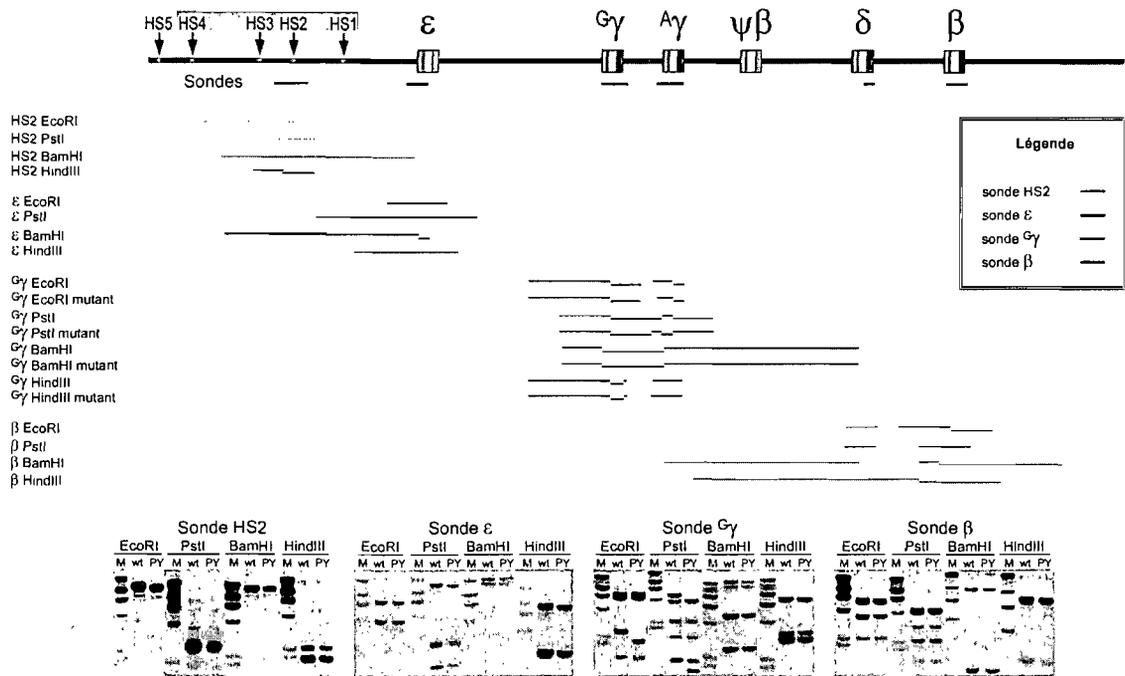


Figure 27: Caractérisation de l'intégrité du mutant $PBGD^{\Delta\gamma}$ -YAC. Analyse par buvardage de type *Southern* des clones *Po5-8*, *Po5-11* et *Po5-12*. Sur la gauche, sont indiquées les différentes digestions utilisées avec les sondes qui sont elles-mêmes indiquées en couleur sous le locus. Les traits de couleur indiquent les fragments attendus pour chaque digestion en fonction de la sonde. Au bas de l'image sont présentés en exemples, les autoradiogrammes du clone $PBGD^{\Delta\gamma}$ -YAC *Po5-12* (PY), en comparaison avec le locus sauvage (wt). M, marqueur de poids moléculaire λ HindIII dont les fragments aperçus correspondent à 23,1 kb, 9,4 kb, 6,5 kb, 4,4 kb, 2,3 kb et 2,0 kb.

3.1.2 Δ -842/-384^A γ -YAC

Recombinaison homologue dans la levure

Les Pop-ins

Après la transformation, 10 clones poussant sur le milieu sans uracile ont été criblés par PCR et tous ont démontré l'intégration du YIp. 3 clones sélectionnés (*Pi29(5)-2-1*, *Pi29(5)-2-3* et *Pi29(5)-2-8*) ont ensuite été caractérisés par buvardage de type *Southern* et tous avaient intégré le YIp au bon endroit et aucun ne présentait de réarrangement. Le clone *Pi29(5)-2-1* a été sélectionné pour générer les *Pop-out*.

Les Pop-out

Compte tenu de la longueur des deux bras d'homologie du YIp intégré dans le YAC, une efficacité de génération de *Pop-out* mutants de l'ordre de 46% avec 54% de révertants était attendue. 15 clones *Pop-out* obtenus avec le *Pi29(5)-2-1* ont été criblés par PCR. 6 clones mutants et 9 révertants ont été obtenus, donc dans la proportion attendue (Fisher: $p = 1,0$). 5 des 6 clones *Pop-out* mutants (les clones *Po29(5)-2-1.1*, *Po29(5)-2-1.2*, *Po29(5)-2-1.7*, *Po29(5)-2-1.9* et *Po29(5)-2-1.13*) ont été sélectionnés pour être caractérisés plus en profondeur.

Caractérisation du YAC mutant

Les 5 mutants sélectionnés ont été caractérisés par buvardage de type *Southern* en utilisant les mêmes sondes réparties le long du locus que pour le PBGD^A γ -YAC (Figure 28 et section 2.4.2, page 106). Tous les clones analysés se sont avérés intègres à l'exception du clone *Po29(5)-2-1.7* qui présentait plusieurs réarrangements dans le locus. Les autres clones ont été conservés pour la génération des souris.

3.1.3 Δ -384/-202^A γ -YAC

Recombinaison homologue dans la levure

Les Pop-ins

Après la transformation, 10 clones poussant sur le milieu sans uracile ont été criblés par PCR et tous ont démontré l'intégration du YIp. 3 clones sélectionnés (*Pi9(14)*-3-2, *Pi9(14)*-3-3 et *Pi9(14)*-3-9) ont été ensuite caractérisés par buvardage de type *Southern* mais seul les clones *Pi9(14)*-3-2 et *Pi9(14)*-3-9 avaient intégré le YIp au bon endroit et ne présentaient aucun réarrangement. Le clone *Pi9(14)*-3-9 a été sélectionné pour générer les *Pop-out*.

Les Pop-out

Compte tenu de la longueur des deux bras d'homologie du YIp intégré dans le YAC, une efficacité de génération de *Pop-out* mutants de l'ordre de 56% avec 44% de révertants était attendue. 15 clones *Pop-out* obtenus avec le *Pi9(14)*-3-9 ont été criblés par PCR. 9 clones mutants, 5 révertants et 1 clone aberrant ont été obtenus ce qui représentait exactement la proportion attendue (Fisher: $p = 1,0$). 4 des 9 clones *Pop-out* mutants (les clones *Po9(14)*-3-9.2, *Po9(14)*-3-9.8, *Po9(14)*-3-9.9 et *Po9(14)*-3-9.10) ont été sélectionnés pour être caractérisés plus en profondeur.

Caractérisation du YAC mutant

Les 4 mutants sélectionnés ont été caractérisés par buvardage de type *Southern* (Figure 28 et section 2.4.3, page 106) et tous les clones analysés se sont avérés intègres à l'exception du clone *Po9(14)*-3-9.8 qui n'avait pas perdu la séquence sauvage de ^A γ . Les clones corrects ont été conservés pour la génération des souris.

3.1.4 Δ -204/-135^A γ -YAC

Recombinaison homologue dans la levure

Les Pop-ins

Après la transformation, 10 clones poussant sur le milieu sans uracile ont été criblés par PCR mais seuls 7 ont démontré l'intégration du YIp. 3 clones sélectionnés (*Pi4(2)*-15-2, *Pi4(2)*-15-5 et *Pi4(2)*-15-8) ont été ensuite caractérisés par buvardage de *Southern* mais seul les clones *Pi4(2)*-15-2 et *Pi4(2)*-15-5 se sont avérés avoir intégré le YIp à l'endroit prévue alors que le clone *Pi4(2)*-15-8 avait intégré le YIp dans la région 5' du promoteur de ^A γ .

Les Pop-out

Compte tenu de la longueur des deux bras d'homologie du YIp intégré dans le YAC, une efficacité de génération de *Pop-out* mutants de l'ordre de 58% et 42% de révertants était attendue. Dans un premier temps, le clone *Pi4(2)*-15-5 a été sélectionné pour générer les *Pop-out* et 15 clones obtenus ont été criblés par PCR. 4 clones mutants et 11 révertants ont été obtenus ce qui, bien que s'éloignant de la proportion attendue, ne représentait pas une différence statistiquement significative (Fisher: $p = 0,1$). Les 4 clones *Pop-out* mutants (*Po4(2)*-15-5.6, *Po4(2)*-15-5.10, *Po4(2)*-15-5.12 et *Po4(2)*-15-5.13) ont été sélectionnés pour la caractérisation subséquente. Toutefois cette analyse a démontré qu'aucun de ces clones n'avait perdu la séquence sauvage du gène ^A γ .

Un deuxième tour de génération de *Pop-out* en utilisant les *Pop-in* *Pi4(2)*-15-2 et *Pi4(2)*-5-8 a donc été effectué. 15 clones obtenus pour chacun des deux *Pop-in* utilisés ont été criblés par PCR. 12 clones mutants et 3 révertants ont été obtenus du clone *Pi4(2)*-15-2 alors que 4 clones mutants et 10 révertants ont été obtenus du clone *Pi4(2)*-5-8. Encore une fois, ces résultats ne divergeaient pas significativement des proportions attendues de 58% et 42% de mutants des *Pop-in* *Pi4(2)*-15-2 et *Pi4(2)*-5-8 respectivement (Fisher: $p = 0,4$; $p =$

0,7). 8 clones mutants (*Po4(2)-15-2.1*, *Po4(2)-15-2.2*, *Po4(2)-15-2.3*, *Po4(2)-15-2.5*, *Po4(2)-15-8.7*, *Po4(2)-15-8.10*, *Po4(2)-15-8.11* et *Po4(2)-15-8.12*) ont été sélectionnés pour être caractérisés.

Caractérisation du YAC mutant

Les 8 mutants sélectionnés ont été caractérisés par buvardage de type *Southern* (Figure 28 et section 2.4.4, page 107) et tous les clones analysés se sont avérés intègres à l'exception du clone *Po4(2)-15-8.12* qui n'avait pas perdu la séquence sauvage de $\Delta\gamma$. Les autres clones ont été conservés pour la génération des souris.

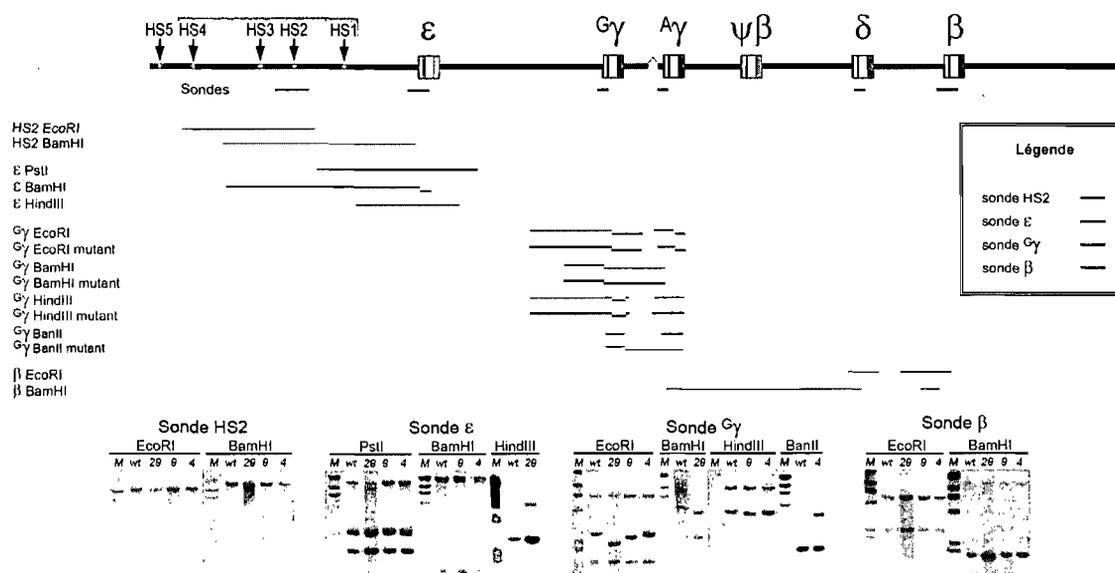


Figure 28: Caractérisation de l'intégrité des mutants $\Delta\gamma$ -YAC. Analyse par buvardage de type *Southern* des clones *Pop-out* pour les trois mutations de délétion. Sur la gauche, sont indiquées les différentes digestions utilisées avec les sondes qui sont indiquées en couleur sous le locus. Les traits de couleur indiquent les fragments attendus pour chaque digestion en fonction de la sonde. Au bas de l'image sont présentés en exemples, les autoradiogrammes des clones Δ -842/-384 $\Delta\gamma$ -YAC *Po29(5)-2-1.2* (**29**), Δ -384/-202 $\Delta\gamma$ -YAC *Po9(14)-3-9.2* (**9**) ainsi que Δ -204/-135 $\Delta\gamma$ -YAC *Po4(2)-15-8.7* (**4**), comparés avec le locus sauvage (wt). M, marqueur de poids moléculaire λ HindIII dont les fragments aperçus correspondent à 23,1 kb, 9,4 kb, 6,5 kb, 4,4 kb, 2,3 kb et 2,0 kb.

3.2 Génération des souris transgéniques

3.2.1 Les souris PBGD^A γ -YAC

Le YAC *Po5-8* (Section 3.1.1, page 129) a été sélectionné pour la microinjection.

Sur les 17 fondatrices obtenues, 7 avaient intégré de 1 à 10 copies entières du locus (Figure 29). Les autres fondatrices démontraient soit des délétions de portions importantes du locus ou des réarrangements dans différentes régions, en plus de présenter dans certains cas une intégration mosaïque caractérisé par un nombre de copies inférieur à 1 (Tableau XVIII, Annexe II, page XIII). Une analyse rapide de l'expression protéique des gènes du transgène sur *UT-PAGE* a permis de détecter les trois chaînes humaines ^A γ , ^G γ et β chez les fondatrices ayant intégré au moins une copie complète du locus alors que les souris qui avaient intégré un locus tronqué, présentaient une expression variable des gènes du locus (Tableau XVIII, Annexe II, page XIII).

Il est à noter, que la lignée PY452 qui possédait entre 5 et 10 copies du YAC ne transmettait le transgène que très difficilement et n'a donc pas été utilisée pour les analyses subséquentes.

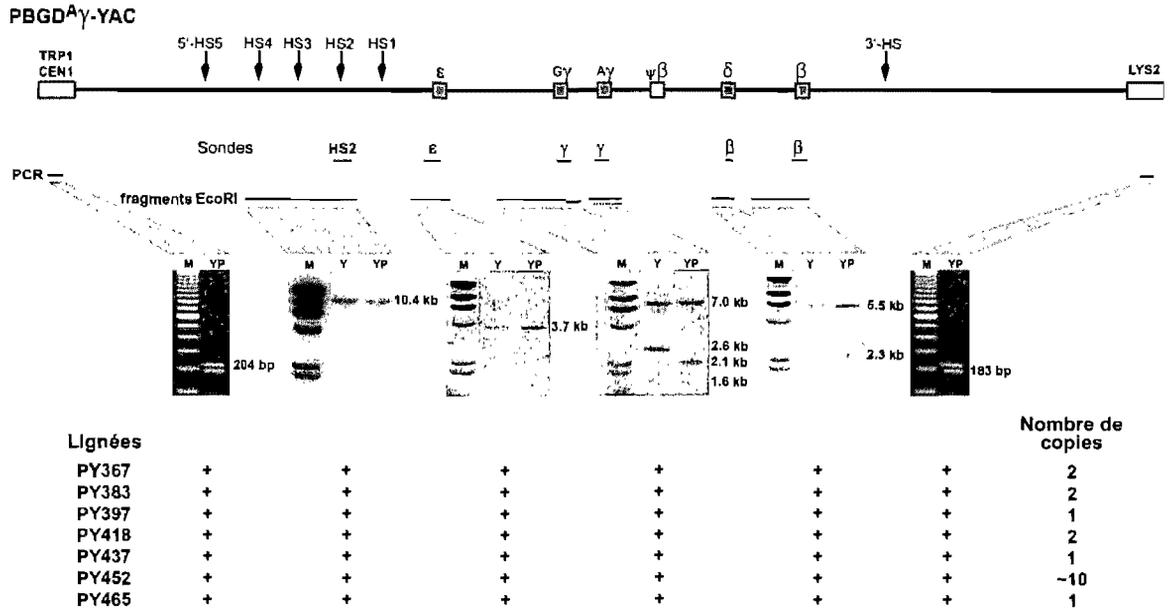


Figure 29: Caractérisation des fondatrices PBGD^Aγ-YAC. Exemples représentatifs de PCR et d'autoradiogrammes démontrant l'intégrité des différentes régions du locus. Les 7 fondatrices ayant le locus intègre sont également présentées (pour une représentation plus détaillée des 17 fondatrices, se référer au Tableau XVIII à l'Annexe II, page XIII). Y: Souris β-YAC ayant le locus humain sauvage; PY: fondatrice de la lignée PY418. M, marqueur de poids moléculaire

3.2.2 Les souris Δ-842/-384^Aγ-YAC

Le YAC *Po29(5)-2-1.2* (Section 3.1.2, page 132) a été sélectionné pour la microinjection.

Les trois fondatrices ΔL7, ΔL11 et ΔL49 avaient intégré une seule copie du locus entier (Tableau VIII). Une analyse rapide de l'expression protéique sur *UT-PAGE*, a permis de détecter la chaîne β chez les trois fondatrices (Tableau VIII).

Tableau VIII: Caractérisation des fondatrices Δ -841/-384^A γ -YAC.

lignées	copies	Intégrité							Expression		
		B.G.	HS2	ϵ	γ	δ	β	B.D.	$\epsilon\gamma$	$\Delta\gamma$	β
Δ L7	1	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	-	-	+				
Δ L11	1	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	-	-	+				
Δ L49	1	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	-	-	+				

Légende du tableau: L'intégrité des différentes régions du locus a été vérifiée par PCR pour les bras gauche (B.G.) et droit (B.D.) du YAC et par buvardage de type *Southern* pour les autres régions: : présent, : présent et intègre. Expression: expression protéique des chaînes de globine au stade adulte, +: expression aisément détectable, -: expression non-détectable.

3.2.3 Les souris Δ -384/-202^A γ -YAC

Le YAC *Po9*(14)-3-9.9 (Section 3.1.3, page 133) a été sélectionné pour la microinjection.

Sur les 3 fondatrices qui ont intégré entre 1 et 2 copies du transgène, seule la souris Δ M56 avait intégré le locus entier alors que les deux autres fondatrices présentaient d'importants réarrangements ou des délétions dans le locus (Tableau IX). Une analyse rapide de l'expression protéique sur *UT-PAGE* a confirmé que seule la souris Δ M56 exprimait la chaîne β au stade adulte, la fondatrice Δ M42 n'exprimant aucune globine. Pour ce qui est de la souris Δ M53 qui était d'haplotype *s/d*, la chaîne β n'a pu être détectée à cause de la présence de la β^{single} endogène. (Tableau IX).

Tableau IX: Caractérisation des fondatrices Δ -384/-202^A γ -YAC.

lignées	copies	Intégrité					Expression		
		HS2	ϵ	$\zeta\gamma$	$\Lambda\gamma$	β	$\zeta\gamma$	$\Lambda\gamma$	β
Δ M42	1	☑	☒	×	☒	×	-	-	-
Δ M53	2	☑	☒	☒	☑	☑	-	-	
Δ M56	2	☑	☑	☑	☑	☑	-	-	+

Légende du tableau: L'intégrité des différentes régions du locus a été vérifiée par buvardage de type *Southern*: ☑: présent et intègre, ☒: présent mais non-intègre (bandes de la mauvaise longueur), ×: absent. Expression: expression protéique des chaînes de globine au stade adulte, +: expression aisément détectable, -: expression non-détectable.

3.3 Analyses d'expression dans les souris transgéniques

3.3.1 Les souris PBGD^A γ -YAC

Pour évaluer l'impact du remplacement du promoteur de la ^A γ -globine sur la commutation de l'hémoglobine fœtale à adulte, l'expression des différents gènes a été mesurée à différents stades du développement, tant au niveau de l'ARN qu'au niveau des protéines.

Expression au niveau ARN

Une analyse par PCR quantitatif en temps réel des gènes embryonnaires, fœtaux et adultes du locus a été effectuée sur trois lignées de souris PBGD^A γ -YAC (PY367, PY418 et PY465) ainsi que sur la lignée contrôle β -YAC qui possède le locus sauvage humain. Cette analyse a été effectuée à différents stades du développement du jour embryonnaire e10,5 au stade adulte (Figure 30). Les trois lignées PBGD^A γ -YAC n'ont démontré aucune différence d'expression pour tous les gènes du locus durant le développement et ce, jusque chez l'adulte (ANOVA: $p > 0,4$).

ε-globine

L'expression du gène embryonnaire ϵ a été mesurée en fonction des globines humaines totales à e10,5 dans le sac vitellin, e12,5, e14,5 et e16,5 dans le foie fœtal, à la naissance dans la rate et chez l'adulte dans la moelle osseuse (Figure 30A-B). L'expression de ϵ à tous les stades du développement dans les lignées PBGD^A γ -YAC, ne différait pas significativement de l'expression observée avec le locus sauvage de la souris β -YAC (ANOVA: $F_{(3,20)} = 1,2$; $p = 0,33$) indiquant de ce fait, que la régulation de ce gène était indépendante de la substitution du promoteur de la ^A γ -globine par le promoteur PBGD.

^A γ - et ^G γ -globine

Utilisant la présence du 5'UTR de PBGD fusionné au messenger ^A γ , les deux gènes fœtaux ont pu être analysés séparément. Telle qu'attendu du promoteur stade-indépendant de PBGD, l'expression de ^A γ dans les souris PBGD^A γ -YAC a été mesurée à un niveau stable tout au long du développement (de e10,5 au stade adulte) ce qui a été démontré par la linéarité de la courbe d'expression (test de non-linéarité de Runs: $p = 0,4$) et par la pente de cette dernière ne déviant pas significativement de la valeur 0 (F -test de pente: $p = 0,34$) (Figure 30B). Cette expression soutenue différait de la répression de l'expression des deux gènes γ , observée dès e12,5 et de l'extinction complète de cette expression vers e14,5 chez la souris β -YAC (Figure 30A). Il est à noter que l'expression du gène non modifié ^G γ devait servir de contrôle interne et était sensé suivre celle du gène équivalent du locus sauvage. Dans les faits toutefois, son expression s'avéra significativement dérégulée. En effet, si à e10,5 l'expression de ^G γ semblait normale, à e12,5 et e14,5 elle était activée 1,5 fois par rapport au niveau mesuré à e10,5 alors que dans le locus sauvage, l'expression mesurée à e12,5 et e14,5 était diminuée de 10 fois et 20 fois respectivement. Finalement juste avant la naissance, une répression progressive de ^G γ a été mesurée chez les souris PBGD^A γ -YAC sans toutefois jamais atteindre l'inactivation complète et ce, même plusieurs mois après la naissance (Figure 30A-B). Cette commutation incomplète des gènes fœtaux a été observée

pour les trois lignées PBGD^Δγ-YAC et différait significativement de celle observée chez les souris β-YAC, qui était complétée dès e14,5 (ANOVA: $F_{(3,20)} = 10,78$; $p \leq 0,0005$).

δ-globine

De même que les gènes fœtaux, les gènes adultes ont démontré un patron d'expression et une cinétique d'activation altérés. Une analyse de l'expression des gènes δ et β en utilisant des amorces spécifiques a été effectuée à tous les stades du développement. À e10,5, l'expression de δ était virtuellement indétectable même après 45 cycles d'amplification par PCR aussi bien pour les souris PBGD^Δγ-YAC que pour les souris β-YAC. De façon intéressante, chez la souris β-YAC, δ-globine était exprimé à un niveau faible mais stable de e12,5 à la naissance, ne représentant toutefois que moins d'un millième de l'expression de β à ces stades. Chez l'adulte, l'expression de δ augmentait de 10 fois pour atteindre les 1-2% normalement attendus (Figure 30C). Dans les souris PBGD^Δγ-YAC, l'expression de δ a été détectée à un niveau stable du jour embryonnaire e12,5 au stade adulte. Cependant, au stade fœtal, le niveau d'expression de δ était ~10 fois inférieur à celui mesuré chez la souris β-YAC. Étant donné que δ n'était pas activé au stade adulte chez les souris PBGD^Δγ-YAC, la différence de ratio δ/β était encore plus grande, ne représentant plus que $\sim 10^{-2}$ de la valeur mesurée chez les β-YAC, démontrant de ce fait que le gène PBGD^Δγ interférait avec le niveau et l'activation ontogénique de δ.

β-globine

L'expression du gène adulte β à e10,5 a été estimée à ~5% des chaînes totales dans les souris β-YAC alors que des niveaux très bas < 0,1% ont été détectés dans les souris PBGD^Δγ-YAC (Figure 30A-B) suggérant un impact du promoteur PBGD sur ce gène dès le stade embryonnaire. Chez les souris β-YAC, dès e12,5 l'expression de β prédominait déjà sur tous les autres gènes et représentait près de 80% des globines humaines. Contrastant avec ce résultat, à e12,5 chez les souris PBGD^Δγ-YAC, β ne faisait que commencer à être

activé. Ce délai observé dans toutes les lignées PBGD^A γ -YAC de e12,5 à la naissance correspondait en fait à une diminution de cinq fois de la cinétique d'activation du gène β par rapport à celle mesurée dans les souris β -YAC. À la naissance, les niveaux d'expression de β (~65%) étaient sous les niveaux mesurés à e12,5 chez la souris β -YAC, démontrant de ce fait un retard de plus de 8 jours. En fait, le niveau maximal de β (80%) n'a été atteint qu'au stade adulte. Toutefois, bien que le patron d'expression de β était sévèrement affecté (ANOVA: $F_{(3,20)} = 11,70$; $p = 0,0005$), l'expression de δ l'était encore plus, suggérant que le promoteur PBGD avait un impact important sur l'expression des deux gènes adultes.

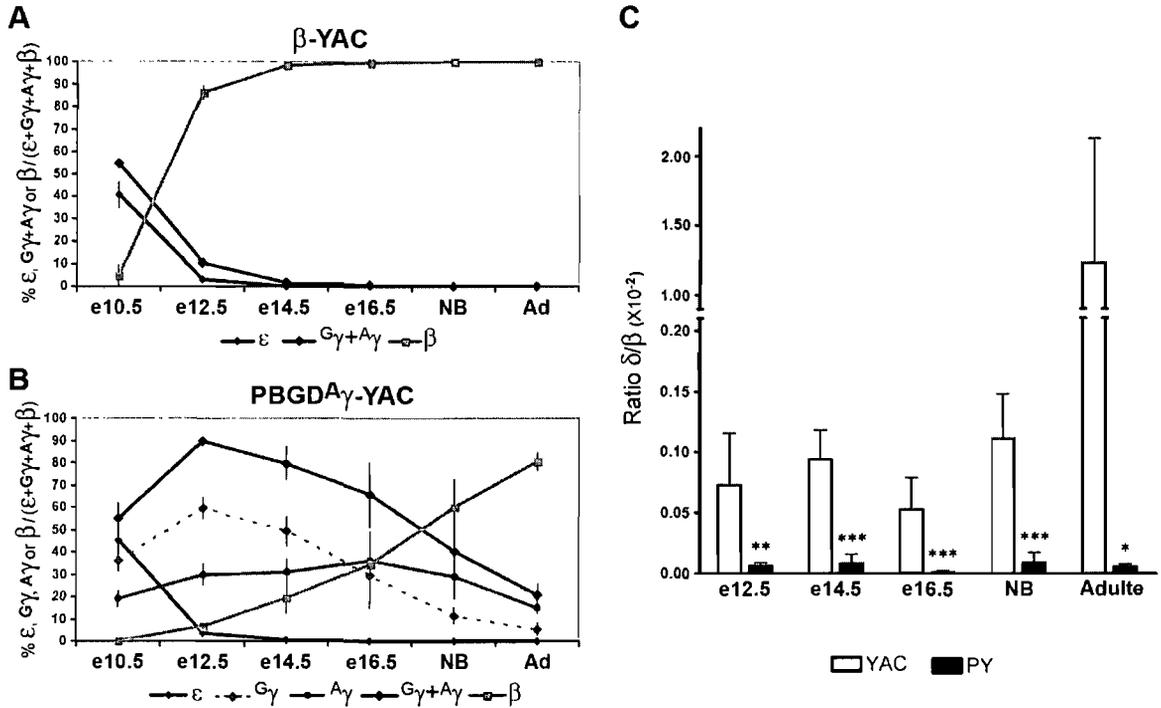


Figure 30: Expression des gènes du locus $PBGD^{A\gamma}$ -YAC (qPCR). L'expression des différents gènes du transgène a été mesurée par RT-PCR quantitatif en temps-réel aux différents stades du développement chez les fœtus de e10,5 à e16,5, chez les nouveau-nés (NB) et chez les adultes (Ad). (A) Patron d'expression des gènes humains ϵ , γ et β dans la souris contrôlée β -YAC. Les résultats représentent la moyenne \pm l'écart type ($N = 3$ pour chaque stade fœtal et $N = 2$ pour les souris adultes). (B) Patron d'expression des gènes ϵ , $G\gamma$, $A\gamma$ et β dans la souris $PBGD^{A\gamma}$ -YAC. Les résultats représentent la moyenne de trois lignées \pm l'écart type ($N = 8$ pour chaque stade fœtal, $N = 7$ pour les nouveau-nés et $N = 8$ pour les adultes; 2-3 souris pour chacune des lignées PY367, PY418 et PY465). (C) Ratio de l'expression du gène δ sur celle du gène β chez les souris β -YAC (YAC) et la moyenne de trois lignées $PBGD^{A\gamma}$ -YAC (PY). Le p statistique a été déterminé à l'aide du t -test de Student: *: $p < 0,002$, **: $p < 0,0005$, ***: $p < 0,0001$.

Expression protéique

Parallèlement à l'analyse du patron d'expression au niveau de l'ARN des globines, le patron d'expression protéique des chaînes de globine du sang périphérique a été mesuré à tous les stades du développement chez les souris β -YAC et les trois lignées PBGD^A γ -YAC ($N = 2-5$ par lignée) (Figure 31A-B) incluant le stade juvénile de p3 à p21 chez les lignées PY367 et PY465 (Figure 31C).

L'expression protéique des souris β -YAC et PBGD^A γ -YAC a démontré ainsi un patron similaire à celui de l'ARN avec un délai toutefois de 4-20 jours dû à la demi-vie accrue des érythrocytes définitifs dans le sang périphérique. Une expression soutenue de la protéine ϵ a été mesurée jusqu'à e14,5 indiquant qu'à cette période, le sang périphérique était toujours constitué surtout d'érythrocytes primitifs produits par le sac vitellin. Par la suite, l'expression de ϵ diminuait à des niveaux à peine détectables dès le jour e16,5 indiquant qu'à ce stade, la commutation cellulaire avait eu lieu et que les cellules circulantes étaient désormais essentiellement d'origine définitive.

Contrairement aux analyses d'ARN, l'analyse de l'expression protéique a permis de distinguer les chaînes $G\gamma$ et $A\gamma$ à la fois chez les PBGD^A γ -YAC et chez les souris contrôles β -YAC (Figure 31A-B). Au jour e10,5, les lignées PBGD^A γ -YAC démontraient des niveaux d'expression de $A\gamma$ et $G\gamma$ similaires à ceux observés dans la lignée β -YAC. Toutefois, dès e12,5 le patron d'expression protéique des chaînes fœtales chez les souris PBGD^A γ -YAC commençait à différer de celui mesuré chez la souris β -YAC. Chez cette dernière, si à e12,5 les niveaux de la protéine $G\gamma$ étaient toujours légèrement plus élevés que ceux de $A\gamma$, à e14,5 ces niveaux étaient devenus égaux (Figure 31A). Par la suite, les chaînes fœtales étaient remplacées graduellement par le chaîne adulte β et dès e16,5, les deux chaînes γ avaient atteint des niveaux à peine détectables. Chez les souris PBGD^A γ -YAC, $G\gamma$ était fortement activé comparativement à $A\gamma$ dès le jour e12,5 et demeurait ainsi la chaîne de globine majoritaire du sang périphérique jusqu'au jour postnatal \sim p10 (Figure

31B-C). Notablement, même à l'âge de plusieurs mois, les deux chaînes γ représentaient 40-50% des chaînes humaines de globine dans le sang périphérique, une contribution bien supérieure à celle mesurée au niveau transcriptionnel dans la moelle osseuse ($p < 10^{-10}$).

De façon concomitante à l'expression soutenue des chaînes fœtales tout au long du développement, un délai de l'activation de β était observable dans les souris PBGD^A γ -YAC comparativement au souris β -YAC. De plus, tout comme au niveau de l'ARN, la cinétique de cette activation était fortement altérée. Le niveau maximal d'expression a été mesuré chez l'adulte mais est demeuré inférieur à celui mesuré au niveau de l'ARN. Alors que la chaîne δ pouvait être détectée chez les souris adultes β -YAC, elle est demeurée indétectable chez les souris PBGD^A γ -YAC à tous les stades du développement suggérant une expression de ce gène sévèrement affectée.

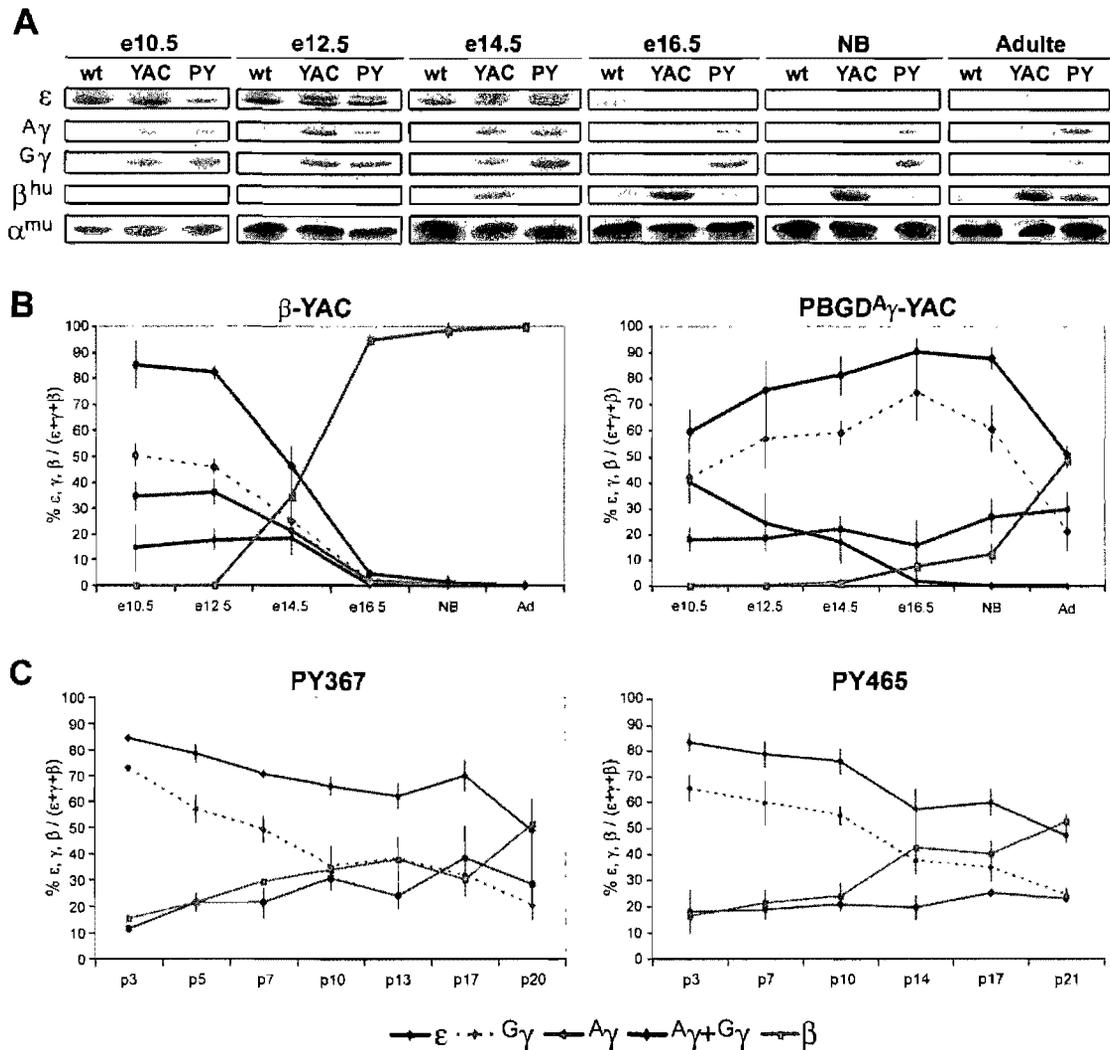


Figure 31: Analyse des chaînes de globine dans les souris $PBGD^{A\gamma}$ -YAC. (A) Les chaînes de globine du sang périphérique des souris sauvages (wt), β -YAC (YAC) et $PBGD^{A\gamma}$ -YAC (PY) à différents stades du développement ont été séparées par *UT-PAGE* pour permettre la discrimination des chaînes $G\gamma$ et $A\gamma$. (A-B) Le patron d'expression au cours du développement (B) et après la naissance (C) des chaînes humaines de globine est rapporté en fonction du pourcentage des chaînes humaines totales. Les résultats représentent la moyenne de \pm l'écart type de 2-5 souris pour les β -YAC et de 9-11 souris provenant de trois lignées (2-5 souris pour chacune des trois lignées PY367, PY418 et PY465) en (B) et 2-4 souris pour chacune des deux lignées PY367 et PY465 en (C). NB: nouveau-né. Ad: adulte.

Distribution cellulaire de l'expression

Pour vérifier si l'expression concomitante de γ et β observée chez les souris adultes PBGD ^{γ} -YAC était due à une expression pancellulaire des deux gènes ou à une expression de chaque gène dans des populations distinctes de cellules, le sang périphérique des sept lignées a été analysé par cytométrie en flux en utilisant des anticorps contre les hémoglobines HbA et HbF. Ces anticorps reconnaissent spécifiquement les tétramères hybrides $\alpha_2^{\text{murin}}\beta_2^{\text{humain}}$ et $\alpha_2^{\text{murin}}\gamma_2^{\text{humain}}$ respectivement sans toutefois reconnaître les hémoglobines murines comme cela est démontré par l'analyse du sang de la souris sauvage (Figure 32). L'expression de la β -globine humaine (HbA⁺) au stade adulte a été détectée dans 70-80% des érythrocytes de toutes les lignées transgéniques, y compris la souris β -YAC. L'expression de la γ -globine humaine (HbF⁺) a été observée dans 85-100% des érythrocytes des sept lignées PBGD ^{γ} -YAC. Le nombre de cellules exprimant γ semblait être influencé par le nombre de copies du transgène puisque les lignées ayant le plus grand nombre de copies présentaient la plus grande proportion de cellules exprimant γ alors que de façon réciproque, les lignées ayant le moins de copies du transgène exhibaient un déplacement de l'intensité du signal de fluorescence dans une petite proportion de cellules mais qui demeurant sous le seuil de détection défini (Figure 32).

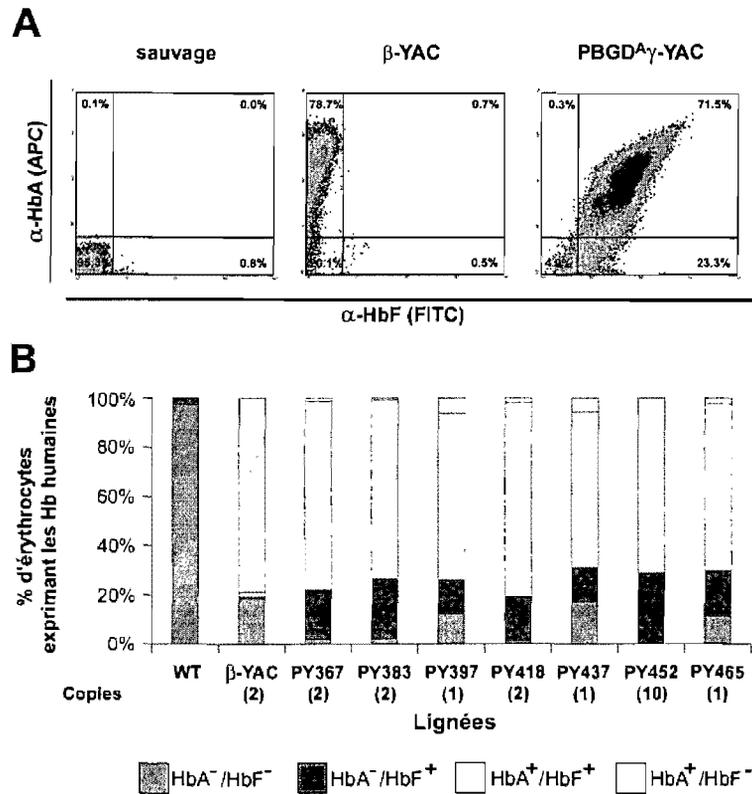


Figure 32: Distribution cellulaire de l'expression de γ et β dans les érythrocytes matures adultes. (A) Exemples représentatifs de graphiques obtenus des analyses par cytométrie en flux de souris sauvages F1(C57Bl/6J X CBA/J), β -YAC et $\text{PBGD}^{\text{A}}\gamma$ -YAC. Des anticorps spécifiques ont été utilisés pour reconnaître spécifiquement les hémoglobines hybrides HbF ($\alpha_2^{\text{murin}}\gamma_2^{\text{humain}}$) et HbA ($\alpha_2^{\text{murin}}\beta_2^{\text{humain}}$) sans détecter les hémoglobines endogènes. (B) Histogrammes de la distribution des cellules exprimant la HbA et/ou la HbF au stade adulte. Les 7 lignées $\text{PBGD}^{\text{A}}\gamma$ -YAC ont été ainsi comparées aux contrôles sauvages et β -YAC. Le pourcentage d'érythrocytes totaux est représenté par la moyenne de 2-4 souris pour chaque lignée.

3.3.2 Les souris PBGD^{Δγ}-YAC/EKLF-nulles

Pour éliminer l'influence de l'activation de β dans la commutation tardive de l'hémoglobine fœtale à adulte observée chez les souris PBGD^{Δγ}-YAC et ainsi déterminer l'étendue de la fonction des gènes fœtaux dans la régulation de cette fonction, le gène EKLF qui est essentiel pour l'expression de β mais dispensable à l'expression des gènes γ a été éliminé par ablation génique dans les souris PBGD^{Δγ}-YAC.

Expression au niveau ARN

Deux lignées PBGD^{Δγ}-YAC ont été croisées avec des souris hétérozygotes pour l'allèle nul du facteur EKLF. Puisque l'ablation homozygote de ce gène est létale à partir de e14,5, les fœtus obtenus de croisements homozygotes PBGD^{Δγ}-YAC/EKLF^{+/-} X EKLF^{+/-} ont été prélevés tôt le matin du 14^e jour embryonnaire puis l'expression des gènes de globine humaine a été mesurée (Figure 33). L'analyse de l'expression des gènes fœtaux et adultes à e14,5 a démontré tel qu'attendu, que la médiane de l'expression de β diminuait de façon proportionnelle à la disparition de EKLF dans les souris hémizygotés et homozygotes pour l'ablation du gène EKLF, passant de 22% des globines humaines chez les souris EKLF^{+/+} à 0,02% chez les EKLF^{-/-} (Kruskal-Wallis: $p < 0,002$). Contrairement à ce qui a été observé avec le gène β , le gène $\Delta\gamma$, sous le contrôle du promoteur PBGD, n'était influencé ni par l'absence de EKLF ni par l'inactivation de β puisque son expression est demeurée stable à ~25% (Kruskal-Wallis: $p > 0,3$) alors que $\Delta\gamma$ présentait une augmentation proportionnelle à la disparition de EKLF passant de 48% chez les EKLF^{+/+} à 76% chez les EKLF^{-/-} (Kruskal-Wallis: $p < 0,005$). Ces résultats démontraient ainsi que l'inactivation de β et/ou la disparition de EKLF pouvaient contrer l'inactivation tardive du gène $\Delta\gamma$ dans le locus PBGD^{Δγ}-YAC.

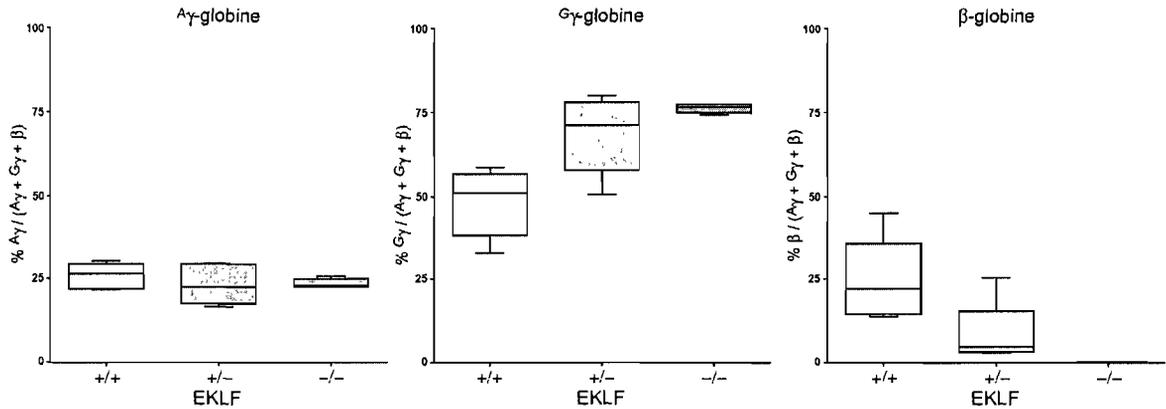


Figure 33: Expression des gènes de globine chez les fœtus PBGD^{Aγ}-YAC/EKLF-nuls. L'ARNm du foie fœtal à e14,5 a été quantifié par RT-PCR quantitatif en temps-réel pour l'expression des gènes G^γ, A^γ et β chez les fœtus PBGD^{Aγ}-YAC⁺ et EKLF^{+/+} (+/+, N = 7), EKLF^{+/-} (+/-, N = 6) et EKLF^{-/-} (-/-, N = 5). L'expression mesurée chez les fœtus de deux lignées PBGD^{Aγ}-YAC (PY367 et PY465) est rapportée en fonction du pourcentage des globines humaines et est représentée par la médiane des deux lignées ± les écarts interquartiles pour chacun des génotypes.

3.3.3 Les souris PBGD^{Aγ}-YAC/Ikaros-nulles

Puisqu'il a été suggéré que le facteur Ikaros et le complexe PYR étaient impliqués de façon indépendante des promoteurs fœtaux dans la régulation de la commutation de l'hémoglobine fœtale à adulte, j'ai décidé d'évaluer l'apport du facteur Ikaros dans la commutation tardive de l'hémoglobine observée chez les souris PBGD^{Aγ}-YAC. Pour ce faire, Ikaros a été éliminé par ablation génique dans les souris PBGD^{Aγ}-YAC.

Expression protéique

Deux lignées PBGD^{Aγ}-YAC ont été croisées avec des souris hétérozygotes pour l'allèle nul du gène endogène Ikaros. Puisque les souris Ikaros^{-/-} survivent difficilement, le sang périphérique de souris âgées de 3 semaines a été prélevé et les chaînes de globine ont été séparées par *UT-PAGE*. (Figure 33). Il est à noter que l'obtention de souris Ikaros^{-/-} s'est

avérée difficile puisque seulement 7 souris (dont 4 étaient négatives pour le transgène PBGD^A γ -YAC) sur 106 étaient de génotype Ikaros^{-/-}. Cette proportion était significativement inférieure à la proportion de 25% attendue d'une transmission mendélienne ($\chi^2_{(2, 13,49)}$; $p < 0,002$) suggérant une mortalité survenant avant le moment du sevrage ou possiblement *in utero*. Ce biais n'était observable toutefois que pour la perte homozygote du gène Ikaros puisque avec 27 souris Ikaros^{+/-} sur 59 souris obtenues de croisements où un seul parent était porteur de l'allèle nul, la transmission ne différait pas des 50% attendus d'une transmission mendélienne (Fisher: $p = 0,7$). L'ablation du facteur Ikaros n'a eu qu'un faible impact sur la commutation de l'hémoglobine fœtale à adulte (Figure 34). En effet, si la proportion de l'expression de PBGD^A γ présentait une augmentation modeste mais significative, passant de ~25% chez les souris Ikaros^{+/+} et Ikaros^{+/-} à ~39% chez les Ikaros^{-/-} ($p < 0,0005$), les proportions de $G\gamma$ et β ne présentaient pas de différences significatives en absence d'Ikaros. Il est à noter que l'ablation génique d'Ikaros n'a pas eu d'effet significatif sur l'expression des deux chaînes adultes endogènes de la souris β^{maj} et β^{min} (Figure 48A, Annexe II, page XIV). Toutefois, les niveaux d'expression des chaînes humaines étaient significativement diminués, tant pour $G\gamma$ que β , alors que l'expression de $A\gamma$ ne présentait qu'une légère diminution statistiquement non significative (Figure 48B, Annexe II, page XIV) expliquant de ce fait, l'apparente activation de $A\gamma$ en comparaison aux deux autres chaînes humaines. Ces résultats suggéraient que si Ikaros ne joue pas un rôle déterminant dans la commutation de l'hémoglobine dans le contexte du locus PBGD^A γ -YAC, il pourrait être important pour l'expression maximale des gènes de globine.

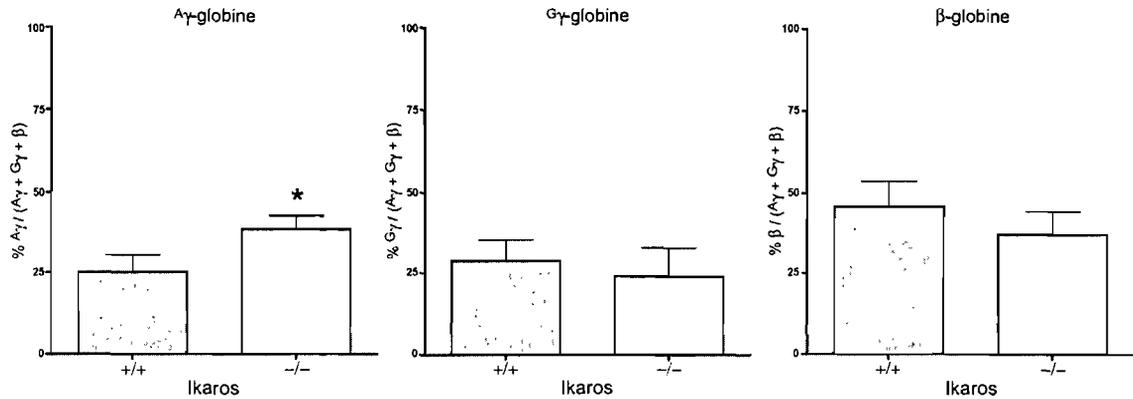


Figure 34: Expression des chaînes de globine chez les souris $PBGD^{\Delta\gamma}$ -YAC/Ikaros-nulles. Les chaînes $G\gamma$ -, $A\gamma$ - et β -globine du sang périphérique ont été quantifiées par *UT-PAGE* chez des souris de trois semaines $PBGD^{\Delta\gamma}$ -YAC⁺ sur fond Ikaros^{+/+} et $+/+$ ($+/+$, $N = 19$) et Ikaros^{-/-} ($-/-$, $N = 3$). L'expression mesurée chez les deux lignées $PBGD^{\Delta\gamma}$ -YAC (PY367 et PY465) est rapportée en fonction du pourcentage des globines humaines et est représentée par la moyenne des deux lignées \pm l'écart type. Le p a été déterminé à l'aide du t -test de Student: *, $p < 0,0005$.

3.3.4 Les souris Δ -842/-384 $A\gamma$ -YAC

Pour évaluer l'impact de la délétion de la région comprise entre -842 et -384 pb du site d'initiation de la transcription du gène $A\gamma$ sur la commutation de l'hémoglobine fœtale à adulte, l'expression au niveau de l'ARN des différents gènes a été mesurée par PCR quantitatif en temps-réel à divers stades du développement chez les souris Δ -842/-384 $A\gamma$ -YAC.

Comme il peut être observé à la Figure 35A, les patrons d'expression des gènes ϵ , γ et β à tous les stades du développement, du jour embryonnaire e10,5 à l'âge de sevrage (\sim 3 semaines), ne différaient pas de façon significative des patrons d'expression de ces gènes dans le locus sauvage β -YAC. De plus, lorsque mesuré au niveau protéique au jours e12,5, si le ratio des chaînes fœtales $A\gamma/G\gamma$ semblait présenter une légère augmentation dans deux des trois lignées Δ -842/-384 $A\gamma$ -YAC (Figure 35B), cette variation n'était toutefois pas

statistiquement significative (ANOVA: $F_{(3,17)} = 1,9$; $p = 0,18$) indiquant que la délétion de cette région du promoteur de la $A\gamma$ -globine n'influence l'expression ni du gène $A\gamma$ lui-même, ni des autres gènes du locus et n'a donc pas d'impact sur la régulation de la commutation de l'hémoglobine fœtale à adulte.

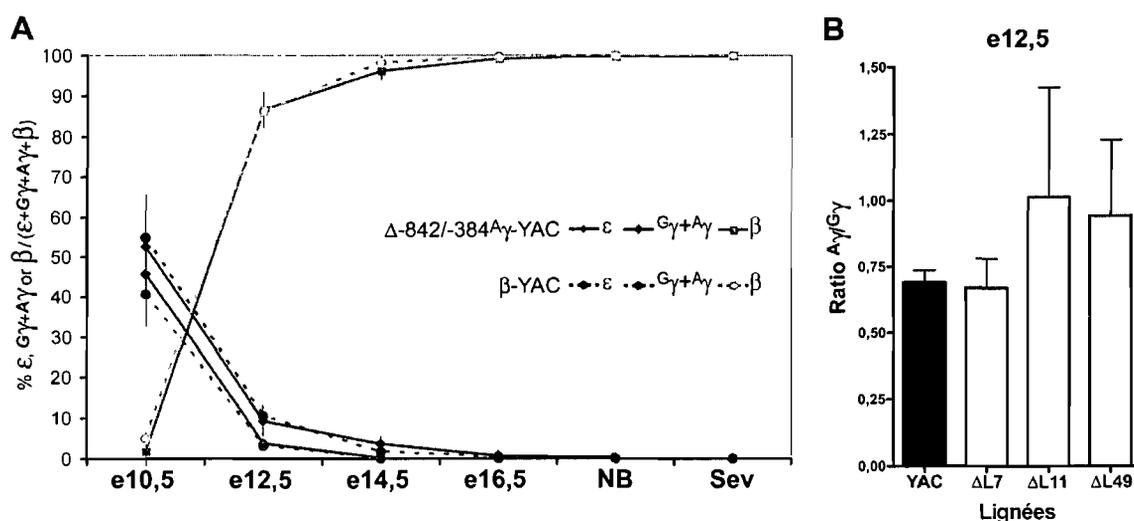


Figure 35: Expression des gènes du locus Δ -842/-384 $A\gamma$ -YAC. (A) Le patron d'expression des différents gènes du transgène a été mesuré par RT-PCR quantitatif en temps-réel à différents stades du développement chez les fœtus de e10,5 à e16,5, chez les nouveau-nés (NB) et chez les souriceaux en âge d'être sevrés (Sev). Les résultats représentent la moyenne des trois lignées $\Delta L7$, $\Delta L11$ et $\Delta L49 \pm$ l'écart type ($N = 9-10$ pour chaque stade et $N = 3-4$ pour chaque lignée) et est comparée à celle de la souris β -YAC ($N = 3$ pour chaque stade). (B) Moyenne \pm écart type des ratios de l'expression de la chaîne $A\gamma$ sur la chaîne $G\gamma$ mesurés par *UT-PAGE* du sang des fœtus Δ -842/-384 $A\gamma$ -YAC ($\Delta L7$, $N = 4$; $\Delta L11$, $N = 4$; $\Delta L49$, $N = 6$ et β -YAC, $N = 4$) prélevés au jour embryonnaire e12,5.

3.3.5 Les souris Δ -384/-202 ^{Δ} γ -YAC

Pour évaluer l'impact de la délétion de la région comprise entre -384 et -202 pb du site d'initiation de la transcription du gène Δ γ sur la commutation de l'hémoglobine fœtale à adulte, l'expression au niveau de l'ARN des différents gènes a été mesurée par PCR quantitatif en temps réel aux stades embryonnaire et fœtal chez les souris Δ -384/-202 ^{Δ} γ -YAC.

Comme il a été observé chez les souris Δ -842/-384 ^{Δ} γ -YAC, le patron d'expression des gènes ϵ , γ et β , de e9,5 à e14,5 jours chez la seule lignée Δ -384/-202 ^{Δ} γ -YAC ayant le locus intègre, ne diffère pas de façon significative du patron d'expression de ces gènes dans le locus sauvage β -YAC (Figure 36). Ce résultat indiquait ainsi que la délétion de cette région du promoteur de la Δ γ -globine n'influçait pas la régulation de la commutation de l'hémoglobine fœtale à adulte.

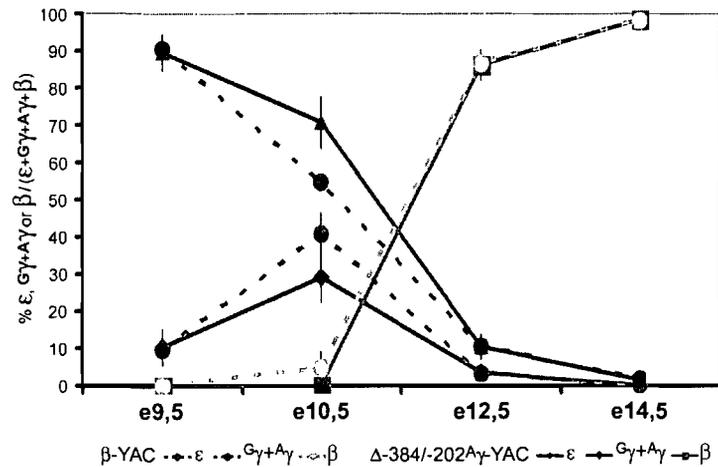


Figure 36: Expression des gènes du locus Δ -384/-202 ^{Δ} γ -YAC. Le patron d'expression des différents gènes du transgène a été mesuré par RT-PCR quantitatif en temps-réel à différents stades du développement chez les fœtus de e9,5 à e14,5 jours. Les résultats représentent la moyenne de la lignée Δ M56 \pm l'écart type ($N = 3-5$ pour chaque stade) comparée à celle de la souris β -YAC ($N = 3$ pour chaque stade).

3.4 Études de l'épigénétique chez la souris PBGD^Δγ-YAC

La différence importante entre les patrons d'expression des gènes de type β-globine chez les souris PBGD^Δγ-YAC comparés à la souris β-YAC m'a mené à analyser la signature épigénétique dans le locus mutant et son rôle potentiel dans la dérégulation de la commutation de l'hémoglobine fœtale à adulte.

3.4.1 Modification des histones dans le locus

Puisque les résultats d'expression à e12,5 indiquaient que les souris PBGD^Δγ-YAC exprimaient essentiellement les gènes fœtaux γ alors que la souris β-YAC exprimait surtout β, l'état de la chromatine dans le locus a été mesuré par immunoprécipitation de la chromatine (ChIP) de l'ADN du foie fœtal à e12,5. Ainsi, les niveaux d'acétylation des histones H3 et H4 ainsi que les niveaux de di-méthylation de la lysine K4 de l'histone H3 ont été mesurés au niveau des deux promoteurs γ et du promoteur β dans deux lignées PBGD^Δγ-YAC (PY367 et PY465) et dans la souris β-YAC (Figure 37). Aucune différence statistiquement significative n'a été mesurée au promoteur ^αγ et PBGD^Δγ des souris PBGD^Δγ-YAC en comparaison aux promoteurs γ de la souris β-YAC (Kruskal-Wallis: $p > 0,1$). Pour ce qui est du promoteur β, mis à part une légère augmentation significative de la di-méthylation en K4 de l'histone H3 chez les souris PBGD^Δγ-YAC en comparaison avec la souris β-YAC aucune autre différence significative n'a été observée (Mann-Whitney: $p > 0,1$). De fait, les promoteurs fœtaux (Friedman: $p = 0,19$) et adultes (Wilcoxon: $p = 0,25$) des souris PBGD^Δγ-YAC ne présentaient aucune tendance générale significative vers une signature d'ouverture ou de fermeture de la chromatine en comparaison à la souris β-YAC.

Une analyse similaire a été effectuée sur les précurseurs érythroïdes (CD71⁺Ter119⁺) triés par FACS des cellules de la moelle osseuse de souris adultes de deux lignées PBGD^Δγ-YAC (PY367 et PY465) et de la lignée β-YAC (Figure 37). Cette fois aucune différence n'a été mesurée pour les trois modifications au niveau du promoteur β

(Mann-Whitney: $p > 0,7$) qui ne présentait aucune tendance générale vers une signature d'ouverture ou de fermeture de la chromatine en comparaison à la souris β -YAC (Wilcoxon: $p = 0,25$). Au niveau des promoteurs γ toutefois, une diminution significative de l'acétylation de l'histone H4 ainsi que de la di-méthylation en K4 de l'histone H3 a été mesurée lorsque comparés aux promoteurs γ de la souris β -YAC qui étaient principalement dues à des diminutions au niveau du promoteur PBGD^A γ . De fait, les promoteurs fœtaux présentaient une signature épigénétique qui tendait significativement vers une fermeture de la chromatine en comparaison au locus sauvage (Friedman: $p < 0,03$) dont la plus grande contribution provenait du promoteur PBGD^A γ (Dunn: $p < 0,05$).

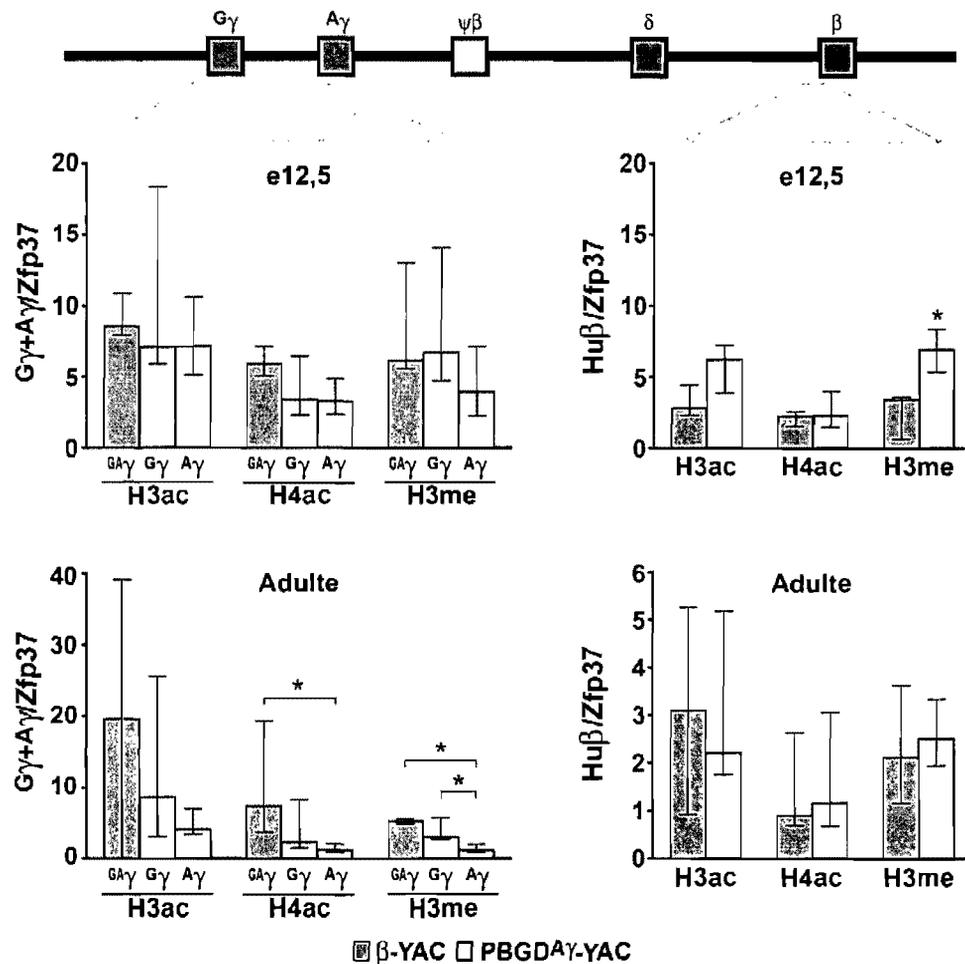


Figure 37: Modifications des histones dans les souris PBGD^{Aγ}-YAC. Une immunoprécipitation de la chromatine (ChIP) a été effectuée sur des fœtus à e12,5 et des adultes PBGD^{Aγ}-YAC (barres blanches) et β-YAC (barres grises) en utilisant des anticorps reconnaissant spécifiquement la forme acétylée des l'histones H3 (H3ac) et H4 (H4ac) ainsi que la forme di-méthylée en K4 de l'histone H3 (H3me). Les fragments des ChIP ont été analysés par qPCR en temps réel aux deux promoteurs fœtaux (^GAγ, -158 à +53) pour la lignée β-YAC ($N = 3$), aux promoteurs ^Gγ (^Gγ, -158 à +53) et PBGD^{Aγ} (^Aγ, -167 à +47) chez les PBGD^{Aγ}-YAC ($N = 6$; PY367, $N = 3$ et PY465, $N = 3$) ainsi qu'au promoteur β (-146 à +36). Les histogrammes représentent la médiane du ratio d'enrichissement relatif au gène contrôle *Zfp37* ± les écarts interquartiles. Le p a été déterminé par le test de Kruskal-Wallis suivi du post-test des comparaisons multiples de Dunn pour les promoteurs fœtaux alors que le test de Mann-Whitney a été utilisé pour le promoteur β: *, $p < 0,05$.

3.4.2 Méthylation 5mC de l'ADN dans le locus

Les niveaux de méthylation des sites de restriction des enzymes sensibles à la méthylation (HpaII, HhaI ou BsaAI) ont été mesurés en évaluant la résistance au clivage de l'ADN du foie d'embryons PBGD^{Δγ}-YAC et β-YAC prélevés au jour embryonnaire e12,5. Puisque le locus de la β-globine est pauvre en séquences CpG, seuls quelques sites situés aux promoteurs (proximaux et distaux) et/ou dans la séquence codante des gènes ^Gγ, ^Aγ, δ et β ont été analysés (Figure 38).

Le site HpaII proximal de ^Gγ (-54 du site d'initiation de la transcription) a démontré un niveau de 5mC environ trois fois inférieur dans les souris PBGD^{Δγ}-YAC (~25%) que dans la souris β-YAC (~80%) ce qui corrèle bien avec l'activité transcriptionnelle de ce gène dans ces deux lignées. À l'inverse, les deux sites distaux HpaII (-3419) et HhaI (-1410) n'ont démontré aucune différence significative bien que le site de méthylation à -3419 semblait présenter une diminution de ~50% dans les souris PBGD^{Δγ}-YAC en comparaison aux souris β-YAC. En fait, cette différence était due à un seul embryon chez qui ce site a été digéré à 100% par l'enzyme alors que les autres embryons présentaient une moyenne du pourcentage de méthylation semblable au pourcentage observé dans la lignée β-YAC. Le gène δ-globine qui n'est pas exprimé à ce stade chez les embryons PBGD^{Δγ}-YAC et β-YAC présentait un niveau de 5mC de plus de 90% au seul site BsaAI présent dans sa séquence codante (+469) et ce, chez tous les embryons analysés. De même, les sites BsaAI situés à -1791 et -268 du promoteur β ne présentaient aucune différence significative entre les locus mutant et sauvage. Toutefois, le site BsaAI situé dans la séquence codante de β démontrait une augmentation modeste mais significative de 5mC chez les embryons PBGD^{Δγ}-YAC (~78%) comparée au locus sauvage (~55%) et corrélant avec l'activité transcriptionnelle.

Il est à noter que les niveaux de 5mC mesurés au site HpaII à +21 dans le promoteur PBGD riche en CpG, étaient beaucoup plus bas que le site HpaII à -54 du gène ^Aγ du locus

sauvage β -YAC et étaient comparables en fait, aux niveaux de 5mC mesurés dans le promoteur G_{γ} du locus mutant (Figure 38).

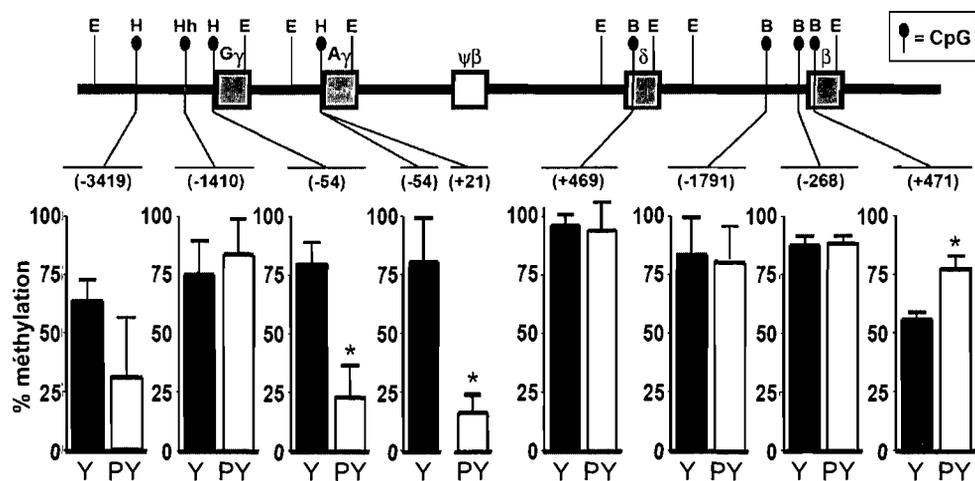


Figure 38: Patron de méthylation de l'ADN dans les souris PBGD^A γ -YAC. Le statut de 5mC des sites HpaII (H), HhaI (Hh) et BsaAI (B) situés entre les sites délimitants EcoRI (E), de l'ADN isolé du foie prélevé à e12,5 de fœtus PBGD^A γ -YAC (PY, $N = 4$; PY367, $N = 2$ et PY418, $N = 2$) et β -YAC (Y, $N = 3$). Les numéros au dessus des histogrammes donnent les coordonnées des sites de restriction relativement au site d'initiation de la transcription de leur gène correspondant. Les histogrammes représentent la moyenne du pourcentage de la fraction méthylée non-digérée \pm l'écart type. Le p a été déterminé à l'aide du t -test de Student. *, $p < 0,002$

3.5 Études mécanistiques chez la souris PBGD^A γ -YAC

Afin de déterminer les mécanismes généraux menant à l'expression soutenue des gènes fœtaux et au retard de l'activation des gènes adultes dans la souris PBGD^A γ -YAC, j'ai décidé d'étudier le comportement du locus mutant par une étude mécanistique permettant de mesurer la proximité relative entre le LCR et les différentes régions du locus.

3.5.1 Capture de conformation chromosomique

Une étude de capture de conformation chromosomique (3C) permettant de déterminer les interactions physiques du LCR avec les différentes régions du locus a été effectuée sur les cellules de foies fœtaux prélevés d'embryons à e12,5 ainsi que sur les précurseurs érythroïdes de la moelle osseuse de souris adultes de deux lignées PBGD^A γ -YAC (PY367 et PY465) et de la lignée contrôle β -YAC.

Chez l'embryon à e12,5

Les interactions génomiques ont été quantifiées en mesurant le taux de réticulation entre les fragments EcoRI, normalisé pour l'efficacité des différents PCR à l'aide du contrôle de ligation et pour la concentration et le nombre de copie sur le pseudogène $\psi\beta$ auquel a été assignée la valeur "1" (Figure 39). Comme contrôle de réticulation non spécifique, l'interaction entre les fragments EcoRI du LCR et de l'exon 15 du gène endogène Pkd1 a été mesurée. Puisque aucun produit de PCR n'a été détecté entre ces deux fragments, ceci confirmait la spécificité des interactions mesurées dans le locus de la β -globine.

Ainsi, à e12,5 dans les fœtus PBGD^A γ -YAC, l'analyse a démontré que le LCR s'associait spécifiquement avec tout le domaine fœtal, incluant les gènes $G\gamma$ et $A\gamma$ ainsi que la région intergénique entre ces deux gènes. Chez les deux lignées PBGD^A γ -YAC, cette association était significativement plus élevée que celle mesurée dans le locus sauvage β -

YAC (ANOVA: $F_{(2,6)} = 52,46$; $p < 0,002$) dans lequel à l'inverse, le LCR se retrouvait exclusivement en association avec tout le domaine adulte, ce qui incluait les gènes δ et β ainsi que leur région intergénique. Cette interaction ne se retrouvait pas chez les lignées PBGD^A γ -YAC (ANOVA: $F_{(2,6)} = 50,81$; $p < 0,002$). Il est à noter que les deux lignées PY367 et PY465 ne présentaient pas de différence significative entre elles, telle que démontré par le post-test des comparaisons multiples de Tukey ($p > 0,05$).

Chez l'adulte

Chez les souris adultes, la situation était partiellement corrigée puisque le LCR se retrouvait désormais à proximité du domaine adulte (δ à β) dans les souris β -YAC et dans les deux lignées PBGD^A γ -YAC (Figure 40). De fait, aucune différence significative n'a été observée entre les lignées PBGD^A γ -YAC et β -YAC au niveau de ce domaine (ANOVA: $F_{(2,6)} = 1,82$; $p = 0,27$). Toutefois, dans le locus mutant PBGD^A γ -YAC, le LCR démontrait toujours une fréquence relative de réticulation assez élevée au niveau du domaine fœtal (^G γ à ^A γ ainsi que leur région intergénique) que l'on ne mesurait pas dans le locus sauvage et ce pour les deux lignées PY367 et PY465 (ANOVA: $F_{(2,6)} = 9,91$; $p < 0,05$).

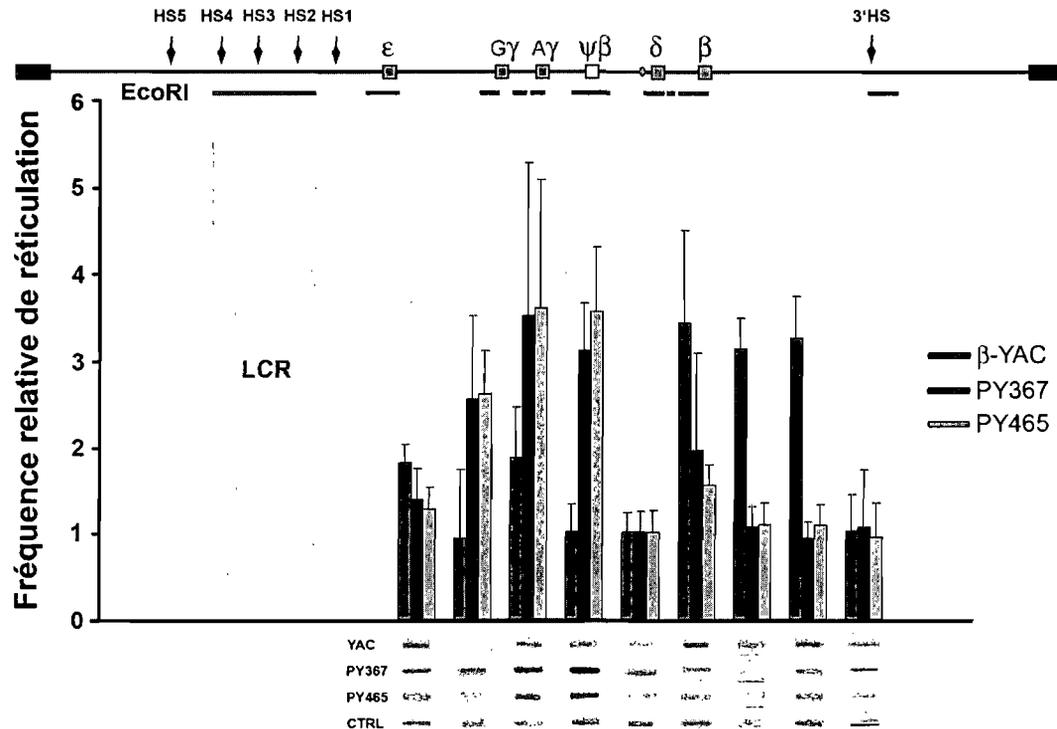


Figure 39: Proximité relative du LCR et des différentes portions du locus à e12,5, mesurée par 3C. La fréquence relative de réticulation entre le fragment EcoRI du LCR (HS2-4) et les différents fragments EcoRI du locus (indiqués par des barres grises sous le locus) a été évaluée en mesurant l'enrichissement des produits de PCR spécifiques. La fréquence relative de réticulation a été calculée en divisant le ratio d'intensité de l'amplicon pour une région donnée sur celle de l'amplicon du contrôle, par le ratio d'intensité de l'amplicon du pseudogène $\psi\beta$ sur celle de l'amplicon du contrôle (voir Méthodes, section 2.6.2, page 127). Sous les histogrammes sont présentés des exemples de produits de PCR représentatifs des différentes régions pour chacune des lignées. Les histogrammes représentent la moyenne de trois pools de 3-4 foies fœtaux \pm l'écart type.

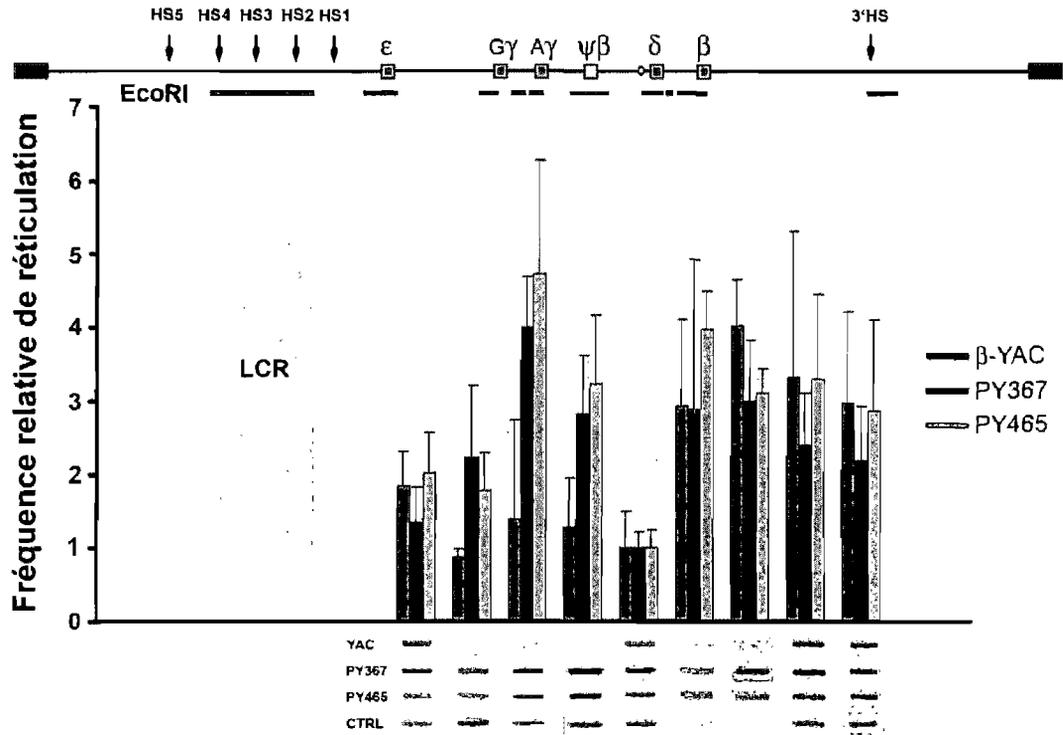


Figure 40: Proximité relative du LCR et des différentes portions du locus chez l'adulte. La fréquence relative de réticulation entre le fragment EcoRI du LCR (HS2-4) et les différents fragments EcoRI du locus a été mesurée dans des cellules Ter119⁺ isolées de la moelle osseuse de souris adultes PBGD^A γ -YAC et β -YAC. Les valeurs des histogrammes représentent la moyenne de trois souris \pm l'écart type.

Chapitre 4

Discussion

Le locus de la β -globine humaine représente un modèle intéressant pour l'étude de la régulation des grands loci multigéniques en général. Ce locus qui comprend cinq gènes fonctionnels exprimés de façon réciproque et hautement régulée au cours du développement a depuis longtemps représenté un grand défi pour les biologistes moléculaires. En effet, bien que chaque gène du locus possède des éléments régulateurs qui lui sont propres et lui assurent un patron d'expression adéquat lors du développement, ces éléments seuls ne parviennent pas à expliquer la régulation fine des différents gènes lors des deux commutations de l'hémoglobine que l'on retrouve chez l'humain et chez les grands primates. Pendant un certain temps, on a bien vu dans le LCR, l'élément du locus qui s'assurait de l'harmonie de la régulation des cinq gènes sous son contrôle,¹⁰⁴ mais aujourd'hui on sait que si le LCR est important pour le bon fonctionnement du locus, son rôle dans la régulation de l'hémoglobine fœtale à adulte est certainement beaucoup plus limité que ce qui avait tout d'abord été proposé.

La régulation de la commutation de l'hémoglobine fœtale à adulte ne peut probablement pas se résumer à un seul maître-régulateur, mais serait plutôt le fait de la collaboration de plusieurs éléments dont leurs fonctions propres contribueraient de façon plus modeste mais interdépendante à cette régulation. On pourrait certes être tenté de penser que chacun des gènes du locus est principalement régulé de façon autonome, tel qu'il a été suggéré par les premières études en transgéniques qui ont démontré que les promoteurs des gènes de globine possédaient la capacité d'être exprimés au stade approprié en l'absence du LCR à l'exception du gène β pour lequel un *enhancer* situé en 3' du gène était également requis.^{94;96;388} Toutefois, les études subséquentes ont démontré que l'expression coordonnée des différents gènes lors des deux commutations de l'hémoglobine ne pouvait s'expliquer simplement par l'inactivation de certains gènes de façon simultanée mais indépendante à l'activation des gènes réciproques et que certains mécanismes de "compétition" semblaient s'appliquer.^{101;103;516;528} De fait, il est devenu de plus en plus évident que la commutation de l'hémoglobine fœtale à adulte, qui se caractérise par

l'inactivation des gènes fœtaux $G\gamma$ et $A\gamma$ et par l'activation coïncidente des gènes adultes β et δ , serait le fait d'une inactivation relativement autonome des gènes fœtaux, dont dépendrait l'activation des gènes adultes.^{103;255;257;259} Ainsi, l'absence d'activité transcriptionnelle du gène β au stade fœtal serait due surtout à un mécanisme de compétition avec les gènes fœtaux pour l'activité du LCR. Ce modèle pointe ainsi vers les promoteurs fœtaux comme étant des éléments régulateurs majeurs de la commutation de l'hémoglobine fœtale à adulte, même si plusieurs questions demeurent sans réponse quant à leur rôle dans cette régulation. La première, soulevée par deux études ayant donné des résultats contradictoires, concerne l'inactivation des gènes adultes par compétition avec les gènes fœtaux et de l'importance de la structure ou de la nature des promoteurs des gènes γ dans cette compétition.^{512;516} En effet, une des questions restait à savoir si cette capacité à inactiver les gènes adultes au stade fœtal était intrinsèque à certains éléments spécialisés des promoteurs γ , à toutes autres parties du gène γ , ou si cette inactivation était simplement dépendante de toute activité transcriptionnelle LCR-dépendante en amont des gènes adultes. La deuxième question concernait la présence de deux gènes fœtaux aux promoteurs virtuellement identiques dans le locus. Si plusieurs études ont démontré que la présence d'un seul de ces gènes en amont des gènes adultes était suffisante pour obtenir une régulation adéquate de ces derniers lors du développement,^{260;262;293;529} l'importance de la duplication des gènes fœtaux dans le recrutement fœtal de leur expression chez les simiens n'était toujours pas comprise.⁵³⁰ De fait, les mécanismes par lesquels ces deux gènes sont co-régulés lors du développement demeuraient à ce jour, inconnus.

L'objectif de cette thèse était principalement de répondre à la première de ces deux questions, c'est-à-dire de déterminer le rôle du promoteur du gène fœtal $A\gamma$ dans la régulation de la commutation de l'hémoglobine fœtale à adulte. L'approche expérimentale développée pour cette étude devait en effet permettre de déterminer si le promoteur $A\gamma$ possédait des éléments précis dont la fonction spécifique serait de contrôler la commutation et d'identifier ces éléments. À l'inverse, elle aurait également permis de déterminer si la

fonction du promoteur fœtal dans la régulation de la commutation n'était due qu'à son activité transcriptionnelle fœtale-spécifique et LCR-dépendante. La modification d'un seul gène fœtal ($^A\gamma$) en gardant l'autre ($^G\gamma$) intacte, m'aura cependant également permis d'obtenir des éléments d'information sur l'implication de la duplication des gènes γ sur le recrutement fœtal de leur expression au cours de l'évolution et les mécanismes par lesquels ils sont co-régulés.

4.1 Impact du promoteur PBGD sur la commutation

La caractérisation fonctionnelle du promoteur fœtal-spécifique du gène de la $^A\gamma$ -globine a été entreprise en le remplaçant dans le contexte du locus entier par le promoteur érythroïde-spécifique mais stade-indépendant du gène de la porphobilinogène désaminase ou PBGD. Cette approche a fait en sorte de remplacer le promoteur de la $^A\gamma$ -globine, soupçonné d'être impliqué directement dans la régulation de la commutation de l'hémoglobine fœtale à adulte, par un promoteur capable de recruter sensiblement les mêmes facteurs de transcriptions et d'être activé par le LCR, mais dont l'expression devait être stable tout au long du développement.^{517;519;521}

Il est important de mentionner que cette modification a été effectuée sans altérer d'autres éléments du locus et que le contrôle utilisé lors de cette étude était une souris comportant le même locus, sans la substitution du promoteur. Ceci faisait en sorte d'assurer que toute divergence mesurée dans l'expression de tous les gènes du locus devait être due strictement à la substitution du promoteur $^A\gamma$ par le promoteur PBGD. De plus, le remplacement de la région du promoteur $^A\gamma$ par une région équivalente de taille identique (887 pb) garantissait l'intégrité de la chromatine en préservant la distance physique entre les différents éléments régulateurs du locus. Cette caractéristique était importante puisqu'il est connu qu'une perturbation de ces distances entre éléments génétiques, peut dérégler leur mode de fonctionnement et leur capacité à interagir, entre autre en modifiant la structure, l'espacement ou le positionnement des nucléosomes.^{531;532}

Il est à noter qu'à la base, la présence du gène fœtal $G\gamma$ non modifié devait servir de contrôle interne permettant d'évaluer l'impact du promoteur PBGD sur l'expression de $A\gamma$. Mais comme nous l'avons vu dans le chapitre précédent, l'activation de $A\gamma$ à tous les stades du développement a eu des répercussions importantes sur presque tous les gènes du locus, y compris son *alter ego* $G\gamma$.

4.1.1 Délai de la commutation de l'hémoglobine fœtale à adulte

La commutation retardée et incomplète de l'hémoglobine fœtale à adulte, provoquée par la présence d'un promoteur actif à tous les stades du développement à la place d'un promoteur actif seulement au stade fœtal, suggère plusieurs implications concernant les mécanismes sous-jacents à ce processus:

Le LCR est recruté par le promoteur PBGD

Le promoteur érythroïde de la porphobilinogène désaminase avait été choisi à la base pour sa capacité à recruter des facteurs similaires à ceux se liant aux promoteurs des gènes de globine, son expression indépendante des stades développementaux et sa capacité à être activé par le LCR du locus de la β -globine.^{517,519-521} Cette dernière caractéristique était importante puisqu'elle allait permettre de déterminer si la régulation de la commutation de l'hémoglobine fœtale à adulte et par conséquent, la compétition entre les promoteurs γ et β était dépendante d'éléments uniques aux gènes γ , ou plutôt secondaire à toute activité transcriptionnelle LCR-dépendante en amont des gènes adulte δ et β .

En accord avec ce qui avait été rapporté par Eleouet *et al.*,⁵¹⁹ le promoteur PBGD dans notre construction pouvait être activé par le LCR et ce, à tous les stades du développement, ce qui a été démontré par l'expression du gène hybride PBGD $A\gamma$ à des niveaux comparables aux gènes de globine du locus. De plus, l'analyse de 3C a clairement démontré que le LCR se retrouvait à proximité du promoteur PBGD indiquant de ce fait, que l'activation de ce promoteur par le LCR se faisait par interactions entre ces deux éléments via le *looping* de la chromatine.

Bien que les interactions LCR-PBGD ne représentent pas une fonction normale de la région de contrôle du locus de la β -globine, connaître quels étaient les éléments du promoteur de PBGD qui étaient impliqués dans le recrutement du LCR donnerait certainement des informations sur le fonctionnement de cet élément. Aussi, bien que cette question aille au-delà des objectifs de cette thèse, il est approprié de se questionner sur l'identité des principaux facteurs impliqués dans ces interactions. De fait, EKLF, GATA-1 et NF-E2 représentent trois candidats intéressants puisqu'ils sont tous capables de se lier au promoteur PBGD⁵¹⁷ et qu'ils ont tous été proposés comme étant impliqués dans les interactions LCR-promoteur dans le locus de la β -globine.^{205;398;406} Toutefois, sur la base des croisements des souris PBGD^{A γ} -YAC avec les souris EKLF-nulles qui n'ont démontré aucune variation de l'expression du gène ^{A γ} , je peux affirmer que ce facteur ne joue, au mieux, qu'un rôle très limité dans le recrutement du LCR au promoteur PBGD. Il est possible cependant, que dans le cas du promoteur PBGD, un autre membre de la famille des KLF puisse avoir une fonction redondante avec EKLF. De façon à pousser plus avant l'étude de ces interactions, il pourrait être intéressant de croiser les souris PBGD^{A γ} -YAC avec des souris nulles pour le gène NF-E2. En effet, un croisement avec ces souris qui sont viables et qui ne présentent qu'une variation négligeable de l'expression des gènes de globine,⁴¹³ pourrait donner une réponse rapide quant à l'implication de ce facteur dans les interactions LCR-PBGD.

Coopération bigénique

Coopération des gènes fœtaux

Contrairement à ce que j'attendais, cette étude a démontré qu'en dépit du fait que la région régulatrice du gène fœtal ^{G γ} , qui n'a pas été modifiée et qui donc, possède vraisemblablement tous les éléments nécessaires à sa répression au stade adulte,³⁰¹ cette répression peut être outrepassée par l'activation du gène ^{A γ} , localisé 4,9 kb en aval. Cette observation est importante puisqu'elle implique en corollaire que si pris individuellement

en association avec le LCR, les gènes fœtaux sont des entités indépendantes capables d'une inactivation autonome au stade adulte, lorsqu'ils sont associés dans un contexte qui leur est plus naturel, leur régulation devient plutôt interdépendante. Supportant l'idée que les deux gènes fœtaux sont capables de coopérer pour être régulés de façon coordonnée, des cas ont été rapportés où l'expression de $G\gamma$ était augmentée suite à la présence de mutations produisant une persistance héréditaire de l'hémoglobine fœtale spécifiques au gène $A\gamma$,^{533;534} ou à l'inverse, qu'une délétion de quatre nucléotides dans le promoteur de $A\gamma$ affectait négativement l'expression de $G\gamma$ au stade fœtal.⁵³⁵ D'ailleurs, une étude au cours de laquelle l'élément DRED du promoteur $A\gamma$ avait été muté dans des souris YAC transgéniques, avait résulté également en une coactivation modeste de $G\gamma$ bien que les auteurs n'aient pu l'expliquer.³⁰¹ Un mécanisme de co-régulation des gènes fœtaux par le LCR est également supporté par une étude qui a démontré que le LCR était capable de partager sa fonction activatrice entre les deux gènes γ d'un même chromosome.⁵⁵⁵ Cependant, puisque l'insertion d'un seul gène γ actif en amont du gène ϵ provoque l'inactivation des deux gènes γ situés à leur position naturelle,²⁵⁸ la position relative des deux gènes γ dans le locus serait critique pour la coopération des gènes fœtaux et leur régulation combinée.

Si la coopération bigénique a été suggérée sur la base de l'activation de $G\gamma$ conséquemment à l'activation de $A\gamma$ par le promoteur PBGD, c'est en observant les résultats de l'analyse de capture de conformation chromosomique que cette coopération prenait tout son sens. En effet, lorsqu'il est recruté au niveau du gène PBGD $A\gamma$, le LCR se retrouve également en association avec le gène $G\gamma$ ainsi qu'avec leur région intergénique. Il est intéressant de noter d'ailleurs que cette association avec la région intergénique est forte, aussi bien au jour embryonnaire e12,5 qu'à l'âge adulte, suggérant que cette association pourrait être impliquée dans l'interaction des deux gènes avec le LCR. Le fait que la proximité relative du LCR avec $G\gamma$, telle que mesurée à e12,5, soit de comparable amplitude à celle du LCR avec $A\gamma$, suggère en fait que le LCR serait capable d'interagir simultanément

avec les deux gènes fœtaux et ce, vraisemblablement via différentes régions ou HS. En effet, un mécanisme de *flip-flop* par lequel le LCR ferait la navette entre les deux gènes fœtaux ne saurait expliquer l'activation concomitante de $G\gamma$ et $A\gamma$ puisqu'il impliquerait que l'affinité de $G\gamma$ pour le LCR soit de force au moins égale à celle de $A\gamma$ malgré les signaux répresseurs présents. Les niveaux d'expression très élevés de $G\gamma$ observés au stade fœtal, supportent d'ailleurs l'hypothèse que le LCR puisse lier et activer les deux gènes fœtaux simultanément.

Dans le modèle que je propose, les deux gènes du domaine fœtal forment un sous-domaine génique capable de coopérer pour être activés de façon concomitante en interagissant simultanément avec l'holocomplexe du LCR via deux ou plusieurs points d'ancrages pouvant être situés dans leur promoteur respectif, dans des régions introniques ou non traduites, et/ou encore dans la régions intergénique. Ainsi certains HS pourraient avoir la fonction de s'attacher au sous-domaine alors que d'autres pourraient avoir la fonction d'activer les deux gènes fœtaux.

Coopération des gènes adultes

Bien qu'elle ne semble pas aussi pertinente que pour les gènes fœtaux puisque que chez l'humain, le gène de la β -globine représente essentiellement le seul gène de globine à être exprimé après la naissance, δ ne représentant jamais plus de 2-3% des chaînes de globine du sang périphérique,¹¹⁶ l'hypothèse d'une coopération bigénique pour les gènes adultes n'en demeure en fait, pas moins importante. D'ailleurs, la délétion de Corfu qui englobe une portion du gène δ ainsi que la région située en amont de ce dernier, résulte en une diminution de l'activité transcriptionnelle du gène β suggérant que les deux gènes adultes intacts sont nécessaires à la régulation de β .¹⁵⁷ De fait, les analyses de 3C présentés dans cette thèse ont démontré une proximité relative importante du gène δ au LCR à chaque fois que β y était associé et ce, indépendamment de l'activité transcriptionnelle de δ .

Ainsi, on peut observer des similitudes évidentes de comportement du LCR avec les gènes adultes et avec les gènes fœtaux.

La raison pour laquelle δ n'est exprimé qu'à des niveaux résiduels chez l'humain est assez bien documentée. La délétion de la boîte CACCC, due à une conversion génique du promoteur δ , fait en sorte que le facteur de transcription essentiel EKLF ne peut initier adéquatement la transcription^{120;536} et ce, même en présence du LCR. Ce qui a d'ailleurs été démontré par des études utilisant des constructions où le gène δ était lié à des éléments du LCR et qui ont échoué à produire des niveaux élevés de δ -globine.⁵³⁷ La question demeurerait toutefois à savoir si les faibles niveaux d'expression de δ étaient dus à son incapacité à recruter le LCR ou à une insensibilité à son activité. Les résultats de notre étude permettent de suggérer que la faible activité de δ ne serait pas due à l'incapacité du gène à interagir avec le LCR mais plutôt à une insensibilité à l'activation par ce dernier, puisque le LCR se retrouve à proximité de δ même lorsque ce dernier n'est pas exprimé. Toutefois, l'hypothèse que les interactions que j'ai observées entre le LCR et δ -globine ne soient pas attribuables directement au promoteur δ mais à d'autres séquences proximales telles que les introns ou encore des régions situées en amont ou en aval du gène ne peut être négligée. En effet, le fragment EcoRI du gène δ utilisé pour l'étude de 3C était de ~2,3 kb et couvrait tout le gène δ ainsi que près de 900 pb de la région promotrice et pourrait donc contenir des éléments encore non-identifiés et impliqués dans le recrutement du LCR. Un de ces éléments potentiels pourrait être l'élément PYR qui bien que n'étant pas inclus dans le fragment, se retrouve tout juste à la limite de ce dernier. D'ailleurs, le fait que l'expression des gènes de globine était diminuée sans préférence en absence d'Ikaros, suggère que ce facteur pourrait être impliqué dans la formation des interactions bigéniques avec le LCR, une fonction qui irait plus loin que sa seule présence dans le complexe PYR. En effet, la présence de sites de liaison d'Ikaros PYR-indépendants près du gène δ mais aussi aux environs des gènes γ ⁴¹⁶ où l'analyse de 3C a identifié de fortes associations avec les HS2-4 du LCR, suggère un rôle de ce facteur dans cette fonction.

Il est d'ailleurs intéressant de noter que la délétion du promoteur β dans le contexte du locus entier, n'empêche pas le recrutement du LCR au domaine adulte à moins qu'elle ne soit associée à une délétion du HS3.⁴⁶³ Ceci suggère en fait que ces deux éléments ne viennent pas en contact direct et donc, que le HS3 se lierait ailleurs dans le domaine adulte.

De fait, si le LCR est capable de s'associer simultanément aux deux gènes fœtaux via différentes régions ou HS, il serait logique de penser qu'il puisse le faire également avec les deux gènes adultes ce qu'ont effectivement démontré mes résultats de 3C. Ainsi, la liaison du LCR avec différents points d'ancrage dans le sous-domaine bigénique adulte pourrait permettre de stabiliser le grand holocomplexe LCR- δ/β .

La coopération bigénique n'est pas obligatoire

On sait depuis longtemps que les gènes de globine peuvent être associés individuellement avec différents éléments du LCR et être ainsi activés.^{346;347} Ceci indique que les différents HS du LCR peuvent individuellement activer un gène qui lui est associé avec une efficacité plus ou moins grande selon le HS. Si certains HS semblent présenter une préférence pour certains gènes plutôt que d'autres,^{347;349} le consensus général est de dire que la fonction des HS n'est pas spécifique à des gènes donnés ce qui soulève la question à savoir quel HS active quel gène dans le contexte d'une coopération bigénique. Des études qui ont déplacé ou ajouté des gènes dans le locus ont démontré également que le LCR complet pouvait activer un seul gène à la fois et donc, qu'il ne nécessitait pas absolument deux gènes pour fonctionner et ce, même si le patron d'expression des gènes déplacés s'en trouvait toujours altéré, étant généralement exprimés plus tôt que lorsque positionnés dans leur contexte naturel.^{258;488;507} Ceci indique que le LCR possède la capacité de s'associer à deux gènes simultanément, mais que cette association serait strictement dépendante de la distance entre ces gènes et de leur position respective dans le locus et que cette association simultanée est nécessaire à une régulation adéquate de ces gènes. Autrement, le LCR possède l'alternative d'aller vers le gène le plus près et de l'activer de façon indépendante

des autres gènes mais avec comme conséquence, une perte de la spécificité temporelle de l'expression. Ces dernières caractéristiques sont également supportées par le fait que la modification de l'orientation du locus ou du LCR affecte le patron d'expression de tous les gènes du locus.^{343;503}

ϵ -globine chez l'humain est un gène atypique

Chez l'humain et chez les autres grands singes, le gène ϵ est le seul gène à être exprimé au stade embryonnaire précoce.¹¹⁵ De ce fait, ce gène ne peut être impliqué dans une coopération bigénique et doit donc fonctionner de façon autonome.²⁵⁰ Ainsi, l'autonomie de ϵ irait bien plus loin que la simple autonomie d'inactivation lors de la commutation de l'hémoglobine embryonnaire à fœtale, mais se caractériserait également par une autonomie dans sa relation avec le LCR. D'ailleurs, il a été démontré que ϵ n'interagissait pas avec le LCR de la même façon que les autres gènes du locus, étant très sensible à la distance entre lui et le LCR.^{254;382} Indubitablement, ϵ représente un gène de globine atypique dont le mode de régulation diffère substantiellement de celui des autres gènes du locus.

La coopération bigénique... un phénomène unique à l'humain?

Si cette coactivation simultanée de deux gènes de globine par le LCR telle que présentée dans cette thèse s'avère exacte chez l'humain, il est peu probable qu'une fonction aussi importante soit exclusive à cette espèce. En effet, au cours de l'évolution le LCR a été remarquablement conservé entre les différentes espèces animales, le nombre et l'identité des HS étant presque toujours maintenus.¹²⁴ Ainsi, la coopération pour l'activité du LCR à un stade donné, de deux gènes ayant la même ontogénie, devrait être la norme plutôt que l'exception. De fait, il est notable que les gènes de type β -globine chez la plupart des espèces semblent fonctionner par paire. Par exemple, dans le locus murin de la β -globine, les deux gènes adultes (β^{maj} et β^{min}) sont exprimés dans les mêmes cellules aux stades fœtal et adulte alors que les gènes embryonnaires (ϵ^{Y} et β_{H1}) sont strictement exprimés dans les

cellules primitives.⁵³⁸ La distance entre ces gènes est plus grande que la distance observée entre les gènes humains ce qui pourrait poser des problèmes stériques inhérents à la capacité du LCR à lier deux gènes simultanément. Toutefois, il est remarquable que la taille globale du LCR chez la souris soit également plus grande que celle du LCR humain. Ce trait pourrait lui permettre en fait, de s'accommoder à la distance accrue entre les gènes des sous-domaines bigéniques (Figure 8, section 1.4.1, page 29). Comme pour le locus humain, peu d'études toutefois ont tenté de déterminer comment les deux gènes embryonnaires et les deux gènes fœtaux/adultes étaient coordonnés entre eux chez la souris. Une étude effectuée par Hu *et al* a démontré cependant que la délétion de l'un ou l'autre des deux gènes embryonnaires du locus, n'avait aucune répercussion marquante sur l'expression du gène restant mais conduisait à l'activation du pseudogène βh_0 , suggérant que la régulation des gènes embryonnaires ne se faisait pas par un mécanisme de compétition pour le LCR.^{539;540} Ceci supporte effectivement l'hypothèse que le LCR peut activer deux gènes simultanément via ses différentes régions ou HS. De plus, il a été démontré que la délétion du gène β^{Maj} de la souris produisait un phénotype thalassémique puisque le gène β^{Min} ne pouvait compenser pour la perte de l'autre gène,^{541;542} supportant également l'hypothèse que ces deux gènes ne sont pas en compétition directe pour le LCR.

Le modèle de la coopération bigénique permet également d'émettre des hypothèses concernant le rôle de la duplication du gène γ dans le recrutement fœtal de ce dernier chez les simiens. En effet, au cours de l'évolution le recrutement fœtal de l'expression du gène γ , et par conséquent, le report au stade adulte de l'expression des gènes adultes, ont coïncidé avec la duplication du gène γ qui a eu lieu avant l'émergence des simiens mais après celle des prosimiens qui ne possèdent qu'un seul gène γ .^{168;530} Chez ces derniers, tout comme chez la plupart des autres mammifères, l'expression des gènes ϵ et γ est strictement embryonnaire, alors que l'expression de β débute au stade fœtal et se poursuit chez l'adulte.¹⁶⁵ Même si la duplication de γ *per se* ne fut pas la cause directe du recrutement fœtal de ce gène,²¹⁵ elle peut avoir conduit à la formation d'un sous-domaine embryonnaire

trigénique incompatible avec la fonction du LCR et donc, au rejet de l'un de ces trois gènes produisant ainsi une redondance au niveau des gènes embryonnaires. Ce serait cette dernière qui aurait finalement permis à la spécialisation des gènes γ d'avoir lieu.^{215;530} De façon intéressante, les études effectuées sur les singes du Nouveau Monde et les singes de l'Ancien Monde suggèrent que les deux gènes d'un sous-domaine n'ont pas à être fonctionnels pour participer à la coopération bigénique. En effet, chez le catarhiniens, le gène δ n'est pas exprimé puisqu'il est réduit à l'état de pseudogène ce qui supporte l'idée que l'interaction δ -LCR présentée ici, n'était pas dépendante de l'activité transcriptionnelle de ce gène et que la fonction de δ pourrait en être une d'ancrage. De même, chez les Platyrrhiniens, le gène γ^1 qui équivaut au gène γ^G chez l'homme, n'est pas fonctionnel, suggérant que la coopération bigénique pourrait être dépendante principalement de la capacité d'ancrage des gènes d'un sous-domaine au LCR plutôt que leur activité transcriptionnelle. Ceci est d'ailleurs supporté par les observations faites chez le singe de nuit *A. azaræ*, chez qui les deux gènes γ ont fusionné pour ne former qu'un seul gène. Chez cette espèce, l'expression fœtale de γ est inférieure à celle observée chez les autres singes du Nouveau Monde et est associée à une activation précoce des gènes adultes. Cette observation peut suggérer soit une expression plus embryonnaire du gène unique γ qui pourrait coopérer avec ϵ , ou soit que la présence de ce gène unique ne peut compétitionner aussi efficacement avec les gènes adultes à cause de l'absence de coopération bigénique.¹⁶⁸ Ces observations supportent en fait l'idée que les gènes de globine fonctionnent mieux en paire et ce, même lorsque l'un des deux gènes n'est pas ou très peu fonctionnel et que leur capacité à agir comme ancrage pour le LCR serait déterminant dans cette fonction.

Une exception possible à ce modèle, pourrait être retrouvée chez les ruminants. En effet, ces animaux qui présentent entre deux et trois commutations de l'hémoglobine ne possèdent pas de gènes étant exprimés de façon concomitante à l'exception de trois gènes embryonnaires chez la chèvre (Section 1.4.5, page 36). Bien que chacun des gènes non-embryonnaires, qu'il soit fœtal, juvénile ou adulte, se retrouve dans un domaine comprenant

également un pseudogène δ (Figure 16, page 37) rien n'indique que ce dernier soit important.⁵⁴³ Toutefois, une autre caractéristique de ce locus atypique est la présence d'un LCR tronqué de deux HS qui possède ainsi seulement les HS1 HS2 et HS3.^{231;232} Il est possible que chez les ruminants, la coopération bigénique n'existe pas et que chaque gène soit régulé de façon monogénique par le LCR.

Les deux gènes en coopération ne sont pas régulés de façon équivalente

Chez plusieurs prosimiens, le locus de la β -globine présente une caractéristique assez unique, puisque le gène δ a subi une conversion par le gène β et possède donc les éléments régulateurs de ce dernier.¹⁶⁵ De fait, δ -globine est exprimé à des niveaux très élevés chez ces primates tout en demeurant néanmoins en deçà des niveaux du gène β (Figure 10, page 31). Si ces deux gènes étaient en compétition pour le LCR, une expression de δ supérieure à celle de β serait normalement attendue puisque le LCR favorise généralement l'expression des gènes les plus proximaux au dépend des gènes distaux. Le modèle de la coopération bigénique apporte toutefois une réponse à ce paradoxe apparent. En effet, si le LCR était capable de lier les deux gènes adultes simultanément pour les activer à l'aide de HS différents, une expression inégale de ces deux gènes, indépendante de leur position respective au LCR serait attendue puisqu'il a été démontré que les différents HS du LCR n'avaient pas tous la même capacité à activer la transcription. Ainsi le gène le plus fortement exprimé serait en fait, activé par les éléments les plus puissants du LCR ou pour lesquels il aurait la plus forte affinité d'interaction alors que l'autre gène serait activé par des éléments moins opérants, ou pour lesquels il aurait moins d'affinité. D'ailleurs, il est intéressant à noter que chez presque toutes les espèces où deux gènes sont exprimés simultanément, l'expression des deux gènes n'est pas parfaitement égale malgré des promoteurs équivalents (e.g. $A\gamma$ et $G\gamma$ chez l'humain ainsi que β^{Maj} et β^{Min} chez la souris).

Compétition transcriptionnelle des gènes fœtaux et adultes

J'ai jusqu'à présent, discuté essentiellement de l'aspect de la coopération bigénique des gènes du locus proposée sur la base des résultats présentés dans cette thèse, sans aborder la question fondamentale de cette dernière, c'est-à-dire à savoir quelle était la fonction du promoteur des gènes fœtaux dans la régulation de l'hémoglobine fœtale à adulte. La raison en est que le modèle de la coopération bigénique sera central à la discussion sur le rôle du promoteur fœtal dans la commutation de l'hémoglobine qui sera présenté ici.

A la base, cette thèse se proposait de répondre à l'ambiguïté concernant l'importance du promoteur de γ dans la compétition de ce gène avec β pour l'activité du LCR. Deux hypothèses étaient proposées: soit le promoteur du gène γ possédait des éléments et donc une fonction spécifique à cette compétition ou soit que toute activité transcriptionnelle LCR-dépendante en amont de β était suffisante à l'inactivation de ce dernier au stade fœtal. Cette étude aura démontré que la réalité n'est probablement pas aussi simple et qu'en fait, ces deux hypothèses pourraient ne former qu'une partie de la réalité.

Inactivation des gènes adultes partiellement spécifique au promoteur γ

Ainsi, tel qu'il était attendu d'un mécanisme de compétition non spécifique aux promoteurs fœtaux, l'activation du gène $^A\gamma$ à tous les stades du développement, obtenue du remplacement de son promoteur par celui de PBGD, a provoqué un important délai de l'activation du gène β en plus de la complète inactivation du gène δ . Ceci suggérait qu'une activité transcriptionnelle qui serait LCR-dépendante et localisée en amont de β , serait suffisante à la compétition avec ce dernier pour l'activité du LCR et donc, que les promoteurs fœtaux ne possèdent pas de fonction intrinsèque d'inactivation de β . Cette observation est supportée par des études qui ont démontré que l'insertion du gène adulte β entre le LCR et ε , en amont de $^G\gamma$ ou même en amont de δ , provoque l'inactivation des gènes situés en aval.^{258,488} Cette analyse plutôt simpliste fait toutefois abstraction de

certaines caractéristiques présentées dans cette thèse. La première de ces caractéristiques étant que, due à la coopération bigénique, $G\gamma$ était fortement activé, surtout au stade fœtal et que cette activation concordait précisément avec le retard d'activation de β . De fait, la cinétique d'activation de β semblait beaucoup plus sensible à la cinétique d'inactivation de $G\gamma$ qu'à la courbe d'expression de $PBGD^{\Delta\gamma}$ qui demeurait stable tout au long du développement. La deuxième caractéristique, corollaire à la première, était que malgré l'expression stable de $\Delta\gamma$ tout au long du développement, la commutation de l'hémoglobine n'était pas abolie mais simplement retardée. Ces deux caractéristiques suggéraient en fait que l'inactivation de β au stade embryonnaire serait dépendante principalement de l'activité transcriptionnelle spécifique aux gènes γ (ici le gène $G\gamma$). Toutefois comme nous le verrons dans la suite de cette discussion, certaines réserves doivent être présentées puisque les mécanismes impliqués pourraient être en fait un peu plus complexes.

Implication des gènes adultes dans la régulation de la commutation

En fait, sur la base de la deuxième caractéristique mentionnée à la section précédente concernant la commutation de l'hémoglobine fœtale à adulte retardée dans la souris $PBGD^{\Delta\gamma}$ -YAC, mais qui pour l'essentiel prend place malgré l'expression soutenue du gène $PBGD^{\Delta\gamma}$, deux hypothèses peuvent être émises pouvant expliquer cette commutation tardive. 1) La commutation qui survient tardivement serait le fait d'une répression autonome accrue du gène fœtal $G\gamma$, qui serait due soit à l'émergence de répresseurs spécifiques tardifs, soit à la disparition tardive d'activateurs fœtaux. 2) La commutation tardive serait le fait d'une augmentation de la compétitivité des gènes adultes provoquée par la présence d'activateurs tardifs ou encore par la disparition tardive de répresseurs spécifiques à β . Si la première hypothèse s'avérait être fondée, elle appuierait le modèle selon lequel la régulation de la commutation de l'hémoglobine fœtale à adulte serait essentiellement contrôlée par les gènes fœtaux ou par le sous-domaine bigénique fœtal. Pour sa part, si la deuxième hypothèse était la bonne, cela indiquerait que si les promoteurs fœtaux jouent un rôle primordial dans la régulation de la commutation, ils ne seraient pas les seuls impliqués

dans cette dernière, puisque les gènes adultes participeraient également à cette régulation, mais de façon plus tardive.

L'ablation du gène EKLF chez la souris PBGD^{Δγ}-YAC a pu fournir une partie de la réponse à cette question. En effet, j'ai démontré que lorsque le gène β était inactivé suite à l'absence de son activateur spécifique principal, l'expression de $\zeta\gamma$ était augmentée de façon inversement proportionnelle. Ceci suggérait fortement que la commutation tardive observée chez les souris PBGD^{Δγ}-YAC n'était pas due à une inactivation autonome accrue du gène fœtal à la naissance, mais plutôt à une potentialisation de l'activité de β à cette période. De fait, cette conclusion est supporté par des études qui ont démontré que l'inactivation des gènes fœtaux ne serait que semi-autonome et que β y aurait un rôle à jouer en étant en compétition avec γ pour le LCR.^{259;324;507}

La question à savoir si cette affinité accrue du promoteur β pour le LCR était due à un ou plusieurs activateurs tardifs ou encore à la disparition tardive de répresseurs demeure sans réponse. Peu de facteurs ont été identifiés pouvant réprimer l'expression de β au stade fœtal. De ceux-ci, le mieux caractérisé est le facteur fœtal-spécifique BP1, qui se lie à une région identifiée comme possédant une fonction de *Silencer*, située en amont du promoteur de β.^{321;335-337} Toutefois, le croisement des souris PBGD^{Δγ}-YAC avec des souris qui sur-exprimaient la protéine humaine BP1 à tous les stades du développement n'a démontré aucune variation d'expression des gènes fœtaux et adultes chez les nouveau-nés (voir résultats ajoutés en preuve à l'Annexe III, page XVI) indiquant de ce fait que dans le locus PBGD^{Δγ}-YAC, la commutation tardive ne semblait pas être provoquée par une dé-répression tardive du gène adulte due à la disparition de ce facteur. L'apparition tardive d'activateurs non-identifiés, dont la fonction demeure inconnue lorsque la commutation prend place normalement chez les souris transgéniques, pourrait donc être en cause.

Ces résultats indiquent que les gènes fœtaux ont un rôle obligatoire et majeur mais non exclusif dans la régulation de la commutation de l'hémoglobine fœtale à adulte puisque les gènes adultes y participent aussi de façon plus modeste.

La commutation du locus humain en souris transgéniques cache certains mécanismes propres à l'humain

Lorsque l'on compare le patron d'expression des gènes humains chez la souris PBGD^A γ -YAC (Figure 30, page 143) avec le patron d'expression de ces gènes chez l'humain (Figure 4, page 15), la ressemblance entre les courbes est frappante. Dès les premières études des gènes de globine en souris transgénique, les chercheurs avaient remarqué que la commutation des gènes humains de globine se produisait en même temps que la commutation de l'hémoglobine embryonnaire à fœtale chez la souris.^{96;487;544} Cette observation suggérait que les facteurs *trans*-régulateurs qui contrôlent cette commutation chez les rongeurs pourraient être les mêmes que ceux qui contrôlent la commutation de l'hémoglobine fœtale à adulte chez l'humain. Toutefois, les résultats que j'ai obtenus suggèrent que, certains signaux qui sont réellement impliqués dans le contrôle de la commutation fœtale à adulte qui a lieu après la naissance chez l'humain, pourraient être non détectés chez la souris, puisque pour une raison inconnue la commutation de l'hémoglobine précède l'établissement de ces signaux. Ainsi, dans la souris PBGD^A γ -YAC, la substitution du promoteur ^A γ pour celui de PBGD a permis d'empêcher la commutation précoce d'avoir lieu, révélant une commutation postnatale possiblement contrôlée par des facteurs qui sont impliqués directement dans cette commutation chez l'humain. Cette observation rend évidemment l'identification de ces facteurs, primordiale à la compréhension des mécanismes qui régulent la commutation de l'hémoglobine fœtale à adulte, et la souris PBGD^A γ -YAC représente certainement un modèle intéressant pour les définir.

Rôle épigénétique mineur dans le délai de la commutation chez la souris PBGD^A γ -YAC

Supportant l'hypothèse de domaines bigéniques fœtaux et adultes, plusieurs études de la chromatine dans le locus de la β -globine ont menés à la conclusion que le locus était divisé en sous-domaines épigénétiques corrélant précisément avec l'activité transcriptionnelle des gènes du locus et ce, tant chez la souris⁴²⁷ que chez l'humain,^{443;557}

voir même chez le poulet,⁴⁴¹ confirmant la conservation de cette caractéristique au cours de l'évolution .

Toutefois, une observation remarquable faite lors de cette étude est que malgré l'impact majeur de l'activation du gène fœtal γ à tous les stades du développement sur la transcription des différents gènes du locus, cette dérégulation n'a eu qu'un effet marginal sur la chromatine dans le locus. Normalement, la commutation de l'hémoglobine fœtale à adulte est associée à une commutation de l'ouverture de la chromatine du domaine fœtal au domaine adulte qui se caractérise par une hyperacétylation des Histones H3 et H4 ainsi qu'une di/tri-méthylation en K4 de l'histone H3.^{441;443} De fait, je m'attendais à retrouver une perturbation de cette signature au niveau des domaines épigénétiques fœtaux et adultes. Pourtant, malgré un patron d'expression opposé des gènes à e12,5, la signature épigénétique des histones ne présentait aucune différence significative entre le domaine fœtal du locus mutant et celui du locus sauvage. Plus encore, bien qu'aucune différence n'ait été mesurée dans le domaine adulte (promoteur β) au niveau de l'acétylation des histones H3 et H4, des niveaux plus élevés de di-méthylation en K4 de l'histone H3 ont été mesurés dans le promoteur β chez la souris PBGD γ -YAC que chez la souris β -YAC, ce qui à première vue semblait contraire à ce qui était attendu. Cette apparente contradiction entre la signature épigénétique et l'activité transcriptionnelle des gènes du locus était même intensifiée chez les adultes où les promoteurs fœtaux du locus sauvage présentaient une tendance significative vers une ouverture de la chromatine en comparaison au locus PBGD γ -YAC. C'est résultats contradictoires et en apparence irréconciliables avec les résultats d'expression trouvent leur explication dans une étude récente qui a démontré que les promoteurs des gènes de globine qui sont transcriptionnellement actifs peuvent présenter des niveaux de modifications épigénétiques très bas lorsque analysés par ChIP.⁵⁵⁶ Ceci serait dû à une déplétion en histone de la région promotrice proximale de ces gènes, ce qui rend leur détection difficile. Ainsi dans notre cas, si une signature d'ouverture de la chromatine a été observée chez les promoteurs moins actifs, c'était vraisemblablement parce que les promoteurs plus actifs possédaient tout simplement moins d'histones. Une

analyse de la signature des histones dans la région codante des gènes et/ou de différentes modifications de la chromatine comme la tri-méthylation en K4 de l'histone H3 auraient possiblement pu fournir des renseignements supplémentaires sur les variations épigénétiques qui prennent place dans le locus au cours du développement. Mes résultats suggèrent toutefois que la potentialisation épigénétique du locus, caractérisé par des marqueurs d'ouverture de la chromatine, ne serait que minimalement dépendante de l'activité transcriptionnelle des gènes ou des interactions avec le LCR. Cette observation est d'ailleurs supportée par des études qui ont démontré que le LCR n'était pas essentiel à l'ouverture de la chromatine dans le locus.^{341;424;444}

À l'inverse de ce qui a été observé au niveau des histones, des niveaux différentiels de 5mC à certains dinucléotides CpG présents dans le promoteur proximal et/ou dans la séquence codante des gènes G_γ et β ont été mesurés entre les souris PBGD^A γ -YAC et les souris β -YAC. Bien que le locus de la β -globine soit pauvre en îlots CpG, la méthylation de CpG isolés peut être associée à une inactivité transcriptionnelle, tel qu'il a été démontré avec certains gènes pauvres en CpG comme les gènes de globine et ce, indépendamment des modifications au niveau des histones.^{545;546} Si cette méthylation/déméthylation des cytosines était responsable de, ou subséquente à l'activité transcriptionnelle des gènes, demeure toutefois à être élucidé. En effet, si le locus de la β -globine ne possède pas de sites canonique pour la liaison des répresseurs MBD pouvant se lier aux CpG-méthylés,⁴⁴⁷ certaines évidences demeurent que la méthylation de l'ADN au niveau des promoteurs γ pourrait être un régulateur actif de la répression de ces gènes,⁴⁵² possiblement via le facteur Sp1 qui est capable de lier préférentiellement le SSE des gènes γ lorsque ce dernier est méthylé.^{298;299} Il se pourrait ainsi que la capacité du LCR à outrepasser les signaux répresseurs de G_γ lorsqu'il est positionné à proximité, se fasse par la prévention de la méthylation de l'ADN dans ce promoteur, suggérant un rôle actif du LCR dans cette fonction.

Coopération bigénique et bicompetition?

Une conséquence importante de la coopération bigénique dans la compétition des promoteurs fœtaux et adultes pour le LCR serait la présence de ce que j'appellerai une bicompetition. En effet, la compétition entre les gènes fœtaux et adultes dans le cadre d'une coopération des gènes d'un même domaine génique implique également une compétition bigénique, c'est-à-dire que les deux gènes fœtaux seraient en compétition avec les deux gènes adultes. Ceci est d'ailleurs supporté par l'observation que de hauts niveaux d'expression des gènes fœtaux dans le cas de HPFH, thalassémies et stress hématopoïétiques, corrôlaient avec une diminution de l'expression des deux gènes adultes β et δ .⁵⁵⁷ Le fait que la délétion de Corfu qui élimine une partie du gène δ résulte en une activation des gènes fœtaux,¹⁵⁷ alors que la délétion de la région intergénique entre $\Lambda\gamma$ et δ ne produit pas cet effet,³⁸⁷ vient supporter encore plus l'idée d'une bicompetition spécifique aux gènes ou à leur région proximale. De plus, puisqu'il est probable que les deux gènes d'un domaine donné soient activés par des régions ou HS différents du LCR, la compétition d'un gène doit se faire, dans une certaine mesure, préférentiellement avec le gène de l'autre domaine avec lequel il partage la même niche dans le LCR. De fait, tel qu'il peut être observé à la Figure 41A (page 186), il est notable que l'expression de $G\gamma$ et celle de β semblent être affectées de façon réciproque, puisque l'effet négatif sur l'expression de β ne diminue que lorsque l'effet positif sur le gène $G\gamma$ décroît. À l'inverse, l'effet négatif mais stable observé sur le gène δ semblerait être la résultante de l'expression stable du gène $\Lambda\gamma$. Ainsi, l'hypothèse la plus simple serait que cette bicompetition ferait intervenir une compétition spécifique entre les gènes $G\gamma$ et β et entre les gènes $\Lambda\gamma$ et δ .

Il est à remarquer que chez les fœtus humains, $G\gamma$ représente la chaîne de globine fœtale prédominante alors que la chaîne $\Lambda\gamma$ est exprimée beaucoup moins fortement. Après la naissance, une commutation des chaînes fœtales a lieu puisque $\Lambda\gamma$ devient la chaîne fœtale prédominante au moment où leur expression globale diminue fortement et qu'elles sont remplacées par les chaînes adultes.¹¹⁷ Supportant l'hypothèse de la bicompetition, une telle commutation des chaînes fœtales serait effectivement attendue si le gène $G\gamma$ devait être

en compétition directe avec le gène β qui possède le potentiel transcriptionnel le plus élevé puisque $\text{A}\gamma$ se retrouverait ainsi en compétition avec le gène plus faible δ .

Le promoteur de PBGD se distingue des promoteurs de la β -spectrine et de la thymidine kinase.

Lorsque Sabatino *et al* ont remplacé le promoteur de la $\text{A}\gamma$ -globine par celui de la β -spectrine dans leur mini-locus, ils ont obtenu une absence complète de compétition entre ce promoteur et celui de β -globine, qui était alors exprimé à tous les stades du développement.⁵¹² Ces résultats, qui sont en opposition avec ceux que j'ai obtenus avec le promoteur PBGD, indiquent que ces deux promoteurs érythroïde-spécifiques ne sont pas équivalents et qu'ils ne peuvent interagir de la même façon avec le LCR du locus de la β -globine. D'ailleurs, l'absence de site de liaison au facteur NF-E2 ou encore la faible affinité de EKLF pour la boîte CACCC du promoteur de la β -spectrine pourrait expliquer l'incapacité de ce promoteur à recruter le LCR, laissant la voie libre à ce dernier pour aller activer β à tous les stades.^{512;553} Il est également possible que l'absence de $\text{G}\gamma$ dans cette construction ait permis une coopération bigénique β -spectrine $\text{A}\gamma$ - β permettant l'expression des deux gènes, même si cette dernière hypothèse est plutôt improbable à cause de l'utilisation d'un μ LCR, vraisemblablement trop petit pour remplir une telle fonction.

À l'inverse, il est plus difficile d'expliquer comment le promoteur de la thymidine kinase utilisé par Anderson *et al* puisse inactiver β au stade fœtal puisque d'une part, aucune étude n'a démontré que ce promoteur pouvait interagir avec des éléments du LCR et que d'autre part, les auteurs n'avaient pas réussi à comparer directement l'expression relative des gènes CAT et β .⁵¹⁶ Toutefois, il est possible que le promoteur de la thymidine kinase soient capable d'interagir faiblement avec le HS2 qui était utilisé dans cette construction et que lors de l'érythropoïèse primitive, où le potentiel transcriptionnel du gène β est le plus faible, cette affinité soit suffisante pour l'inactivation de β par compétition. Par la suite, lors de l'érythropoïèse définitive, alors que le gène β voit son

potentiel transcriptionnel augmenter, cette faible affinité du promoteur de la thymidine kinase pour le HS2 deviendrait insuffisante pour cette compétition. Ceci supporterait encore une fois un modèle où les gènes adultes seraient impliqués dans la régulation de la commutation tel que je le propose dans cette thèse.

4.1.2 Un modèle

Ainsi, cette thèse propose que la régulation de l'hémoglobine fœtale à adulte dans le locus de la β -globine soit contrôlée principalement par les promoteurs fœtaux via une coopération bigénique des deux gènes γ pour l'activité du LCR. Ce dernier peut ainsi se lier à un domaine génique et y activer les deux gènes simultanément, possiblement aidé en cela par le facteur Ikaros et autres facteurs impliqués dans les interactions avec le LCR comme GATA-1 et NF-E2 (Figure 41B). De cette coopération, découle une bi-compétition pour le LCR entre les gènes du domaine fœtal et ceux du domaine adulte. Si l'expression de ces derniers est tributaire à l'inactivation relativement autonome des gènes du domaine fœtal, cette inactivation est en fait potentialisée par l'activation des promoteurs adultes, sans doute par le recrutement d'activateurs tardifs spécifiques ou par la disparition tardive de répresseurs faibles spécifiques aux gènes adultes. Étant donné la capacité du LCR à interagir de façon concomitante avec deux gènes d'un même domaine, ces gènes sont activés par différentes régions du LCR et donc, de façon non équivalente. Cette non-équivalence se reflète d'ailleurs dans l'expression inégale de deux gènes ayant des promoteurs équivalents, ainsi que dans la façon dont ils vont être en compétition avec les gènes de l'autre domaine. En effet, ils seront plus sensibles à la capacité à attirer le LCR du gène qui partage la même niche qu'eux, qu'à celle du gène avec qui ils ne sont pas en compétition directe (Figure 41A-B). La nature de cette bicompetition est évidente dans le cadre du locus mutant PBGD^A γ -YAC puisque ^G γ et β semblent être en compétition directe et que PBGD^A γ et δ le semblent tout autant. Toutefois, la question à savoir si cette spécificité ^G γ - β et ^A γ - δ s'applique également au locus sauvage, n'est soutenue que par le fait

que le ratio des gènes fœtaux G_{γ}^A subit une commutation après la naissance chez l'humain suggérant effectivement que G_{γ} et β sont en compétition pour le même site du LCR.

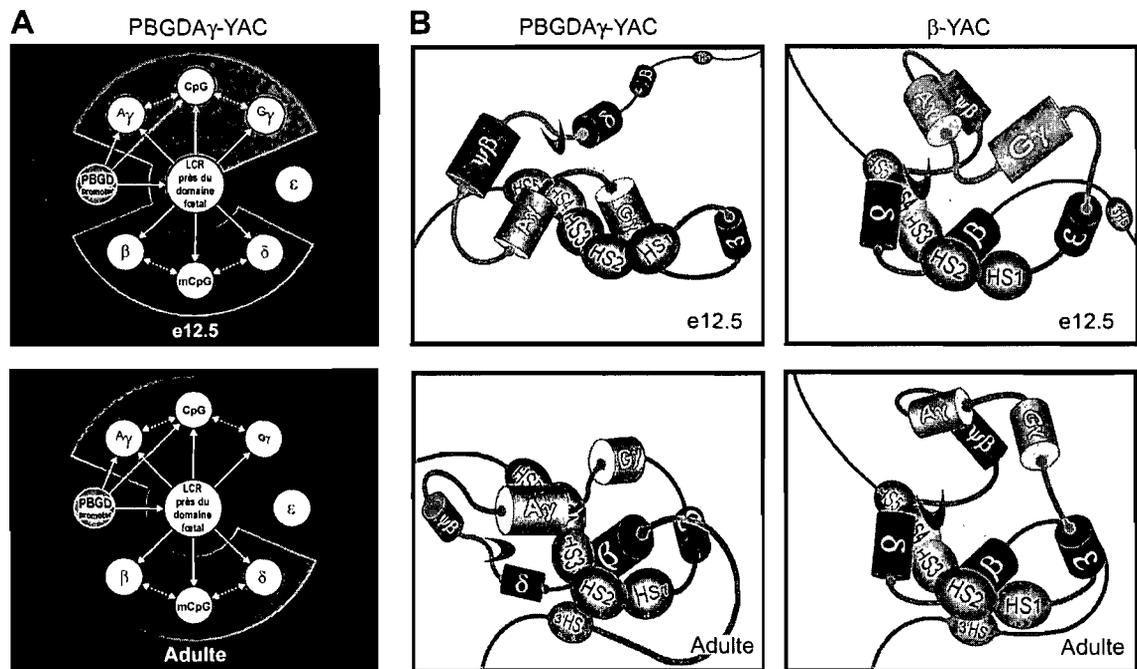


Figure 41: Modèle de la régulation des gènes de globine dans le locus $PBGD^A$ -YAC. (A) Impact du promoteur PBGD sur la proximité du LCR aux gènes fœtaux, sur l'expression des gènes du locus et sur l'état de la méthylation de l'ADN dans les domaines fœtaux et adultes. Vert: impact positif, Rouge: impact négatif. Flèche blanche: impact direct démontré, Flèche grise: impact direct non démontré (pouvant être un effet indirect), Flèche pointillée: Sens de l'impact non déterminé (lequel des événements est la cause et lequel est la conséquence demeure inconnu). (B) Modèle tridimensionnel de la coopération bigénique proposée pour les locus $PBGD^A$ -YAC et β -YAC chez les embryons à e12,5 et chez les adultes. \blacktriangleleft : Ikaros (PYR).

4.2 Impact des délétions sur la commutation

Ayant démontré que les promoteurs foetaux possédaient une fonction importante dans la commutation de l'hémoglobine foetale à adulte, la caractérisation des éléments spécifiques impliqués dans cette fonction devenait des plus pertinentes.

4.2.1 La délétion Δ -842/-384

La région comprise entre -842 et -384 du site d'initiation de la transcription du gène $\Lambda\gamma$ a déjà été démontrée comme possédant une activité de *Silencer* et représentait donc une région candidate importante.²⁵⁷ En effet, comme je l'ai démontré, c'est principalement l'inactivation autonome des gènes foetaux qui est responsable de l'activation temporelle adéquate des gènes adultes. À notre surprise, la délétion de cette région n'a eu aucun impact sur la commutation de l'hémoglobine foetale à adulte et ce, pour tous les gènes du locus, y compris $\Lambda\gamma$ lui-même. Ce résultat était d'autant plus surprenant qu'une étude récente effectuée par le groupe du Dr Kenneth Peterson a démontré que le promoteur commençant à la position -730 d'un gène $\Lambda\gamma^m$ inséré en 5' du gène ϵ était réprimé correctement chez l'adulte alors que le promoteur commençant à la position -378 inséré au même endroit menait à une expression soutenue du gène $\Lambda\gamma^m$ à ce stade.³¹² Ces résultats suggèrent que la région -730 à -378 comporte une fonction répressive de l'expression du gène $\Lambda\gamma$ lors de l'érythropoïèse définitive, ce qui est appuyé par le fait que la même étude a démontré également que la mutation directe du site GATA-1 situé à -570 du site d'initiation de la transcription de $\Lambda\gamma$ à sa position naturelle, abrogeaient aussi la répression de ce gène au stade adulte.³¹²

Toutefois, il est à noter que chez le mutant utilisé par l'équipe du Dr. Peterson, le gène $\Lambda\gamma^m$ positionné devant ϵ ne possède aucun des éléments régulateurs situés en amont de la région -730 du site d'initiation de la transcription et qu'il se retrouve donc dans un contexte génomique différent, ayant donc potentiellement des éléments régulateurs

différents.³¹² De même, il est possible que la délétion relativement importante de cette région dans mon modèle (458 pb) puisse avoir repositionné des éléments *cis*-répresseurs situés dans la région 5' du promoteur de γ^A , qui normalement sont trop éloignés du site d'initiation de la transcription pour interférer avec elle mais qui, après la délétion, se retrouvent suffisamment rapprochés pour réprimer le gène. Ceci expliquerait que la mutation ponctuelle du site GATA-1 à la position -570 résulte en une activation du gène γ^A lors de l'érythropoïèse définitive³¹² alors que la délétion complète de la région ne produit aucun effet. De fait, comme on peut le voir à la Figure 42, une région située en amont du promoteur de la γ^A -globine, reconnue comme ayant une certaine activité répressive et comprenant des sites de liaison à GATA-1 et YY1,³¹³ se retrouve repositionnée à proximité de la région normalement attribuée au *Silencer* de γ^A . Le plus troublant est la ressemblance de ce site GATA-1 avec celui normalement situé à -570, également à proximité d'un site YY1,³¹¹ suggérant que la coopération de GATA-1 et YY1 à cette position du promoteur pourrait être critique pour la régulation temporelle de γ^A . À l'inverse, lorsque le site GATA-1 à -570 est muté de façon à empêcher sa liaison spécifique, cette région se trouve dépourvue de l'activité de régulation négative de la transcription alors que les sites distaux GATA-1 et YY1, toujours intacts, se retrouvent à une position possiblement trop éloignée du site d'initiation de la transcription pour contribuer significativement à la répression de γ^A au stade adulte et pour reconstituer le *Silencer* principal de ce gène.

Ainsi, si cette étude semblait indiquer au départ que cette région n'était pas impliquée dans la régulation de l'hémoglobine fœtale à adulte, Une analyse plus stricte du promoteur mutant a indiqué que cette conclusion ne peut en être tirée. En effet, et à la lumière des travaux de Costa *et al*,³¹² cette région représente toujours un candidat majeur pour cette régulation. Ce que nos résultats suggèrent toutefois, c'est que la fonction de cette région serait dépendante de tout site GATA-1 adjacent à un site YY1 et situé dans un intervalle restreint du promoteur. Par quels mécanismes cette région permet l'inactivation du gène γ au stade adulte, demeurent incertains. Il a été démontré que GATA-1 était capable d'agir comme répresseur en séquestrant les séquences activatrices distales loin de

leurs gènes respectifs ou encore comme activateur lorsque son site est plus proximal au site d'initiation de la transcription.⁵⁴⁷ Cette double activité positive et négative est également une caractéristique du facteur YY1 qui est proposé comme étant un commutateur de l'expression génique et dont la fonction dépendrait du contexte moléculaire dans lequel il se retrouve et des partenaires avec lesquels il interagit.^{548;549} De quelle façon YY1 produit sa fonction répressive n'est pas très bien comprise⁵⁵⁰ mais une coopération avec le facteur GATA-1 a déjà été démontrée comme ayant une fonction inhibitrice de la transcription,^{283;551} supportant l'idée que cette coopération pourrait être également impliquée dans la régulation des gènes γ .

Modèle

En bref, le modèle que je propose est que des sites GATA-1 et YY1 situés dans une région définie du promoteur de γ -globine, sont importants pour la régulation de la commutation de l'hémoglobine fœtale à adulte. Lorsque ces sites sont situés à une distance trop importante, l'action répressive ne peut être suffisamment efficace pour inactiver le gène situé en aval. Ainsi, à cause du repositionnement des sites distaux de liaison à GATA-1 et YY1 à une distance suffisamment rapprochée suite à la délétion de la région -842 à -384, la fonction *Silencer* de cette région a été reconstituée dans les souris Δ -842/-384 γ -YAC ce qui expliquerait la divergence de nos résultats et ceux de Costa *et al* (Figure 42).

dans la régulation de la commutation de l'hémoglobine fœtale à adulte dans les souris transgéniques possédant le locus humain de la β -globine. Ces résultats suggéraient ainsi que la fonction du complexe PE γ dans cette régulation serait plutôt limitée.

4.3 Perspectives

Si cette thèse a permis de définir le rôle du promoteur fœtal $^A\gamma$ dans la régulation de la commutation de l'hémoglobine fœtale à adulte ainsi que de proposer un nouveau mécanisme pour la régulation concertée des gènes d'un même domaine dans le locus, elle a également soulevé certaines questions qui demeurent à être investiguées.

4.3.1 Définition des mécanismes de la coopération bigénique

Le modèle d'une coopération bigénique telle qu'elle a été mise en évidence dans cette thèse par les études d'expression et de capture de conformation chromosomique, demandera certainement à être investigué en profondeur. Également, l'implication du facteur Ikaros dans la formation des interactions simultanées entre le LCR et les deux gènes d'un même domaine devra être confirmée ou infirmée par des études de 3C chez les souris PBGD $^A\gamma$ -YAC/Ikaros $^{-/-}$, parallèles à des études d'expression effectuées à tous les stades du développement. En analyse supplémentaire, des croisements avec des souris NF-E2-nulles pourraient permettre de vérifier l'importance de ce facteur tant dans la coopération bigénique que dans la capacité du promoteur PBGD $^A\gamma$ à recruter le LCR.

4.3.2 Caractérisation de la bicompetition

L'élément proposé dans cette thèse ayant le plus besoin d'une caractérisation expérimentale est la possible bicompetition définie par la compétition spécifique de chaque gène d'un domaine avec un gène précis du domaine opposé dans le cadre d'un mécanisme de coopération bigénique. Puisque cette bicompetition implique la liaison des deux gènes d'un domaine donné avec différentes régions ou HS du LCR, l'étape inévitable de cette étude impliquera d'effectuer une analyse de 3C inversée par rapport à celle présentée dans

cette thèse et qui permettrait de déterminer la proximité relative des gènes avec différentes régions du locus, dont chacun des HS pris individuellement. En démontrant ainsi par exemple que les deux gènes foetaux se retrouvent à proximité de régions distinctes du LCR, comparables aux régions liées par les gènes adultes, la nature de la bicompetition pourrait être déterminée.

4.3.3 Évaluation du rôle de la région -204/-135

La production des souris transgéniques avec le locus mutant Δ -204/-135 ^{γ} -YAC demeure également un objectif principal dans cette étude. Cette région du promoteur du gène ^{γ} possède plusieurs sites de liaison à des facteurs de transcription impliqués dans la régulation de ce gène, dont la boîte CACCC située à ~145 pb en amont du site d'initiation et qui est importante pour l'activation du gène au stade foetal. Cette boîte CACCC a d'ailleurs été proposée comme étant un des éléments clés pour la répression du gène β par le promoteur γ au stade foetal.⁵¹⁴ Il sera intéressant de mesurer l'impact de l'inactivation de ^{γ} , qui est attendue de cette délétion, sur l'expression des autres gènes du locus lors de la commutation de l'hémoglobine foetale à adulte. En fait, si l'inactivation de ^{γ} au stade foetal était obtenue, cette étude donnerait de précieuses informations concernant la coopération bigénique et la bicompetition proposées dans cette thèse.

Conclusion

Cette étude visait à déterminer la fonction du promoteur foetal γ^A dans la régulation de l'hémoglobine foetale à adulte. Dans un premier temps, j'ai démontré que les promoteurs foetaux sont effectivement impliqués de façon importante mais non-exclusive dans cette régulation, implication illustrée par le délai de la commutation occasionné par l'activation des gènes foetaux suite au remplacement du promoteur foetal de γ^A par le promoteur stade-indépendant de la porphobilinogène désaminase. Dans un second temps, j'ai démontré que les mécanismes par lesquels les gènes de globine sont régulés par le LCR passeraient par une coopération bigénique des gènes d'un même sous-domaine épigénétique. En effet, l'activation du gène non modifié γ^G suite à l'activation forcée de γ^A , ne peut s'expliquer que par le potentiel du LCR à s'associer avec et à activer simultanément les deux gènes. Cette observation a d'ailleurs été supportée par une analyse de capture de conformation chromosomique qui a mis en évidence la capacité du LCR à interagir avec les deux gènes du domaine foetal ou du domaine adulte et ce même si l'un des deux gènes était peu exprimé. Cette capacité du LCR à interagir avec un gène de façon indépendante de la transcription de ce dernier, soulève d'ailleurs des questions concernant la nature des liens qui unissent le LCR aux gènes. Que pour chacun des promoteurs, certains facteurs spécifiques à ces derniers soient impliqués dans cette interaction est fortement vraisemblable. Toutefois, les résultats obtenus avec l'ablation génique du facteur Ikaros semblent identifier ce dernier comme un facteur potentiel impliqué de façon générale dans la formation de ces interactions.

En plus de la coopération bigénique des gènes de globine et conséquemment à celle-ci, j'ai proposé également au fil de cette thèse l'existence possible d'un mécanisme de bicompetition, qui ferait en sorte que chacun des deux gènes d'un domaine soient en compétition spécifique avec l'un des deux gènes du domaine opposé pour une région particulière du LCR, et proposé que dans le cas du locus mutant PBGD γ^A -YAC, cette bicompetition se ferait entre γ^G et β et entre γ^A et δ .

La caractérisation des régions du promoteur de γ^A par génération de délétions dans ce dernier n'aura toutefois pas permis d'identifier la région critique pour la fonction de

régulation de ce promoteur dans la commutation de l'hémoglobine fœtale à adulte. La délétion de la région -384 à -202, comprenant un site de liaison au complexe γ PE dont la fonction demeurerait controversée n'a eu aucun impact sur l'expression des gènes du locus lors du développement, incluant le gène γ lui-même. Ceci indique que cette région ne possède pas de fonction significative, à tout le moins dans la régulation de la première phase de la commutation de l'hémoglobine qui est observée pour le locus humain dans un contexte en souris transgénique.

De façon plus surprenante, la délétion de la région -842 à -384 de γ , qui était reconnue pour avoir une fonction de répression au stade adulte grâce à un site de liaison au facteur GATA-1, n'a également démontré aucun impact significatif sur la régulation des gènes du locus. Toutefois, une analyse attentive des séquences situées en amont de la délétion a permis de démontrer qu'un site GATA-1 très semblable au site impliqué dans la répression, se retrouvait repositionné à proximité du gène, reconstituant possiblement la fonction *Silencer* de cette région, supportant l'idée que cette région serait importante pour la régulation de la commutation de l'hémoglobine fœtale à adulte. Des études subséquentes permettront de déterminer le rôle des régions proximales, généralement associées à l'activation du gène au stade fœtal, sur la régulation de la commutation.

Finalement, j'ai également soulevé quelques questions quant à la fidélité avec laquelle la commutation de l'hémoglobine fœtale à adulte du locus humain peut être reproduite chez la souris. En effet, la présence d'une commutation tardive chez la souris PBGD γ -YAC suggère que certains éléments régulateurs, cachés chez la souris, pourraient jouer un rôle significatif chez l'humain.

Bibliographie

1. Silverman L, Glick D: Measurement of protein concentration by quantitative electron microscopy. *J.Cell Biol.* 1969; 40: 773-8
2. Hardison R: Hemoglobins from bacteria to man: evolution of different patterns of gene expression. *J.Exp.Biol.* 1998; 201 (Pt 8): 1099-117
3. Wittenberg JB, Bolognesi M, Wittenberg BA, Guertin M: Truncated hemoglobins: A new family of hemoglobins widely distributed in bacteria, unicellular eukaryotes and plants. *J.Biol.Chem.* 2001; 277: 871-4
4. Frey AD, Kallio PT: Bacterial hemoglobins and flavohemoglobins: versatile proteins and their impact on microbiology and biotechnology. *FEMS Microbiol. Rev.* 2003; 27: 525-45
5. Kundu S, Trent JT, III, Hargrove MS: Plants, humans and hemoglobins. *Trends Plant Sci.* 2003; 8: 387-93
6. Shikama K, Matsuoka A, Iwaasa H: The unique structures of protozoan myoglobin and yeast hemoglobin: an evolutionary diversity. *Int.J.Biochem.Cell Biol.* 1995; 27: 1107-15
7. Respiratory pigments in animals: relation structure-function. Edited by Lamy J, Truchot J-P, Gilles R. Berlin, Springer-Verlag, 1985, pp 1-175
8. Terwilliger NB: Functional adaptations of oxygen-transport proteins. *J.Exp.Biol.* 1998; 201: 1085-98
9. Vinogradov SN, Walz DA, Pohajdak B, Moens L, Kapp OH, Suzuki T, Trotman CN: Adventitious variability? The amino acid sequences of nonvertebrate globins. *Comp.Biochem.Physiol. B* 1993; 106: 1-26
10. Miranda JJ, Maillett DH, Soman J, Olson JS: Thermoglobin, oxygen-avid hemoglobin in a bacterial hyperthermophile. *J.Biol.Chem.* 2005; 280: 36754-61
11. Kleinschmidt T, Keyl HG, Braunitzer G: Comparison of insect hemoglobins (Erythrocruorins) from *Chironomus thummi thummi* and *Chironomus thummi piger* (*Diptera*). The primary structure of the monomeric hemoglobin CTP III. *Biol. Chem.Hoppe-Seyler* 1989; 370: 839-45
12. Ohno S, Morrison M: Multiple gene loci for the monomeric hemoglobin of the hagfish (*Eptatretus stoutii*). *Science* 1966; 154: 1034-5

13. Kao WY, Hankeln T, Schmidt ER, Bergtrom G: Sequence and evolution of the gene for the monomeric globin I and its linkage to genes coding for dimeric globins in the insect *Chironomus thummi*. *J.Mol.Evol.* 1995; 40: 354-61
14. Kawano K, Iwasaki N, Suzuki T: Notable diversity in hemoglobin expression patterns among species of the deep-sea clam, *Calypptogena*. *Cell Mol.Life Sci.* 2003; 60: 1952-6
15. Tarricone C, Galizzi A, Coda A, Ascenzi P, Bolognesi M: Unusual structure of the oxygen-binding site in the dimeric bacterial hemoglobin from *Vitreoscilla* sp. *Structure.* 1997; 5: 497-507
16. Pathania R, Navani NK, Gardner AM, Gardner PR, Dikshit KL: Nitric oxide scavenging and detoxification by the *Mycobacterium tuberculosis* hemoglobin, HbN in *Escherichia coli*. *Mol.Microbiol.* 2002; 45: 1303-14
17. Fago A, Giangiacomo L, D'Avino R, Carratore V, Romano M, Boffi A, Chiancone E: Hagfish hemoglobins: structure, function, and oxygen-linked association. *J.Biol.Chem.* 2001; 276: 27415-23
18. Zal F, Lallier FH, Wall JS, Vinogradov SN, Toulmond A: The multi-hemoglobin system of the hydrothermal vent tube worm *Riftia pachyptila*. I. Reexamination of the number and masses of its constituents. *J.Biol.Chem.* 1996; 271: 8869-74
19. Kirschvink JL, Gaidos EJ, Bertani LE, Beukes NJ, Gutzmer J, Maepa LN, Steinberger RE: Paleoproterozoic snowball earth: extreme climatic and geochemical global change and its biological consequences. *Proc.Natl.Acad.Sci.U.S.A* 2000; 97: 1400-5
20. Downie JA: Legume hemoglobins: symbiotic nitrogen fixation needs bloody nodules. *Curr.Biol.* 2005; 15: R196-R198
21. Wittenberg JB, Wittenberg BA: Myoglobin function reassessed. *J.Exp.Biol.* 2003; 206: 2011-20
22. Jensen FB: Red blood cell pH, the Bohr effect, and other oxygenation-linked phenomena in blood O₂ and CO₂ transport. *Acta Physiol.Scand.* 2004; 182: 215-27
23. Coates ML: Hemoglobin function in the vertebrates: an evolutionary model. *J.Mol.Evol.* 1975; 6: 285-307
24. Wells RM: Evolution of hemoglobin function: molecular adaptations to environment. *Clin.Exp.Pharmacol.Physiol.* 1999; 26: 591-5

25. Mairbaurl H: Red blood cell function in hypoxia at altitude and exercise. *Int.J. Sports Med.* 1994; 15: 51-63
26. Giardina B, Mosca D, De Rosa MC: The Bohr effect of hæmoglobin in vertebrates: an example of molecular adaptation to different physiological requirements. *Acta Physiol.Scand.* 2004; 182: 229-44
27. Fantoni A, Farace MG, Gambari R: Embryonic hemoglobins in man and other mammals. *Blood* 1981; 57: 623-33
28. Kitchen H, Brett I: Embryonic and fetal hemoglobin in animals. *Ann.N.Y.Acad.Sci.* 1974; 241: 653-71
29. Bunn HF: Evolution of mammalian hemoglobin function. *Blood* 1981; 58: 189-97
30. Blunt MH, Kitchens JL, Mayson SM, Huisman TH: Red cell 2,3-diphosphoglycerate and oxygen affinity in newborn goats and sheep. *Proc.Soc.Exp.Biol.Med.* 1971; 138: 800-3
31. Bard H, Teasdale F: Red cell oxygen affinity, hemoglobin type, 2,3-diphosphoglycerate, and pH as a function of fetal development. *Pediatrics* 1979; 64: 483-7
32. Tyuma I, Shimizu K: Different response to organic phosphates of human fetal and adult hemoglobins. *Arch.Biochem.Biophys.* 1969; 129: 404-5
33. Kaltsoya A, Fessas P, Stavropoulos A: Hemoglobins of early human embryonic development. *Science* 1966; 153: 1417-8
34. Capp GL, Rigas DA, Jones RT: Hemoglobin Portland 1: a new human hemoglobin unique in structure. *Science* 1967; 157: 65-6
35. Capp GL, Rigas DA, Jones RT: Evidence for a new hæmoglobin chain (ζ -chain). *Nature* 1970; 228: 278-80
36. Gale RE, Clegg JB, Huehns ER: Human embryonic hæmoglobins Gower 1 and Gower 2. *Nature* 1979; 280: 162-4
37. He Z, Russell JE: Expression, purification, and characterization of human hemoglobins Gower-1 ($\zeta_2\epsilon_2$), Gower-2 ($\alpha_2\epsilon_2$), and Portland-2 ($\zeta_2\beta_2$) assembled in complex transgenic-knockout mice. *Blood* 2001; 97: 1099-105
38. Beauchemin H, Blouin MJ, Trudel M: Differential regulatory and compensatory responses in hematopoiesis/erythropoiesis in α - and β -globin hemizygous mice. *J.Biol.Chem.* 2004; 279: 19471-80

39. Teichmann LK: Über die Kristallisation der organischen Bestandteile des Bluts. *Z.Rat.Med.* 1853; 3: 375-99. Cité dans Edsall JT: Blood and hemoglobin: The evolution of knowledge of functional adaptation in a biochemical system. *J.His.Biol.* 1972; 5: 205-57
40. Stokes GG: On the Reduction and Oxidation of the Colouring Matter of the Blood. *Proc.Roy.Soc.Lond.* 1864; 13: 355-64
41. Hoppe-Seyler F: Ueber die chemischen und optischen Eigenschaften des Blutfarbstoffs. *Arch.F.path.Anat.U.Physiol.(Virchows Archiv.)* 1864; 29: 233-5
42. Preyer WT: Die Blutkrystalle. Jena, 1871, pp 1-200. Cité dans Lankester ER: A contribution to the knowledge of hæmoglobin. *Proc.Roy.Soc.Lond.* 1872-1873; 21: 70-81
43. Henderson LJ: Blood: a study in general physiology. New Heaven, Yale University Press, 1928, pp 1-297
44. Neel JV: The Inheritance of Sickle Cell Anemia. *Science* 1949; 110: 64-6
45. Pauling L, Itano HA, et al: Sickle cell anemia a molecular disease. *Science* 1949; 110: 543-9
46. Rhinesmith HS, Schroeder WA, Pauling L: The N-Terminal Amino Acid Residues of Normal Adult Human Hemoglobin: A Quantitative Study of Certain Aspects of Sanger's DNP-Method. *J.Am.Chem.Soc.* 1957; 79: 609-15
47. Rhinesmith HS, Schroeder WA, Martin N: The N-Terminal Sequence of the α Chains¹ of Normal Adult Human Hemoglobin. *J.Am.Chem.Soc.* 1958; 80: 3358-61
48. Vinograd JR, Hutchison WD, Schroeder WA: C14-Hybrids of human hemoglobins. II. The identification of the aberrant chain in human hemoglobin S. *J.Am.Chem.Soc.* 1959; 81: 3168-9
49. Ingram VM: Abnormal human hæmoglobins. III. The chemical difference between normal and sickle cell hæmoglobins. *Biochim.Biophys.Acta* 1959; 36: 402-11
50. Perutz MF, Rossmann MG, Cullis AF, Muirhead H, Will G, North ACT: Structure of hæmoglobin. A three-dimensional Fourier synthesis at 5.5Å resolution, obtained by X-ray analysis. *Nature* 1960; 185: 416-22
51. Perutz MF: X-ray analysis of hæmoglobin, Nobel lectures, Chemistry 1942-1962. Amsterdam, Elsevier Publishing Company, 1964, pp 653-73

52. Mayow J: *Tertius de respiratione fœtus in utero et ovo, Tractatus quinque medico-physici*. 1674. Traduction anglaise publiée dans: On the respiration of the fœtus in the uterus and in the egg, Medico-physical works being a translation of *tractatus quinque medico-physici*. Edinburgh, The Alembic Club, 1907, pp 211-27
53. Zuntz N: Ueber die respiration des säugethier-fœtus. *Pflugers Archiv*. 1877; 14: 605-27. Cité dans Edsall JT: Blood and hemoglobin: The evolution of knowledge of functional adaptation in a biochemical system. *J.His.Biol*. 1972; 5: 205-57
54. Huggett ASG: Fœtal blood-gas tension and gas transfusion through the placenta of the goat. *J.Physiol*. 1927; 62: 373-84
55. McCarthy EF: A comparison of fœtal and maternal hæmoglobins in the goat. *J.Physiol*. 1933; 80: 206-12
56. Hall FG: A spectroscopic comparison of fœtal and maternal blood of the rabbit and goat. *J.Physiol*. 1934; 82: 33-7
57. Hall FG: Hæmoglobin function in the developing chick. *J.Physiol*. 1934; 83: 222-8
58. McCarthy EF: The oxygen affinity of human maternal and foetal hæmoglobin. *J.Physiol*. 1943; 102: 55-61
59. Haselhorst G, Stromberger K: Über den gasgehalt des nabelschnurblutes vor und nach der geburt des kindes und über den gasaustausch in der plazenta. *Z.Geburtsch. Gynäk* 1931; 100: 49-58
60. Allen DW, Wyman J, Jr., Smith CA: The oxygen equilibrium of fetal and adult human hemoglobin. *J.Biol.Chem*. 1953; 203: 81-7
61. Greenwald I: A new type of phosphoric acid compound isolated from blood, with some remarks on the effect of substitution on the rotation of *l*-glyceric acid. *J.Biol.Chem*. 1925; 63: 339-49
62. Chanutin A, Curnish RR: Effect of organic and inorganic phosphates on the oxygen equilibrium of human erythrocytes. *Arch.Biochem.Biophys*. 1967; 121: 96-102
63. Benesch R, Benesch RE, Yu CI: Reciprocal binding of oxygen and diphosphoglycerate by human hemoglobin. *Proc.Natl.Acad.Sci.U.S.A* 1968; 59: 526-32
64. Huehns ER, Flynn FV, Butler EA, Beaven GH: Two new hæmoglobin variants in a very young human embryo. *Nature* 1961; 189: 496-7

65. Scott AF, Bunn HF, Brush AH: Functional aspects of hemoglobin evolution in the mammals. *J.Mol.Evol.* 1976; 8: 311-6
66. Scott AF, Bunn HF, Brush AH: The phylogenetic distribution of red cell 2,3 diphosphoglycerate and its interaction with mammalian hemoglobins. *J.Exp.Zool.* 1977; 201: 269-88
67. Watson JD, Crick FH: Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 1953; 171: 737-8
68. Kelly TJ, Jr., Smith HO: A restriction enzyme from *Hemophilus influenzae*. II. *J.Mol.Biol.* 1970; 51: 393-409
69. Smith HO, Wilcox KW: A restriction enzyme from *Hemophilus influenzae*. I. Purification and general properties. *J.Mol.Biol.* 1970; 51: 379-91
70. Smith HO, Nathans D: Letter: A suggested nomenclature for bacterial host modification and restriction systems and their enzymes. *J.Mol.Biol.* 1973; 81: 419-23
71. Temin HM, Mizutani S: RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* 1970; 226: 1211-3
72. Baltimore D: RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* 1970; 226: 1209-11
73. Cohen SN, Chang AC: Recircularization and autonomous replication of a sheared R-factor DNA segment in *Escherichia coli* transformants. *Proc.Natl.Acad.Sci. U.S.A* 1973; 70: 1293-7
74. Cohen SN, Chang AC, Boyer HW, Helling RB: Construction of biologically functional bacterial plasmids in vitro. *Proc.Natl.Acad.Sci.U.S.A* 1973; 70: 3240-4
75. Rougeon F, Kourilsky P, Mach B: Insertion of a rabbit β -globin gene sequence into an *E. coli* plasmid. *Nucleic Acids Res.* 1975; 2: 2365-78
76. Jeffreys AJ, Flavell RA: The rabbit β -globin gene contains a large large insert in the coding sequence. *Cell* 1977; 12: 1097-108
77. Flavell RA, Waalwijk C, Jeffreys AJ: The structure of the rabbit β -globin gene. *Biochem.Soc.Trans.* 1978; 6: 742-6

78. Tilghman SM, Curtis PJ, Tiemeier DC, Leder P, Weissmann C: The intervening sequence of a mouse β -globin gene is transcribed within the 15S beta-globin mRNA precursor. *Proc.Natl.Acad.Sci.U.S.A* 1978; 75: 1309-13
79. Leder P, Tilghman SM, Tiemeier DC, Polsky FI, Seidman JG, Edgell MH, Enquist LW, Leder A, Norman B: The cloning of mouse globin and surrounding gene sequences in bacteriophage lambda. *Cold Spring Harb.Symp.Quant.Biol.* 1978; 42 Pt 2: 915-20
80. O'Farrell PZ, Cordell B, Valenzuela P, Rutter WJ, Goodman HM: Structure and processing of yeast precursor tRNAs containing intervening sequences. *Nature* 1978; 274: 438-45
81. van den BJ, van OA, Mantei N, Schambock A, Grosveld G, Flavell RA, Weissmann C: Comparison of cloned rabbit and mouse β -globin genes showing strong evolutionary divergence of two homologous pairs of introns. *Nature* 1978; 276: 37-44
82. Bernard O, Hozumi N, Tonegawa S: Sequences of mouse immunoglobulin light chain genes before and after somatic changes. *Cell* 1978; 15: 1133-44
83. Catterall JF, O'Malley BW, Robertson MA, Staden R, Tanaka Y, Brownlee GG: Nucleotide sequence homology at 12 intron-exon junctions in the chick ovalbumin gene. *Nature* 1978; 275: 510-3
84. Abelson J: RNA processing and the intervening sequence problem. *Annu.Rev. Biochem.* 1979; 48: 1035-69
85. Leder P: Discontinuous genes. *N.Engl.J.Med.* 1978; 298: 1079-81
86. Smithies O, Blechl AE, niston-Thompson K, Newell N, Richards JE, Slightom JL, Tucker PW, Blattner FR: Cloning human fetal γ globin and mouse α -type globin DNA: characterization and partial sequencing. *Science* 1978; 202: 1284-9
87. Lawn RM, Fritsch EF, Parker RC, Blake G, Maniatis T: The isolation and characterization of linked δ - and β -globin genes from a cloned library of human DNA. *Cell* 1978; 15: 1157-74
88. Deisseroth A, Nienhuis A, Lawrence J, Giles R, Turner P, Ruddle FH: Chromosomal localization of human β globin gene on human chromosome 11 in somatic cell hybrids. *Proc.Natl.Acad.Sci.U.S.A* 1978; 75: 1456-60

89. Ramirez F, Burns AL, Mears JG, Spence S, Starkman D, Bank A: Isolation and characterization of cloned human fetal globin genes. *Nucleic Acids Res.* 1979; 7: 1147-62
90. Proudfoot NJ, Baralle FE: Molecular cloning of human ϵ -globin gene. *Proc.Natl. Acad.Sci.U.S.A* 1979; 76: 5435-9
91. Fritsch EF, Shen CK, Lawn RM, Maniatis T: The organization of repetitive sequences in mammalian globin gene clusters. *Cold Spring Harb.Symp.Quant.Biol.* 1981; 45 Pt 2: 761-5
92. Shen SH, Smithies O: Human globin $\psi\beta^2$ is not a globin-related sequence. *Nucleic Acids Res.* 1982; 10: 7809-18
93. Chada K, Magram J, Raphael K, Radice G, Lacy E, Costantini F: Specific expression of a foreign β -globin gene in erythroid cells of transgenic mice. *Nature* 1985; 314: 377-80
94. Costantini F, Radice G, Magram J, Stamatoyannopoulos G, Papayannopoulou T, Chada K: Developmental regulation of human globin genes in transgenic mice. *Cold Spring Harb.Symp.Quant.Biol.* 1985; 50: 361-70
95. Townes TM, Lingrel JB, Chen HY, Brinster RL, Palmiter RD: Erythroid-specific expression of human β -globin genes in transgenic mice. *EMBO J.* 1985; 4: 1715-23
96. Kollias G, Wrighton N, Hurst J, Grosveld F: Regulated expression of human γ -, β -, and hybrid $\gamma\beta$ -globin genes in transgenic mice: manipulation of the developmental expression patterns. *Cell* 1986; 46: 89-94
97. Tuan D, Solomon W, Li Q, London IM: The " β -like-globin" gene domain in human erythroid cells. *Proc.Natl.Acad.Sci.U.S.A* 1985; 82: 6384-8
98. Forrester WC, Takegawa S, Papayannopoulou T, Stamatoyannopoulos G, Groudine M: Evidence for a locus activation region: the formation of developmentally stable hypersensitive sites in globin-expressing hybrids. *Nucleic Acids Res.* 1987; 15: 10159-77
99. Blom van AG, Hanscombe O, Grosveld F, Greaves DR: The β -globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner. *Cell* 1989; 56: 969-77

100. Grosveld F, van Assendelft GB, Greaves DR, Kollias G: Position-independent, high-level expression of the human β -globin gene in transgenic mice. *Cell* 1987; 51: 975-85
101. Ryan TM, Behringer RR, Martin NC, Townes TM, Palmiter RD, Brinster RL: A single erythroid-specific DNase I super-hypersensitive site activates high levels of human β -globin gene expression in transgenic mice. *Genes Dev.* 1989; 3: 314-23
102. Talbot D, Collis P, Antoniou M, Vidal M, Grosveld F, Greaves DR: A dominant control region from the human β -globin locus conferring integration site-independent gene expression. *Nature* 1989; 338: 352-5
103. Behringer RR, Ryan TM, Palmiter RD, Brinster RL, Townes TM: Human γ - to β -globin gene switching in transgenic mice. *Genes Dev.* 1990; 4: 380-9
104. Orkin SH: Globin gene regulation and switching: circa 1990. *Cell* 1990; 63: 665-72
105. Wahl GM, Lewis KA, Ruiz JC, Rothenberg B, Zhao J, Evans GA: Cosmid vectors for rapid genomic walking, restriction mapping, and gene transfer. *Proc.Natl.Acad. Sci.U.S.A* 1987; 84: 2160-4
106. Burke DT, Carle GF, Olson MV: Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* 1987; 236: 806-12
107. Barton MC, Hoekstra MF, Emerson BM: Site-directed, recombination-mediated mutagenesis of a complex gene locus. *Nucleic Acids Res.* 1990; 18: 7349-55
108. Gaensler KM, Burmeister M, Brownstein BH, Taillon-Miller P, Myers RM: Physical mapping of yeast artificial chromosomes containing sequences from the human β -globin gene region. *Genomics* 1991; 10: 976-84
109. Schedl A, Beermann F, Thies E, Montoliu L, Kelsey G, Schutz G: Transgenic mice generated by pronuclear injection of a yeast artificial chromosome. *Nucleic Acids Res.* 1992; 20: 3073-7
110. Peterson KR, Clegg CH, Huxley C, Josephson BM, Haugen HS, Furukawa T, Stamatoyannopoulos G: Transgenic mice containing a 248-kb yeast artificial chromosome carrying the human β -globin locus display proper developmental control of human globin genes. *Proc.Natl.Acad.Sci.U.S.A* 1993; 90: 7593-7
111. Gaensler KM, Kitamura M, Kan YW: Germ-line transmission and developmental regulation of a 150-kb yeast artificial chromosome containing the human β -globin locus in transgenic mice. *Proc.Natl.Acad.Sci.U.S.A* 1993; 90: 11381-5

112. Furukawa T, Navas PA, Josephson BM, Peterson KR, Papayannopoulou T, Stamatoyannopoulos G: Coexpression of ϵ , $G\gamma$ and $A\gamma$ globin mRNA in embryonic red blood cells from a single copy β -YAC transgenic mouse. *Blood Cells Mol.Dis.* 1995; 21: 168-78
113. Tavassoli M: Embryonic and fetal hemopoiesis: an overview. *Blood Cells* 1991; 17: 269-81
114. Hecht F, Motulsky AG, Lemire RJ, Shepard TE: Predominance of hemoglobin Gower 1 in early human embryonic development. *Science* 1966; 152: 91-2
115. Peschle C, Mavilio F, Care A, Migliaccio G, Migliaccio AR, Salvo G, Samoggia P, Petti S, Guerriero R, Marinucci M, .: Hæmoglobin switching in human embryos: asynchrony of $\zeta \rightarrow \alpha$ and $\epsilon \rightarrow \gamma$ -globin switches in primitive and definite erythropoietic lineage. *Nature* 1985; 313: 235-8
116. Neumeyer P, Betke K: Hæmoglobin A2 in newborn infants of different maturity. *Eur.J.Pediatr.* 1987; 146: 598-600
117. Alter BP: The $G\gamma:A\gamma$ composition of fetal hemoglobin in fetuses and newborns. *Blood* 1979; 54: 1158-63
118. Hardison RC: Comparison of the β -like globin gene families of rabbits and humans indicates that the gene cluster 5'- ϵ - γ - δ - β -3' predates the mammalian radiation. *Mol. Biol.Evol.* 1984; 1: 390-410
119. Hardies SC, Edgell MH, Hutchison CA, III: Evolution of the mammalian β -globin gene cluster. *J.Biol.Chem.* 1984; 259: 3748-56
120. Tang DC, Ebb D, Hardison RC, Rodgers GP: Restoration of the CCAAT box or insertion of the CACCC motif activates δ -globin gene expression. *Blood* 1997; 90: 421-7
121. Ristaldi MS, Casula S, Porcu S, Marongiu MF, Pirastu M, Cao A: Activation of the δ -globin gene by the β -globin gene CACCC motif. *Blood Cells Mol.Dis.* 1999; 25: 193-209
122. Inagaki K, Inagaki J, Dumoulin A, Padovan JC, Chait BT, Popowicz A, Manning LR, Manning JM: Expression and properties of recombinant HbA2 ($\alpha_2\delta_2$) and hybrids containing δ - β sequences. *J.Protein Chem.* 2000; 19: 649-62
123. Kamuzora H: Recent developments in foetal hæmoglobin research. *Humangenetik.* 1975; 30: 197-205

124. Hardison R, Slightom JL, Gumucio DL, Goodman M, Stojanovic N, Miller W: Locus control regions of mammalian β -globin gene clusters: combining phylogenetic analyses and experimental results to gain functional insights. *Gene* 1997; 205: 73-94
125. Bungert J, Tanimoto K, Patel S, Liu Q, Fear M, Engel JD: Hypersensitive site 2 specifies a unique function within the human β -globin locus control region to stimulate globin gene transcription. *Mol.Cell Biol.* 1999; 19: 3062-72
126. Navas PA, Li Q, Peterson KR, Swank RA, Rohde A, Roy J, Stamatoyannopoulos G: Activation of the β -like globin genes in transgenic mice is dependent on the presence of the β -locus control region. *Hum.Mol.Genet.* 2002; 11: 893-903
127. Navas PA, Swank RA, Yu M, Peterson KR, Stamatoyannopoulos G: Mutation of a transcriptional motif of a distant regulatory element reduces the expression of embryonic and fetal globin genes. *Hum.Mol.Genet.* 2003; 12: 2941-8
128. Navas PA, Peterson KR, Li Q, McArthur M, Stamatoyannopoulos G: The 5'HS4 core element of the human β -globin locus control region is required for high-level globin gene expression in definitive but not in primitive erythropoiesis. *J.Mol.Biol.* 2001; 312: 17-26
129. Fang X, Sun J, Xiang P, Yu M, Navas PA, Peterson KR, Stamatoyannopoulos G, Li Q: Synergistic and additive properties of the β -globin locus control region (LCR) revealed by 5'HS3 deletion mutations: implication for LCR chromatin architecture. *Mol.Cell Biol.* 2005; 25: 7033-41
130. Antequera F, Bird A: Number of CpG islands and genes in human and mouse. *Proc. Natl.Acad.Sci.U.S.A* 1993; 90: 11995-9
131. Bulger M, van Doorninck JH, Saitoh N, Telling A, Farrell C, Bender MA, Felsenfeld G, Axel R, Groudine M, von Doorninck JH: Conservation of sequence and structure flanking the mouse and human β -globin loci: the β -globin genes are embedded within an array of odorant receptor genes. *Proc.Natl.Acad.Sci.U.S.A* 1999; 96: 5129-34
132. Bulger M, Bender MA, van Doorninck JH, Wertman B, Farrell CM, Felsenfeld G, Groudine M, Hardison R: Comparative structural and functional analysis of the olfactory receptor genes flanking the human and mouse beta-globin gene clusters. *Proc.Natl.Acad.Sci.U.S.A* 2000; 97: 14560-5
133. Li Q, Stamatoyannopoulos G: Hypersensitive site 5 of the human β locus control region functions as a chromatin insulator. *Blood* 1994; 84: 1399-401

134. Li Q, Zhang M, Han H, Rohde A, Stamatoyannopoulos G: Evidence that DNase I hypersensitive site 5 of the human β -globin locus control region functions as a chromosomal insulator in transgenic mice. *Nucleic Acids Res.* 2002; 30: 2484-91
135. Xiang P, Han H, Barkess G, Olave I, Fang X, Yin W, Stamatoyannopoulos G, Li Q: Juxtaposition of the HPFH2 enhancer is not sufficient to reactivate the γ -globin gene in adult erythropoiesis. *Hum.Mol.Genet.* 2005; 14: 3047-56
136. Bender MA, Byron R, Ragozy T, Telling A, Bulger M, Groudine M: Flanking HS-62.5 and 3' HS1, and regions upstream of the LCR, are not required for β -globin transcription. *Blood* 2006; 108: 1395-401
137. Long Q, Bengra C, Li C, Kutlar F, Tuan D: A long terminal repeat of the human endogenous retrovirus ERV-9 is located in the 5' boundary area of the human β -globin locus control region. *Genomics* 1998; 54: 542-55
138. Tuan D, Pi W, Zhu X, Wu M, Wang Y, Ling J: Long-range function of the ERV-9 LTR in transcriptional regulation of the globin genes. Résumé présenté à: the 15th conference on hemoglobin switching. St. John's College, Oxford, UK, 2006,
139. Sarnaik SA: Thalassemia and related hemoglobinopathies. *Indian J.Pediatr.* 2005; 72: 319-24
140. Olivieri NF: The β -thalassemias. *N.Engl.J.Med.* 1999; 341: 99-109
141. Hardison RC, Chui DH, Giardine B, Riemer C, Patrinos GP, Anagnou N, Miller W, Wajcman H: HbVar: A relational database of human hemoglobin variants and thalassemia mutations at the globin gene server. *Hum.Mutat.* 2002; 19: 225-33
142. Giardine B, van BS, Kaimakis P, Riemer C, Miller W, Samara M, Kollia P, Anagnou NP, Chui DH, Wajcman H, Hardison RC, Patrinos GP: HbVar database of human hemoglobin variants and thalassemia mutations: 2007 update. *Hum.Mutat.* 2007; 28: 206
143. Forget BG: Molecular basis of hereditary persistence of fetal hemoglobin. *Ann.N.Y. Acad.Sci.* 1998; 850: 38-44
144. Watson J: The significance of the paucity of sickle cells in newborn Negro infants. *Am.J.Med.Sci.* 1948; 215: 419-23
145. Cooley TB, Lee P: Series of cases of splenomegaly in children with anaemia and peculiar bone changes. *Trans.Am.Pediatr.Soc.* 1925; 37: 29-33

146. Loukopoulos D: Thalassemia: genotypes and phenotypes. *Ann.Hematol.* 1991; 62: 145-50
147. Huisman TH: A short review of human γ -globin gene anomalies. *Acta Haematol.* 1987; 78: 80-4
148. Bouva MJ, Harteveld CL, van DP, Giordano PC: Known and new δ globin gene mutations and their diagnostic significance. *Haematologica* 2006; 91: 129-32
149. Rooks H, Bergounioux J, Game L, Close JP, Osborne C, Best S, Senior T, Height S, Thompson R, Hadzic N, Fraser P, Bolton-Maggs P, Thein SL: Heterogeneity of the $\epsilon\gamma\delta\beta$ -thalassaemias: characterization of three novel English deletions. *Br.J. Haematol.* 2005; 128: 722-9
150. Taramelli R, Kioussis D, Vanin E, Bartram K, Groffen J, Hurst J, Grosveld FG: $\gamma\delta\beta$ -thalassaemias 1 and 2 are the result of a 100 kbp deletion in the human β -globin cluster. *Nucleic Acids Res.* 1986; 14: 7017-29
151. Diaz-Chico JC, Huang HJ, Juricic D, Efremov GD, Wadsworth LD, Huisman TH: Two new large deletions resulting in $\epsilon\gamma\delta\beta$ -thalassemia. *Acta Haematol.* 1988; 80: 79-84
152. Rouyer-Fessard P, Scott MD, Leroy-Viard K, Garel MC, Bachir D, Galacteros F, Beuzard Y: Fate of α -hemoglobin chains and erythrocyte defects in β -thalassemia. *Ann.N.Y.Acad.Sci.* 1990; 612: 106-17
153. Schrier SL: Thalassemia: pathophysiology of red cell changes. *Annu.Rev.Med.* 1994; 45: 211-8
154. Shiokawa S, Yamada H, Takihara Y, Matsunaga E, Ohba Y, Yamamoto K, Fukumaki Y: Molecular analysis of Japanese $\delta\beta$ -thalassemia. *Blood* 1988; 72: 1771-6
155. Mastropietro F, Modiano G, Cappabianca M, Foglietta E, D'Asero C, Mezzabotta M, Ponzini D, Maffei L, Amato A, Lerone M, Grisanti P, Di Biagio P, Rinaldi S, Bianco I: Factors regulating Hb F synthesis in thalassaemic diseases. *BMC.Blood Disord.* 2002; 2: 2
156. Blouin MJ, Beauchemin H, Wright A, De Paepe M, Sorette M, Bleau AM, Nakamoto B, Ou CN, Stamatoyannopoulos G, Trudel M: Genetic correction of sickle cell disease: insights using transgenic mouse models. *Nat.Med.* 2000; 6: 177-82

157. Chakalova L, Osborne CS, Dai YF, Goyenechea B, Metaxotou-Mavromati A, Kattamis A, Kattamis C, Fraser P: The Corfu $\delta\beta$ thalassemia deletion disrupts γ -globin gene silencing and reveals post-transcriptional regulation of HbF expression. *Blood* 2005; 105: 2154-60
158. Svasti S, Paksua S, Nuchprayoon I, Winichagoon P, Fucharoen S: Characterization of a novel deletion causing $(\delta\beta)^0$ -thalassemia in a Thai family. *Am.J.Hematol.* 2007; 82: 155-61
159. Coleman MB, Adams JG, III, Steinberg MH, Plonczynski MW, Harrell AH, Castro O, Winter WP: $G\gamma^A\gamma(\beta^+)$ hereditary persistence of fetal hemoglobin: the $G\gamma$ -158 C \rightarrow T mutation in cis to the -175 T \rightarrow C mutation of the $A\gamma$ -globin gene results in increased $G\gamma$ -globin synthesis. *Am.J.Hematol.* 1993; 42: 186-90
160. Stoming TA, Stoming GS, Lanclos KD, Fei YJ, Altay C, Kutlar F, Huisman TH: An $A\gamma$ type of nondeletional hereditary persistence of fetal hemoglobin with a T \rightarrow C mutation at position -175 to the cap site of the $A\gamma$ globin gene. *Blood* 1989; 73: 329-33
161. Zertal-Zidani S, Merghoub T, Ducrocq R, Gerard N, Satta D, Krishnamoorthy R: A novel C \rightarrow A transversion within the distal CCAAT motif of the $G\gamma$ -globin gene in the Algerian $G\gamma\beta^+$ -hereditary persistence of fetal hemoglobin. *Hemoglobin* 1999; 23: 159-69
162. Bouva MJ, Hartevelde CL, Bakker-Verweij G, van DP, Giordano PC: $G\gamma$ -37 (A \rightarrow T): a new nondeletional hereditary persistence of fetal hemoglobin determinant associated with the rare codon 91(+T) δ^0 -thalassemia. *Hemoglobin* 2006; 30: 371-7
163. Goodman M, Moore GW: Phylogeny of Hemoglobin. *Sys.Zool.* 1973; 22: 508-32
164. Benz EJ, Jr.: Hemoglobin switching in animals. *Tex.Rep.Biol.Med.* 1980; 40: 111-23
165. Tagle DA, Koop BF, Goodman M, Slightom JL, Hess DL, Jones RT: Embryonic ϵ and γ globin genes of a prosimian primate (*Galago crassicaudatus*). Nucleotide and amino acid sequences, developmental regulation and phylogenetic footprints. *J.Mol. Biol.* 1988; 203: 439-55
166. Tagle DA, Slightom JL, Jones RT, Goodman M: Concerted evolution led to high expression of a prosimian primate δ globin gene locus. *J.Biol.Chem.* 1991; 266: 7469-80

167. Martin SL, Vincent KA, Wilson AC: Rise and fall of the δ globin gene. *J.Mol.Biol.* 1983; 164: 513-28
168. Johnson RM, Buck S, Chiu C, Schneider H, Sampaio I, Gage DA, Shen TL, Schneider MP, Muniz JA, Gumucio DL, Goodman M: Fetal globin expression in New World monkeys. *J.Biol.Chem.* 1996; 271: 30298
169. Vincent KA, Wilson AC: Evolution and transcription of old world monkey globin genes. *J.Mol.Biol.* 1989; 207: 465-80
170. Bauer C, Tamm R, Petschow D, Bartels R, Bartels H: Oxygen affinity and allosteric effects of embryonic mouse haemoglobins. *Nature* 1975; 257: 333-4
171. Kohn B, Reilly MP, Asakura T, Giger U: Polymorphism of feline β -globins studied by reversed-phase high-performance liquid chromatography. *Am.J.Vet.Res.* 1998; 59: 830-5
172. Kohn B, Henthorn PS, Rajpurohit Y, Reilly MP, Asakura T, Giger U: Feline adult β -globin polymorphism reflected in restriction fragment length patterns. *J.Hered.* 1999; 90: 177-81
173. Wesolowski SR, Raney NE, Ernst CW: Developmental changes in the fetal pig transcriptome. *Physiol.Genomics* 2004; 16: 268-74
174. Kleihauer E, Stoffler G: Embryonic hemoglobins of different animal species. Quantitative and qualitative data about production and properties of hemoglobins during early developmental stages of pig, cattle and sheep. *Mol.Gen.Genet.* 1968; 101: 59-69
175. Zinkl J, Kaneko JJ: Erythrocyte 2,3-diphosphoglycerate in normal and porphyric fetal, neonatal and adult cattle. *Comp.Biochem.Physiol. A* 1973; 45: 699-704
176. Schimenti JC, Duncan CH: Structure and organization of the bovine β -globin genes. *Mol.Biol.Evol.* 1985; 2: 514-25
177. Schimenti JC, Duncan CH: Concerted evolution of the cow $\epsilon 2$ and $\epsilon 4$ β -globin genes. *Mol.Biol.Evol.* 1985; 2: 505-13
178. Luse DS, Haynes JR, VanLeeuwen D, Schon EA, Cleary ML, Shapiro SG, Lingrel JB, Roeder RG: Transcription of the β -like globin genes and pseudogenes of the goat in a cell-free system. *Nucleic Acids Res.* 1981; 9: 4339-54

179. Serreri E, Hadjisterkotis E, Ei Sherbini S, Palici di Suni M, Manca L, Pirastu M, Masala B: Structure and organization of Caprinæ globin genes. *Z.Jagdwiss.* 2002; 48: 274-80
180. Braend M, Tucker EM: Hemoglobin types in Saanen goats and Barbary sheep: genetic and comparative aspects. *Biochem.Genet.* 1988; 26: 511-8
181. Rubino, G., Petazzi, F., Lacinio, R., Caroli, A., and Pieragostini, E. Hematological patterns and globin genotypes in *Gentile di puglia* sheep. Communication personelle: 2004.
182. Huisman TH, Lewis JP, Blunt MH, Adams HR, Miller A, Dozy AM, Boyd EM: Hemoglobin C in newborn sheep and goats: a possible explanation for its function and biosynthesis. *Pediatr.Res.* 1969; 3: 189-98
183. Hammerberg B, Brett I, Kitchen H: Ontogeny of hemoglobins in sheep. *Ann.N.Y. Acad.Sci.* 1974; 241: 672-82
184. Young NS, Benz EJ, Jr., Kantor JA, Kretschmer P, Nienhuis AW: Hemoglobin switching in sheep: only the γ gene is in the active conformation in fetal liver but all the γ and β genes are in the active conformation in bone marrow. *Proc.Natl.Acad. Sci.U.S.A* 1978; 75: 5884-8
185. Garner KJ, Lingrel JB: Structural organization of the β -globin locus of B-haplotype sheep. *Mol.Biol.Evol.* 1988; 5: 134-40
186. Zanjani ED, McGlave PB: Study of hemoglobin switching in sheep: is switching regulated by environmental influences? *Prog.Clin.Biol.Res.* 1983; 134: 523-31
187. Wheeler D, Hope R, Cooper SB, Dolman G, Webb GC, Bottema CD, Gooley AA, Goodman M, Holland RA: An orphaned mammalian β -globin gene of ancient evolutionary origin. *Proc.Natl.Acad.Sci.U.S.A* 2001; 98: 1101-6
188. Aguilera G, Bielawski JP, Yang Z: Evolutionary rate variation among vertebrate β globin genes: implications for dating gene family duplication events. *Gene* 2006; 380: 21-9
189. Murphy WS, Metcalfe J, Hoversland AS, Dhindsa DS: Postnatal changes in blood respiratory characteristics in an American opossum (*Didelphis virginiana*). *Respir. Physiol.* 1977; 29: 73-80

190. Koop BF, Goodman M: Evolutionary and developmental aspects of two hemoglobin β -chain genes (ϵ^M and β^M) of opossum. *Proc.Natl.Acad.Sci.U.S.A* 1988; 85: 3893-7
191. De Leo AA, Wheeler D, Lefevre C, Cheng JF, Hope R, Kuliwaba J, Nicholas KR, Westerman M, Graves JA: Sequencing and mapping hemoglobin gene clusters in the Australian model dasyurid marsupial *Sminthopsis macroura*. *Cytogenet.Genome Res.* 2005; 108: 333-41
192. Rae C, Sweeney KJE, Krockenberger AK, Agar NS, Gallagher CH, Kuchel PW: Comparison of the 1H and ^{31}P NMR Spectra of Erythrocytes and Plasma from some Australian Native Animals: Bandicoot, Echidna, Koala, Little Penguin, Tammar Wallaby, Tasmanian Devil, Tree Kangaroo and Wombat. *Comp.Haematol. Int.* 1993; 3: 71-80
193. Isaacks R, Nicol S, Sallis J, Zeidler R, Kim HD: Erythrocyte phosphates and hemoglobin function in monotremes and some marsupials. *Am.J.Physiol.*1984; 246: R236-R241
194. Lee MH, Shroff R, Cooper SJ, Hope R: Evolution and molecular characterization of a β -globin gene from the Australian Echidna *Tachyglossus aculeatus* (*Monotremata*). *Mol.Phylogenet.Evol.* 1999; 12: 205-14
195. Grigg GC, Wells RMG, Beard LA: Allosteric control of oxygen binding by h emoglobin during embryonic development in the crocodile *Crocodylus porosus*: the role of red cell organic phosphates and carbon dioxide. *J.Exp.Biol.* 19930; 175: 15-32
196. Isaacks RE, Harkness DR: 2,3-Diphosphoglycerate in erythrocytes of chick embryos. *Science* 1975; 189: 393-4
197. Chiang ST, Tzeng WF, King SC, Lee PY: Respiratory function of blood of the avians. *Chin J.Physiol.*1978; 22: 133-40
198. Landes GM, Villeponteau B, Pribyl TM, Martinson HG: Hemoglobin switching in chickens. Is the switch initiated post-transcriptionally? *J.Biol.Chem.* 1982; 257: 11008-14
199. Choi OR, Engel JD: Developmental regulation of β -globin gene switching. *Cell* 1988; 55: 17-26

200. Schnek AG, Paul C, Léonis J: Evolution and adaptation of avian and crocodilian hemoglobins, *Respiratory pigments in animals: relation structure-function*. Edited by Lamy J, Truchot J-P, Gilles R. Berlin, Springer-Verlag, 1985, pp 141-58
201. Hardison RC: New views of evolution and regulation of vertebrate β -like globin gene clusters from an orphaned gene in marsupials. *Proc.Natl.Acad.Sci.U.S.A* 2001; 98: 1327-9
202. Leder P, Hansen JN, Konkel D, Leder A, Nishioka Y, Talkington C: Mouse globin system: a functional and evolutionary analysis. *Science* 1980; 209: 1336-42
203. Hill A, Hardies SC, Phillips SJ, Davis MG, Hutchison CA, III, Edgell MH: Two mouse early embryonic β -globin gene sequences. Evolution of the nonadult β -globins. *J.Biol.Chem.* 1984; 259: 3739-47
204. Epner E, Reik A, Cimbara D, Telling A, Bender MA, Fiering S, Enver T, Martin DI, Kennedy M, Keller G, Groudine M: The β -globin LCR is not necessary for an open chromatin structure or developmentally regulated transcription of the native mouse β -globin locus. *Mol.Cell* 1998; 2: 447-55
205. Sawado T, Igarashi K, Groudine M: Activation of β -major globin gene transcription is associated with recruitment of NF-E2 to the β -globin LCR and gene promoter. *Proc.Natl.Acad.Sci.U.S.A* 2001; 98: 10226-31
206. Stamatoyannopoulos G, Raich N, Ebens A, Josephson B, Nakamoto B, Constantoulakis P, Costantini F, Papayannopoulou T, Enver T: Analysis of human γ -to- β switching in transgenic mice. *Ann.N.Y.Acad.Sci.* 1990; 612: 127-33
207. Peterson KR, Li QL, Clegg CH, Furukawa T, Navas PA, Norton EJ, Kimbrough TG, Stamatoyannopoulos G: Use of yeast artificial chromosomes (YACs) in studies of mammalian development: production of β -globin locus YAC mice carrying human globin developmental mutants. *Proc.Natl.Acad.Sci.U.S.A* 1995; 92: 5655-9
208. Goodman M, Koop BF, Czelusniak J, Weiss ML: The η -globin gene. Its long evolutionary history in the β -globin gene family of mammals. *J.Mol.Biol.* 1984; 180: 803-23
209. Phillips SJ, Hardies SC, Jahn CL, Edgell MH, Hutchison CA, III: The complete nucleotide sequence of a β -globin-like structure, β_{h2} , from the [Hbb]d mouse BALB/c. *J.Biol.Chem.* 1984; 259: 7947-54

210. Hutchison CA, III, Hardies SC, Padgett RW, Weaver S, Edgell MH: The mouse globin pseudogene β_3 is descended from a premammalian δ -globin gene. *J.Biol. Chem.* 1984; 259: 12881-9
211. Weaver S, Haigwood NL, Hutchison CA, III, Edgell MH: DNA fragments of the *Mus musculus* β globin haplotypes Hbb^s and Hbb^d. *Proc.Natl.Acad.Sci.U.S.A* 1979; 76: 1385-9
212. Russell ES, Blake SL, McFarland EC: Characterization and strain distribution of four alleles at the hemoglobin β -chain structural locus in the mouse. *Biochem. Genet.* 1972; 7: 313-30
213. Wawrzyniak CJ, Lewis SE, Popp RA: Differential expression of murine adult hemoglobins in early ontogeny. *Prog.Clin.Biol.Res.* 1985; 191: 421-31
214. Trimborn T, Gribnau J, Grosveld F, Fraser P: Mechanisms of developmental control of transcription in the murine α - and β -globin loci. *Genes Dev.* 1999; 13: 112-24
215. TomHon C, Zhu W, Millinoff D, Hayasaka K, Slightom JL, Goodman M, Gumucio DL: Evolution of a fetal expression pattern via *cis* changes near the γ globin gene. *J.Biol.Chem.* 1997; 272: 14062-6
216. Johnson RM, Gumucio D, Goodman M: Globin gene switching in primates. *Comp. Biochem.Physiol.A.Mol.Integr.Physiol.* 2002; 133: 877-83
217. Jeffreys AJ, Barrie PA, Harris S, Fawcett DH, Nugent ZJ, Boyd AC: Isolation and sequence analysis of a hybrid δ -globin pseudogene from the brown lemur. *J.Mol. Biol.* 1982; 156: 487-503
218. Barrie PA, Jeffreys AJ, Scott AF: Evolution of the β -globin gene cluster in man and the primates. *J.Mol.Biol.* 1981; 149: 319-36
219. Giebel LB, van S, V, Slightom JL, Spritz RA: Nucleotide sequence, evolution, and expression of the fetal globin gene of the spider monkey *Ateles geoffroyi*. *Proc.Natl. Acad.Sci.U.S.A* 1985; 82: 6985-9
220. Goodman M, Slightom JL, Gumucio DL: Molecular evolution in the β -globin gene family of mammals: emergence of redundant genes, important new genes, and new expression patterns, Gene families: structure, function, genetics and evolution. Edited by Holmes RS, Lim HA. New Jersey, World Scientific Pub. Co., 1996, pp 43-52

221. Boyer SH, Crosby EF, Fuller GL, Noyes AN, Adams JG: The structure and biosynthesis of hemoglobins A and A2 in the new world primate *Ateles paniscus*: a preliminary account. *Ann.N.Y.Acad.Sci.* 1969; 165: 360-77
222. Goodman M, Koop BF, Czelusniak J, Fitch DH, Tagle DA, Slightom JL: Molecular phylogeny of the family of apes and humans. *Genome* 1989; 31: 316-35
223. Goodman M, Tagle DA, Fitch DH, Bailey W, Czelusniak J, Koop BF, Benson P, Slightom JL: Primate evolution at the DNA level and a classification of hominoids. *J.Mol.Evol.* 1990; 30: 260-6
224. Raaum RL, Sterner KN, Noviello CM, Stewart CB, Disotell TR: Catarrhine primate divergence dates estimated from complete mitochondrial genomes: concordance with fossil and nuclear DNA evidence. *J.Hum.Evol.* 2005; 48: 237-57
225. Johnson RM, Buck S, Chiu CH, Gage DA, Shen TL, Hendrickx AG, Gumucio DL, Goodman M: Humans and old world monkeys have similar patterns of fetal globin expression. *J.Exp.Zool.* 2000; 288: 318-26
226. Breathnach CS: Foetal and neonatal hæmoglobins in sheep and goats. *Q.J.Exp. Physiol.Cogn.Med.Sci.* 1964; 49: 277-89
227. Barker JE, Pierce JE, Kefauver BC, Nienhuis AW: Hemoglobin switching in sheep and goats: induction of hemoglobin C synthesis in cultures of sheep fetal erythroid cells. *Proc.Natl.Acad.Sci.U.S.A* 1977; 74: 5078-82
228. Mueggler PA, Carpenter S, Black JA: Postnatal regulation of 2,3-DPG in sheep erythrocytes. *Am.J.Physiol.* 1983; 245: H506-H512
229. Breepoel PM, Kreuzer F, Hazevoet M: Interaction of organic phosphates with bovine hemoglobin. II. Oxygen binding equilibria of newborn and adult hemoglobin. *Pflugers Archiv.* 1981; 389: 227-35
230. Townes TM, Fitzgerald MC, Lingrel JB: Triplication of a four-gene set during evolution of the goat β -globin locus produced three genes now expressed differentially during development. *Proc.Natl.Acad.Sci.U.S.A* 1984; 81: 6589-93
231. Li QL, Zhou B, Powers P, Enver T, Stamatoyannopoulos G: β -globin locus activation regions: conservation of organization, structure, and function. *Proc. Natl.Acad.Sci.U.S.A* 1990; 87: 8207-11
232. Li Q, Zhou B, Powers P, Enver T, Stamatoyannopoulos G: Primary structure of the goat β -globin locus control region. *Genomics* 1991; 9: 488-99

233. Schon EA, Cleary ML, Haynes JR, Lingrel JB: Structure and evolution of goat γ -, β^C - and β^A -globin genes: three developmentally regulated genes contain inserted elements. *Cell* 1981; 27: 359-69
234. Shapiro SG, Moshirfar M: Structure of the goat $\psi\beta^y$ β -globin pseudogene. Analysis of goat pseudogene evolutionary patterns. *J.Mol.Biol.* 1989; 209: 181-9
235. Garner KJ, Lingrel JB: A comparison of the β^A -and β^B -globin gene clusters of sheep. *J.Mol.Evol.* 1989; 28: 175-84
236. Benz EJ, Jr., Kretschmer PJ, Geist CE, Kantor JA, Turner PH, Nienhuis AW: Hemoglobin switching in sheep. Synthesis, cloning, and characterization of DNA sequences coding for the β^B , β^C , and γ -globin mRNAs. *J.Biol.Chem.* 1979; 254: 6880-8
237. Kretschmer PJ, Coon HC, Davis A, Harrison M, Nienhuis AW: Hemoglobin switching in sheep. Isolation of the fetal γ -globin gene and demonstration that the fetal γ - and adult β^A -globin genes lie within eight kilobase segments of homologous DNA. *J.Biol.Chem.* 1981; 256: 1975-82
238. Agar NS, Harley JD: In vitro metabolism of red blood cells from newborn and adult goats. *Biol.Neonate* 1976; 28: 113-7
239. Perk K, Danon D: Comparative structure analysis of foetal and adult type red cells in newborn and adult cattle. *Res.Vet.Sci.* 1965; 6: 442-6
240. Lavrijzen K, Verwilghen RL: The synthesis of adult hemoglobins during hepatic erythropoiesis in the calf fetus. *Hemoglobin* 1983; 7: 159-79
241. Huang Y, Liu DP, Feng DX, Wu M, Shen W, Tang Y, Tang XB, Liang CC: No existence of translocus balancer to coordinate the expression and regulation of human hemoglobin genes in transgenic mice study. *Int J.Biochem.Cell Biol.* 2004; 36: 1261-5
242. Yu X, Kong Y, Dore LC, Abdulmalik O, Katein AM, Zhou S, Choi JK, Gell D, Mackay JP, Gow AJ, Weiss MJ: An erythroid chaperone that facilitates folding of α -globin subunits for hemoglobin synthesis. *J.Clin.Invest.* 2007; 117: 1856-65
243. Shen W, Liu DP, Liang CC: The regulatory network controlling β -globin gene switching. *Mol.Biol.Rep.* 2001; 28: 175-83

244. Chakalova L, Carter D, Debrand E, Goyenechea B, Horton A, Miles J, Osborne C, Fraser P: Developmental regulation of the β -globin gene locus. *Prog.Mol.Subcell. Biol.* 2005; 38: 183-206
245. Li XG, Liu DP, Liang CC: Beyond the locus control region: new light on β -globin locus regulation. *Int.J.Biochem.Cell Biol.* 2001; 33: 914-23
246. Harju S, McQueen KJ, Peterson KR: Chromatin structure and control of β -like globin gene switching. *Exp.Biol.Med.(Maywood.)* 2002; 227: 683-700
247. Peterson KR: Hemoglobin switching: new insights. *Curr.Opin.Hematol.* 2003; 10: 123-9
248. Stamatoyannopoulos G: Control of globin gene expression during development and erythroid differentiation. *Exp.Hematol.* 2005; 33: 259-71
249. Qiu C, Hanson E, Olivier E, Inada M, Kaufman DS, Gupta S, Bouhassira EE: Differentiation of human embryonic stem cells into hematopoietic cells by coculture with human fetal liver cells recapitulates the globin switch that occurs early in development. *Exp.Hematol.* 2005; 33: 1450-8
250. Raich N, Enver T, Nakamoto B, Josephson B, Papayannopoulou T, Stamatoyannopoulos G: Autonomous developmental control of human embryonic globin gene switching in transgenic mice. *Science* 1990; 250: 1147-9
251. Raich N, Papayannopoulou T, Stamatoyannopoulos G, Enver T: Demonstration of a human ϵ -globin gene silencer with studies in transgenic mice. *Blood* 1992; 79: 861-4
252. Zitnik G, Hines P, Stamatoyannopoulos G, Papayannopoulou T: Murine erythro-leukemia cell line GM979 contains factors that can activate silent chromosomal human γ -globin genes. *Proc.Natl.Acad.Sci.U.S.A* 1991; 88: 2530-4
253. Peterson KR, Zitnik G, Huxley C, Lowrey CH, Gnirke A, Leppig KA, Papayannopoulou T, Stamatoyannopoulos G: Use of yeast artificial chromosomes (YACs) for studying control of gene expression: correct regulation of the genes of a human β -globin locus YAC following transfer to mouse erythroleukemia cell lines. *Proc.Natl.Acad.Sci.U.S.A* 1993; 90: 11207-11
254. Ling J, Ainol L, Zhang L, Yu X, Pi W, Tuan D: HS2 enhancer function is blocked by a transcriptional terminator inserted between the enhancer and the promoter. *J.Biol.Chem.* 2004; 279: 51704-13

255. Dillon N, Grosveld F: Human γ -globin genes silenced independently of other genes in the β -globin locus. *Nature* 1991; 350: 252-4
256. Morley BJ, Abbott CA, Wood WG: Regulation of human fetal and adult globin genes in mouse erythroleukemia cells. *Blood* 1991; 78: 1355-63
257. Stamatoyannopoulos G, Josephson B, Zhang JW, Li Q: Developmental regulation of human γ -globin genes in transgenic mice. *Mol.Cell Biol.* 1993; 13: 7636-44
258. Harju S, Navas PA, Stamatoyannopoulos G, Peterson KR: Genome architecture of the human β -globin locus affects developmental regulation of gene expression. *Mol. Cell Biol.* 2005; 25: 8765-78
259. Yu M, Han H, Xiang P, Li Q, Stamatoyannopoulos G: Autonomous silencing as well as competition controls γ -globin gene expression during development. *Mol. Cell Biol.* 2006; 26: 4775-81
260. Lloyd JA, Krakowsky JM, Crable SC, Lingrel JB: Human γ - to β -globin gene switching using a mini construct in transgenic mice. *Mol.Cell Biol.* 1992; 12: 1561-7
261. Guy LG, Mei Q, Perkins AC, Orkin SH, Wall L: Erythroid Krüppel-like factor is essential for β -globin gene expression even in absence of gene competition, but is not sufficient to induce the switch from γ -globin to β -globin gene expression. *Blood* 1998; 91: 2259-63
262. Ryan TM, Sun CW, Ren J, Townes TM: Human γ -globin gene promoter element regulates human β -globin gene developmental specificity. *Nucleic Acids Res.* 2000; 28: 2736-40
263. Hanscombe O, Whyatt D, Fraser P, Yannoutsos N, Greaves D, Dillon N, Grosveld F: Importance of globin gene order for correct developmental expression. *Genes Dev.* 1991; 5: 1387-94
264. Peterson KR, Stamatoyannopoulos G: Role of gene order in developmental control of human γ - and β -globin gene expression. *Mol.Cell Biol.* 1993; 13: 4836-43
265. Ebb D, Tang DC, Drew L, Chin K, Berg PE, Rodgers GP: Identification of upstream regulatory elements that repress expression of adult β -like globin genes in a primitive erythroid environment. *Blood Cells Mol.Dis.* 1998; 24: 356-69
266. Ikuta T, Kan YW: In vivo protein-DNA interactions at the β -globin gene locus. *Proc.Natl.Acad.Sci.U.S.A* 1991; 88: 10188-92

267. Reddy PM, Shen CK: Protein-DNA interactions in vivo of an erythroid-specific, human β -globin locus enhancer. *Proc.Natl.Acad.Sci.U.S.A* 1991; 88: 8676-80
268. Duan Z, Stamatoyannopoulos G, Li Q: Role of NF-Y in in vivo regulation of the γ -globin gene. *Mol.Cell Biol.* 2001; 21: 3083-95
269. Horak CE, Mahajan MC, Luscombe NM, Gerstein M, Weissman SM, Snyder M: GATA-1 binding sites mapped in the β -globin locus by using mammalian ChIP-Chip analysis. *Proc.Natl.Acad.Sci.U.S.A* 2002; 99: 2924-9
270. Donze D, Townes TM, Bieker JJ: Role of erythroid Krüppel-like factor in human γ -to β -globin gene switching. *J.Biol.Chem.* 1995; 270: 1955-9
271. Filipe A, Li Q, Deveaux S, Godin I, Romeo PH, Stamatoyannopoulos G, Mignotte V: Regulation of embryonic/fetal globin genes by nuclear hormone receptors: a novel perspective on hemoglobin switching. *EMBO J.* 1999; 18: 687-97
272. Liberati C, Cera MR, Secco P, Santoro C, Mantovani R, Ottolenghi S, Ronchi A: Cooperation and competition between the binding of COUP-TFII and NF-Y on human ϵ - and γ -globin gene promoters. *J.Biol.Chem.* 2001; 276: 41700-9
273. Tanabe O, McPhee D, Kobayashi S, Shen Y, Brandt W, Jiang X, Campbell AD, Chen YT, Chang C, Yamamoto M, Tanimoto K, Engel JD: Embryonic and fetal β -globin gene repression by the orphan nuclear receptors, TR2 and TR4. *EMBO J.* 2007; 26: 2295-306
274. Gong QH, Stern J, Dean A: Transcriptional role of a conserved GATA-1 site in the human ϵ -globin gene promoter. *Mol.Cell Biol.* 1991; 11: 2558-66
275. Yu CY, Motamed K, Chen J, Bailey AD, Shen CK: The CACC box upstream of human embryonic ϵ globin gene binds Sp1 and is a functional promoter element in vitro and in vivo. *J.Biol.Chem.* 1991; 266: 8907-15
276. Gumucio DL, Shelton DA, Bailey WJ, Slightom JL, Goodman M: Phylogenetic footprinting reveals unexpected complexity in *trans* factor binding upstream from the ϵ -globin gene. *Proc.Natl.Acad.Sci.U.S.A* 1993; 90: 6018-22
277. Chae JH, Kim CG: CP2 binding to the promoter is essential for the enhanced transcription of globin genes in erythroid cells. *Mol.Cells* 2003; 15: 40-7
278. Asano H, Li XS, Stamatoyannopoulos G: FKLf, a novel Krüppel-like factor that activates human embryonic and fetal β -like globin genes. *Mol.Cell Biol.* 1999; 19: 3571-9

279. Ronchi AE, Bottardi S, Mazzucchelli C, Ottolenghi S, Santoro C: Differential binding of the NFE3 and CP1/NFY transcription factors to the human γ - and ϵ -globin CCAAT boxes. *J.Biol.Chem.* 1995; 270: 21934-41
280. Ronchi A, Bellorini M, Mongelli N, Mantovani R: CCAAT-box binding protein NF-Y (CBF, CP1) recognizes the minor groove and distorts DNA. *Nucleic Acids Res.* 1995; 23: 4565-72
281. Cao SX, Gutman PD, Dave HP, Schechter AN: Identification of a transcriptional silencer in the 5'-flanking region of the human ϵ -globin gene. *Proc.Natl.Acad.Sci. U.S.A* 1989; 86: 5306-9
282. Gutman PD, Cao SX, Dave HP, Mittelman M, Schechter AN: Binding of erythroid and non-erythroid nuclear proteins to the silencer of the human ϵ -globin-encoding gene. *Gene* 1992; 110: 197-203
283. Raich N, Clegg CH, Grofti J, Romeo PH, Stamatoyannopoulos G: GATA1 and YY1 are developmental repressors of the human ϵ -globin gene. *EMBO J.* 1995; 14: 801-9
284. Trepicchio WL, Dyer MA, Baron MH: Developmental regulation of the human embryonic β -like globin gene is mediated by synergistic interactions among multiple tissue- and stage-specific elements. *Mol.Cell Biol.* 1993; 13: 7457-68
285. Trepicchio WL, Dyer MA, Baron MH: A novel developmental regulatory motif required for stage-specific activation of the ϵ -globin gene and nuclear factor binding in embryonic erythroid cells. *Mol.Cell Biol.* 1994; 14: 3763-71
286. Yan, Z. J., Chen, Y. D., and Qian, R. L. Interactions between HMG proteins and 5'-flanking cis-acting elements of human ϵ -globin gene. *Sheng Wu Hua Xue.Yu Sheng Wu Wu Li Xue.Bao.(Shanghai)* 1998; 30: 471-476 (Résumé).
287. Dyer MA, Hayes PJ, Baron MH: The HMG domain protein SSRP1/PREIIBF is involved in activation of the human embryonic β -like globin gene. *Mol.Cell Biol.* 1998; 18: 2617-28
288. Yan, Z. J., Chen, Y. D., and Qian, R. L. Identification of an Embryonic Erythroid Nuclear Factor Binding to the Positive Stage-specific Regulatory Element of the Human ϵ -Globin Gene. *Sheng Wu Hua Xue.Yu Sheng Wu Wu Li Xue.Bao. (Shanghai)* 1997; 29: 287-293 (Résumé)

289. Wen J, Huang S, Rogers H, Dickinson LA, Kohwi-Shigematsu T, Noguchi CT: SATB1 family protein expressed during early erythroid differentiation modifies globin gene expression. *Blood* 2005; 105: 3330-9
290. Yan ZJ, Qian RL: The 5'-flanking *cis*-acting elements of the human ϵ -globin gene associates with the nuclear matrix and binds to the nuclear matrix proteins. *Cell Res.* 1998; 8: 209-18
291. Li Q, Blau CA, Clegg CH, Rohde A, Stamatoyannopoulos G: Multiple ϵ -promoter elements participate in the developmental control of ϵ -globin genes in transgenic mice. *J.Biol.Chem.* 1998; 273: 17361-7
292. Hou CH, Huang J, Qian RL: Identification of a NF- κ B site in the negative regulatory element (ϵ -NRAII) of human ϵ -globin gene and its binding protein NF- κ B p50 in the nuclei of K562 cells. *Cell Res.* 2002; 12: 79-82
293. Jane SM, Ney PA, Vanin EF, Gumucio DL, Nienhuis AW: Identification of a stage selector element in the human γ -globin gene promoter that fosters preferential interaction with the 5' HS2 enhancer when in competition with the β -promoter. *EMBO J.* 1992; 11: 2961-9
294. Ristaldi MS, Drabek D, Gribnau J, Poddie D, Yannoutsos N, Cao A, Grosveld F, Imam AM: The role of the -50 region of the human γ -globin gene in switching. *EMBO J.* 2001; 20: 5242-9
295. Zhou W, Clouston DR, Wang X, Cerruti L, Cunningham JM, Jane SM: Induction of human fetal globin gene expression by a novel erythroid factor, NF-E4. *Mol.Cell Biol.* 2000; 20: 7662-72
296. Zhao Q, Cumming H, Cerruti L, Cunningham JM, Jane SM: Site-specific acetylation of the fetal globin activator NF-E4 prevents its ubiquitination and regulates its interaction with the histone deacetylase, HDAC1. *J.Biol.Chem.* 2004; 279: 41477-86
297. Bottardi S, Santoro C: CTF/NF-1 binds the stage selector element of the human γ -globin gene promoter. *Biochem.Biophys.Res.Comm.* 1995; 215: 874-80
298. Jane SM, Gumucio DL, Ney PA, Cunningham JM, Nienhuis AW: Methylation-enhanced binding of Sp1 to the stage selector element of the human γ -globin gene promoter may regulate development specificity of expression. *Mol.Cell Biol.* 1993; 13: 3272-81

299. Sengupta PK, Lavelle D, DeSimone J: Increased binding of Sp1 to the γ -globin gene promoter upon site-specific cytosine methylation. *Am.J.Hematol.* 1994; 46: 169-72
300. Katsube T, Fukumaki Y: A role for the distal CCAAT box of the γ -globin gene in Hb switching. *J.Biochem.(Tokyo)* 1995; 117: 68-76
301. Omori A, Tanabe O, Engel JD, Fukamizu A, Tanimoto K: Adult stage γ -globin silencing is mediated by a promoter direct repeat element. *Mol.Cell Biol.* 2005; 25: 3443-51
302. Liberati C, Ronchi A, Lievens P, Ottolenghi S, Mantovani R: NF-Y organizes the γ -globin CCAAT boxes region. *J.Biol.Chem.* 1998; 273: 16880-9
303. Ronchi A, Berry M, Raguz S, Imam A, Yannoutsos N, Ottolenghi S, Grosveld F, Dillon N: Role of the duplicated CCAAT box region in γ -globin gene regulation and hereditary persistence of fetal haemoglobin. *EMBO J.* 1996; 15: 143-9
304. Asano H, Li XS, Stamatoyannopoulos G: FKLf-2: a novel Krüppel-like transcriptional factor that activates globin and other erythroid lineage genes. *Blood* 2000; 95: 3578-84
305. Zhang P, Basu P, Redmond LC, Morris PE, Rupon JW, Ginder GD, Lloyd JA: A functional screen for Krüppel-like factors that regulate the human γ -globin gene through the CACCC promoter element. *Blood Cells Mol.Dis.* 2005; 35: 227-35
306. Fischer KD, Haese A, Nowock J: Cooperation of GATA-1 and Sp1 can result in synergistic transcriptional activation or interference. *J.Biol.Chem.* 1993; 268: 23915-23
307. Lee JS, Ngo H, Kim D, Chung JH: Erythroid Krüppel-like factor is recruited to the CACCC box in the β -globin promoter but not to the CACCC box in the γ -globin promoter: the role of the neighboring promoter elements. *Proc.Natl.Acad.Sci.U.S.A* 2000; 97: 2468-73
308. Liu LR, Du ZW, Zhao HL, Liu XL, Huang XD, Shen J, Ju LM, Fang FD, Zhang JW: T to C substitution at -175 or -173 of the γ -globin promoter affects GATA-1 and Oct-1 binding in vitro differently but can independently reproduce the hereditary persistence of fetal hemoglobin phenotype in transgenic mice. *J.Biol. Chem.* 2005; 280: 7452-9
309. Lloyd JA, Case SS, Ponce E, Lingrel JB: Positive transcriptional regulation of the human γ -globin gene. γ PE is a novel nuclear factor with multiple binding sites near the gene. *J.Biol.Chem.* 1994; 269: 19385-93

310. Case SS, Huber P, Lloyd JA: The γ PE complex contains both SATB1 and HOXB2 and has positive and negative roles in human γ -globin gene regulation. *DNA Cell Biol.* 1999; 18: 805-17
311. Yant SR, Zhu W, Millinoff D, Slightom JL, Goodman M, Gumucio DL: High affinity YY1 binding motifs: identification of two core types (ACAT and CCAT) and distribution of potential binding sites within the human β globin cluster. *Nucleic Acids Res.* 1995; 23: 4353-62
312. Costa FC, Harju S, Fedosyuk H, Zelenchuk L, Neades R, Peterson KR: Silencing of γ -globin gene expression during adult definitive erythropoiesis is mediated by GATA-1 binding. Résumé présenté à: the 15th conference on hemoglobin switching. St. John's College, Oxford, UK, 2006
313. Gumucio DL, Heilstedt-Williamson H, Gray TA, Tarle SA, Shelton DA, Tagle DA, Slightom JL, Goodman M, Collins FS: Phylogenetic footprinting reveals a nuclear protein which binds to silencer sequences in the human γ and ϵ globin genes. *Mol. Cell Biol.* 1992; 12: 4919-29
314. Zhu W, TomHon C, Mason M, Campbell T, Sheldon E, Richards N, Goodman M, Gumucio DL: Analysis of linked human ϵ and γ transgenes: effect of locus control region hypersensitive sites 2 and 3 or a distal YY1 mutation on stage-specific expression patterns. *Blood* 1999; 93: 3540-9
315. Amrolia PJ, Cunningham JM, Ney P, Nienhuis AW, Jane SM: Identification of two novel regulatory elements within the 5'-untranslated region of the human γ -globin gene. *J.Biol.Chem.* 1995; 270: 12892-8
316. Foley HA, Ofori-Acquah SF, Yoshimura A, Critz S, Baliga BS, Pace BS: Stat3 β Inhibits γ globin gene expression in erythroid cells. *J.Biol.Chem.* 2002;
317. Bieker JJ, Southwood CM: The erythroid Krüppel-like factor transactivation domain is a critical component for cell-specific inducibility of a β -globin promoter. *Mol.Cell Biol.* 1995; 15: 852-60
318. Asano H, Stamatoyannopoulos G: Activation of β -globin promoter by erythroid Krüppel-like factor. *Mol.Cell Biol.* 1998; 18: 102-9
319. Marini MG, Asunis I, Porcu L, Salgo MG, Loi MG, Brucchiatti A, Cao A, Moi P: The distal β -globin CACCC box is required for maximal stimulation of the β -globin gene by EKLf. *Br.J.Haematol* 2004; 127: 114-7

320. deBoer E, Antoniou M, Mignotte V, Wall L, Grosveld F: The human β -globin promoter; nuclear protein factors and erythroid specific induction of transcription. *EMBO J.* 1988; 7: 4203-12
321. Berg PE, Williams DM, Qian RL, Cohen RB, Cao SX, Mittelman M, Schechter AN: A common protein binds to two silencers 5' to the human β -globin gene. *Nucleic Acids Res.* 1989; 17: 8833-52
322. Baysal E, Ribeiro ML, Huisman TH: Binding of nuclear factors to the proximal and distal CACC motifs of the β -globin gene promoter: implications for the -101 (C \rightarrow T) 'silent' β -thalassemia mutation. *Acta Haematol.* 1994; 91: 16-20
323. Lee JS, Lee CH, Chung JH: Studying the recruitment of Sp1 to the β -globin promoter with an in vivo method: protein position identification with nuclease tail (PIN*POINT). *Proc.Natl.Acad.Sci.U.S.A* 1998; 95: 969-74
324. Foley KP, Engel JD: Individual stage selector element mutations lead to reciprocal changes in β - vs. ϵ -globin gene transcription: genetic confirmation of promoter competition during globin gene switching. *Genes Dev.* 1992; 6: 730-44
325. Delvoeye NL, Destroismaisons NM, Wall LA: Activation of the β -globin promoter by the locus control region correlates with binding of a novel factor to the CAAT box in murine erythroleukemia cells but not in K562 cells. *Mol.Cell Biol.* 1993; 13: 6969-83
326. Wall L, Destroismaisons N, Delvoeye N, Guy LG: CAAT/enhancer-binding proteins are involved in β -globin gene expression and are differentially expressed in murine erythroleukemia and K562 cells. *J.Biol.Chem.* 1996; 271: 16477-84
327. Morgan WD, Williams GT, Morimoto RI, Greene J, Kingston RE, Tjian R: Two transcriptional activators, CCAAT-box-binding transcription factor and heat shock transcription factor, interact with a human hsp70 gene promoter. *Mol.Cell Biol.* 1987; 7: 1129-38
328. Gordon CT, Fox VJ, Najdovska S, Perkins AC: C/EBP δ and C/EBP γ bind the CCAAT-box in the human β -globin promoter and modulate the activity of the CACC-box binding protein, EKLF. *Biochim.Biophys.Acta* 2005; 1729: 74-80
329. Stuve LL, Myers RM: A directly repeated sequence in the β -globin promoter regulates transcription in murine erythroleukemia cells. *Mol.Cell Biol.* 1990; 10: 972-81

330. Stuve LL, Myers RM: Identification and characterization of a β -globin promoter-binding factor from murine erythroleukemia cells. *Mol.Cell Biol.* 1993; 13: 4311-22
331. Leach KM, Vieira KF, Kang SH, Aslanian A, Teichmann M, Roeder RG, Bungert J: Characterization of the human β -globin downstream promoter region. *Nucleic Acids Res.* 2003; 31: 1292-301
332. Crusselle-Davis VJ, Vieira KF, Zhou Z, Anantharaman A, Bungert J: Antagonistic regulation of β -globin gene expression by helix-loop-helix proteins USF and TFII-I. *Mol.Cell Biol.* 2006; 26: 6832-43
333. Pountney D, Trugnan G, Bourgeois M, Beaumont C: The identification of ferritin in the nucleus of K562 cells, and investigation of a possible role in the transcriptional regulation of adult β -globin gene expression. *J.Cell Sci.* 1999; 112 (Pt 6): 825-31
334. Broyles RH, Belegu V, DeWitt CR, Shah SN, Stewart CA, Pye QN, Floyd RA: Specific repression of β -globin promoter activity by nuclear ferritin. *Proc.Natl. Acad.Sci.U.S.A* 2001; 98: 9145-50
335. Berg PE, Mittelman M, Elion J, Labie D, Schechter AN: Increased protein binding to a -530 mutation of the human β -globin gene associated with decreased β -globin synthesis. *Am.J.Hematol.* 1991; 36: 42-7
336. Drew LR, Tang DC, Berg PE, Rodgers GP: The role of *trans*-acting factors and DNA-bending in the silencing of human β -globin gene expression. *Nucleic Acids Res.* 2000; 28: 2823-30
337. Chase MB, Fu S, Haga SB, Davenport G, Stevenson H, Do K, Morgan D, Mah AL, Berg PE: BP1, a homeodomain-containing isoform of DLX4, represses the β -globin gene. *Mol.Cell Biol.* 2002; 22: 2505-14
338. Fu S, Stevenson H, Strovel JW, Haga SB, Stamberg J, Do K, Berg PE: Distinct functions of two isoforms of a homeobox gene, BP1 and DLX7, in the regulation of the β -globin gene. *Gene* 2001; 278: 131-9
339. Forrester WC, Epner E, Driscoll MC, Enver T, Brice M, Papayannopoulou T, Groudine M: A deletion of the human β -globin locus activation region causes a major alteration in chromatin structure and replication across the entire β -globin locus. *Genes Dev.* 1990; 4: 1637-49
340. Li G, Lim KC, Engel JD, Bungert J: Individual LCR hypersensitive sites cooperate to generate an open chromatin domain spanning the human β -globin locus. *Genes Cells* 1998; 3: 415-29

341. Reik A, Telling A, Zitnik G, Cimborá D, Epner E, Groudine M: The locus control region is necessary for gene expression in the human β -globin locus but not the maintenance of an open chromatin structure in erythroid cells. *Mol.Cell Biol.* 1998; 18: 5992-6000
342. Cimborá DM, Schubeler D, Reik A, Hamilton J, Francastel C, Epner EM, Groudine M: Long-distance control of origin choice and replication timing in the human β -globin locus are independent of the locus control region. *Mol.Cell Biol.* 2000; 20: 5581-91
343. Tanimoto K, Liu Q, Bungert J, Engel JD: Effects of altered gene order or orientation of the locus control region on human β -globin gene expression in mice. *Nature* 1999; 398: 344-8
344. Molete JM, Petrykowska H, Bouhassira EE, Feng YQ, Miller W, Hardison RC: Sequences flanking hypersensitive sites of the β -globin locus control region are required for synergistic enhancement. *Mol.Cell Biol.* 2001; 21: 2969-80
345. Jackson DA, McDowell JC, Dean A: β -globin locus control region HS2 and HS3 interact structurally and functionally. *Nucleic Acids Res.* 2003; 31: 1180-90
346. Fraser P, Hurst J, Collis P, Grosveld F: DNaseI hypersensitive sites 1, 2 and 3 of the human β -globin dominant control region direct position-independent expression. *Nucleic Acids Res.* 1990; 18: 3503-8
347. Fraser P, Pruzina S, Antoniou M, Grosveld F: Each hypersensitive site of the human β -globin locus control region confers a different developmental pattern of expression on the globin genes. *Genes Dev.* 1993; 7: 106-13
348. Milot E, Strouboulis J, Trimborn T, Wijgerde M, de Boer E, Langeveld A, Tan-Un K, Vergeer W, Yannoutsos N, Grosveld F, Fraser P: Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. *Cell* 1996; 87: 105-14
349. Fedosyuk H, Peterson KR: Deletion of the human β -globin LCR 5'HS4 or 5'HS1 differentially affects β -like globin gene expression in β -YAC transgenic mice. *Blood Cells Mol.Dis.* 2007; 39: 44-55
350. Feng D, Kan YW: The binding of the ubiquitous transcription factor Sp1 at the locus control region represses the expression of β -like globin genes. *Proc.Natl. Acad.Sci.U.S.A* 2005; 102: 9896-900
351. Morley BJ, Abbott CA, Sharpe JA, Lida J, Chan-Thomas PS, Wood WG: A single β -globin locus control region element (5' hypersensitive site 2) is sufficient for

- developmental regulation of human globin genes in transgenic mice. *Mol.Cell Biol.* 1992; 12: 2057-66
352. Roberts NA, Sloane-Stanley JA, Sharpe JA, Stanworth SJ, Wood WG: Globin gene switching in transgenic mice carrying HS2-globin gene constructs. *Blood* 1997; 89: 713-23
 353. Jia CP, Huang SZ, Yan JB, Xiao YP, Ren ZR, Zeng YT: Effects of human locus control region elements HS2 and HS3 on human β -globin gene expression in transgenic mouse. *Blood Cells Mol.Dis.* 2003; 31: 360-9
 354. Talbot D, Grosveld F: The 5'HS2 of the globin locus control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NF-E2 binding sites. *EMBO J.* 1991; 10: 1391-8
 355. Kong S, Bohl D, Li C, Tuan D: Transcription of the HS2 enhancer toward a *cis*-linked gene is independent of the orientation, position, and distance of the enhancer relative to the gene. *Mol.Cell Biol.* 1997; 17: 3955-65
 356. Cavalleco R, Tuan D: Modulatory subdomains of the HS2 enhancer differentially regulate enhancer activity in erythroid cells at different developmental stages. *Blood Cells Mol.Dis.* 1997; 23: 8-26
 357. Elnitski L, Li J, Noguchi CT, Miller W, Hardison R: A negative *cis*-element regulates the level of enhancement by hypersensitive site 2 of the β -globin locus control region. *J.Biol.Chem.* 2001; 276: 6289-98
 358. Peterson KR, Clegg CH, Navas PA, Norton EJ, Kimbrough TG, Stamatoyannopoulos G: Effect of deletion of 5'HS3 or 5'HS2 of the human β -globin locus control region on the developmental regulation of globin gene expression in β -globin locus yeast artificial chromosome transgenic mice. *Proc.Natl.Acad.Sci. U.S.A* 1996; 93: 6605-9
 359. Fiering S, Epner E, Robinson K, Zhuang Y, Telling A, Hu M, Martin DI, Enver T, Ley TJ, Groudine M: Targeted deletion of 5'HS2 of the murine β -globin LCR reveals that it is not essential for proper regulation of the β -globin locus. *Genes Dev.* 1995; 9: 2203-13
 360. Hu X, Bulger M, Bender MA, Fields J, Groudine M, Fiering S: Deletion of the core region of 5' HS2 of the mouse β -globin locus control region reveals a distinct effect in comparison with human β -globin transgenes. *Blood* 2006; 107: 821-6

361. Zhang SB, Qian RL: The interaction between the human β -globin locus control region and nuclear matrix. *Cell Res.* 2002; 12: 411-6
362. Elnitski L, Miller W, Hardison R: Conserved E boxes function as part of the enhancer in hypersensitive site 2 of the β -globin locus control region. Role of basic helix-loop-helix proteins. *J.Biol.Chem.* 1997; 272: 369-78
363. Lam LT, Bresnick EH: Identity of the β -globin locus control region binding protein HS2NF5 as the mammalian homolog of the notch-regulated transcription factor suppressor of hairless. *J.Biol.Chem.* 1998; 273: 24223-31
364. Johnson KD, Norton JE, Bresnick EH: Requirements for utilization of CREB binding protein by hypersensitive site two of the β -globin locus control region. *Nucleic Acids Res.* 2002; 30: 1522-30
365. Zhou D, Pawlik KM, Ren J, Sun CW, Townes TM: Differential binding of erythroid Krüppel-like factor to embryonic/fetal globin gene promoters during development. *J.Biol.Chem.* 2006; 281: 16052-7
366. Navas PA, Josephson B, Furukawa T, Stamatoyannopoulos G, Li Q: The position of integration affects expression of the γ -globin-encoding gene linked to HS3 in transgenic mice. *Gene* 1995; 160: 165-71
367. Jackson JD, Petrykowska H, Philipsen S, Miller W, Hardison R: Role of DNA sequences outside the cores of DNase hypersensitive sites (HSs) in functions of the β -globin locus control region. Domain opening and synergism between HS2 and HS3. *J.Biol.Chem.* 1996; 271: 11871-8
368. Molete JM, Petrykowska H, Sigg M, Miller W, Hardison R: Functional and binding studies of HS3.2 of the β -globin locus control region. *Gene* 2002; 283: 185-97
369. Hug BA, Wesselschmidt RL, Fiering S, Bender MA, Epner E, Groudine M, Ley TJ: Analysis of mice containing a targeted deletion of β -globin locus control region 5' hypersensitive site 3. *Mol.Cell Biol.* 1996; 16: 2906-12
370. Bungert J, Dave U, Lim KC, Lieu KH, Shavit JA, Liu Q, Engel JD: Synergistic regulation of human β -globin gene switching by locus control region elements HS3 and HS4. *Genes Dev.* 1995; 9: 3083-96
371. Philipsen S, Pruzina S, Grosveld F: The minimal requirements for activity in transgenic mice of hypersensitive site 3 of the β globin locus control region. *EMBO J.* 1993; 12: 1077-85

372. Pruzina S, Antoniou M, Hurst J, Grosveld F, Philipsen S: Transcriptional activation by hypersensitive site three of the human β -globin locus control region in murine erythroleukemia cells. *Biochim.Biophys.Acta* 1994; 1219: 351-60
373. Shelton DA, Stegman L, Hardison R, Miller W, Bock JH, Slightom JL, Goodman M, Gumucio DL: Phylogenetic footprinting of hypersensitive site 3 of the β -globin locus control region. *Blood* 1997; 89: 3457-69
374. Tuan D, Abelovich A, Lee-Oldham M, Lee D: Identification of regulatory elements of human β -like globin genes, Developmental control of globin gene expression. Edited by Stamatoyannopoulos G, Nienhuis AW. New York, A.R. Liss., 1987, pp 211-20
375. Walters M, Kim C, Gelinas R: Characterization of a DNA binding activity in DNase I hypersensitive site 4 of the human globin locus control region. *Nucleic Acids Res.* 1991; 19: 5385-93
376. Lowrey CH, Bodine DM, Nienhuis AW: Mechanism of DNase I hypersensitive site formation within the human globin locus control region. *Proc.Natl.Acad.Sci.U.S.A* 1992; 89: 1143-7
377. Stamatoyannopoulos JA, Goodwin A, Joyce T, Lowrey CH: NF-E2 and GATA binding motifs are required for the formation of DNase I hypersensitive site 4 of the human β -globin locus control region. *EMBO J.* 1995; 14: 106-16
378. Tanimoto K, Sugiura A, Omori A, Felsenfeld G, Engel JD, Fukamizu A: Human β -globin locus control region HS5 contains CTCF- and developmental stage-dependent enhancer-blocking activity in erythroid cells. *Mol.Cell Biol.* 2003; 23: 8946-52
379. Wai AW, Gillemans N, Raguz-Bolognesi S, Pruzina S, Zafarana G, Meijer D, Philipsen S, Grosveld F: HS5 of the human β -globin locus control region: a developmental stage-specific border in erythroid cells. *EMBO J.* 2003; 22: 4489-500
380. Ling J, Pi W, Yu X, Bengra C, Long Q, Jin H, Seyfang A, Tuan D: The ERV-9 LTR enhancer is not blocked by the HS5 insulator and synthesizes through the HS5 site non-coding, long RNAs that regulate LTR enhancer function. *Nucleic Acids Res.* 2003; 31: 4582-96
381. Kim A, Kiefer CM, Dean A: Distinctive signatures of histone methylation in transcribed coding and noncoding human β -globin sequences. *Mol.Cell Biol.* 2007; 27: 1271-9

382. Shimotsuma M, Matsuzaki H, Tanabe O, Campbell AD, Engel JD, Fukamizu A, Tanimoto K: Linear distance from the locus control region determines ϵ -globin transcriptional activity. *Mol.Cell Biol.* 2007; 27: 5664-72
383. Fleenor DE, Kaufman RE: Characterization of the DNase I hypersensitive site 3' of the human β globin gene domain. *Blood* 1993; 81: 2781-90
384. O'Neill D, Bornschlegel K, Flamm M, Castle M, Bank A: A DNA-binding factor in adult hematopoietic cells interacts with a pyrimidine-rich domain upstream from the human δ -globin gene. *Proc.Natl.Acad.Sci.U.S.A* 1991; 88: 8953-7
385. O'Neill DW, Schoetz SS, Lopez RA, Castle M, Rabinowitz L, Shor E, Krawchuk D, Goll MG, Renz M, Seelig HP, Han S, Seong RH, Park SD, Agalioti T, Munshi N, Thanos D, Erdjument-Bromage H, Tempst P, Bank A: An ikaros-containing chromatin-remodeling complex in adult-type erythroid cells. *Mol.Cell Biol.* 2000; 20: 7572-82
386. O'Neill D, Yang J, Erdjument-Bromage H, Bornschlegel K, Tempst P, Bank A: Tissue-specific and developmental stage-specific DNA binding by a mammalian SWI/SNF complex associated with human fetal-to-adult globin gene switching. *Proc.Natl.Acad.Sci.U.S.A* 1999; 96: 349-54
387. Gaensler KM, Zhang Z, Lin C, Yang S, Hardt K, Flebbe-Rehwaldt L: Sequences in the γ - δ intergenic region are not required for stage-specific regulation of the human β -globin gene locus. *Proc.Natl.Acad.Sci.U.S.A* 2003; 100: 3374-9
388. Trudel M, Costantini F: A 3' enhancer contributes to the stage-specific expression of the human β -globin gene. *Genes Dev.* 1987; 1: 954-61
389. Grosveld F, Antoniou M, van Assendelft GB, de BE, Hurst J, Kollias G, MacFarlane F, Wrighton N: The regulation of expression of human β -globin genes. *Prog.Clin.Biol.Res.* 1987; 251: 133-44
390. Buzina A, Aladjem MI, Kolman JL, Wahl GM, Ellis J: Initiation of DNA replication at the human β -globin 3' enhancer. *Nucleic Acids Res.* 2005; 33: 4412-24
391. Bodine DM, Ley TJ: An enhancer element lies 3' to the human γ globin gene. *EMBO J.* 1987; 6: 2997-3004
392. Gribnau J, Diderich K, Pruzina S, Calzolari R, Fraser P: Intergenic transcription and developmental remodeling of chromatin subdomains in the human β -globin locus. *Mol.Cell* 2000; 5: 377-86

393. Ling J, Pi W, Bollag R, Zeng S, Keskinetepe M, Saliman H, Krantz S, Whitney B, Tuan D: The solitary long terminal repeats of ERV-9 endogenous retrovirus are conserved during primate evolution and possess enhancer activities in embryonic and hematopoietic cells. *J.Virol.* 2002; 76: 2410-23
394. Donovan-Peluso M, Acuto S, Swanson M, Dobkin C, Bank A: Expression of human γ -globin genes in human erythroleukemia (K562) cells. *J.Biol.Chem.* 1987; 262: 17051-7
395. Miller IJ, Bieker JJ: A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Krüppel family of nuclear proteins. *Mol.Cell Biol.* 1993; 13: 2776-86
396. Armstrong JA, Bieker JJ, Emerson BM: A SWI/SNF-related chromatin remodeling complex, E-RC1, is required for tissue-specific transcriptional regulation by EKLf in vitro. *Cell* 1998; 95: 93-104
397. Zhang W, Bieker JJ: Acetylation and modulation of erythroid Krüppel-like factor (EKLf) activity by interaction with histone acetyltransferases. *Proc.Natl.Acad. Sci.U.S.A* 1998; 95: 9855-60
398. Lee JS, Lee CH, Chung JH: The β -globin promoter is important for recruitment of erythroid Krüppel-like factor to the locus control region in erythroid cells. *Proc. Natl.Acad.Sci.U.S.A* 1999; 96: 10051-5
399. Drissen R, Palstra RJ, Gillemans N, Splinter E, Grosveld F, Philipsen S, de Laat W: The active spatial organization of the β -globin locus requires the transcription factor EKLf. *Genes Dev.* 2004; 18: 2485-90
400. Nuez B, Michalovich D, Bygrave A, Ploemacher R, Grosveld F: Defective haematopoiesis in fetal liver resulting from inactivation of the EKLf gene. *Nature* 1995; 375: 316-8
401. Chen X, Bieker JJ: Stage-specific repression by the EKLf transcriptional activator. *Mol.Cell Biol.* 2004; 24: 10416-24
402. Letting DL, Chen YY, Rakowski C, Reedy S, Blobel GA: Context-dependent regulation of GATA-1 by friend of GATA-1. *Proc.Natl.Acad.Sci.U.S.A* 2004; 101: 476-81
403. Welch JJ, Watts JA, Vakoc CR, Yao Y, Wang H, Hardison RC, Blobel GA, Chodosh LA, Weiss MJ: Global regulation of erythroid gene expression by transcription factor GATA-1. *Blood* 2004; 104: 3136-47

404. Walters M, Martin DI: Functional erythroid promoters created by interaction of the transcription factor GATA-1 with CACCC and AP-1/NFE-2 elements. *Proc.Natl. Acad.Sci.U.S.A* 1992; 89: 10444-8
405. Im H, Grass JA, Johnson KD, Kim SI, Boyer ME, Imbalzano AN, Bieker JJ, Bresnick EH: Chromatin domain activation via GATA-1 utilization of a small subset of dispersed GATA motifs within a broad chromosomal region. *Proc.Natl. Acad.Sci.U.S.A* 2005; 102: 17065-70
406. Vakoc CR, Letting DL, Gheldof N, Sawado T, Bender MA, Groudine M, Weiss MJ, Dekker J, Blobel GA: Proximity among distant regulatory elements at the β -globin locus requires GATA-1 and FOG-1. *Mol.Cell* 2005; 17: 453-62
407. Li Q, Clegg C, Peterson K, Shaw S, Raich N, Stamatoyannopoulos G: Binary transgenic mouse model for studying the trans control of globin gene switching: evidence that GATA-1 is an in vivo repressor of human ϵ gene expression. *Proc. Natl.Acad.Sci.U.S.A* 1997; 94: 2444-8
408. Kotkow KJ, Orkin SH: Dependence of globin gene expression in mouse erythroleukemia cells on the NF-E2 heterodimer. *Mol.Cell Biol.* 1995; 15: 4640-7
409. Forsberg EC, Downs KM, Bresnick EH: Direct interaction of NF-E2 with hypersensitive site 2 of the β -globin locus control region in living cells. *Blood* 2000; 96: 334-9
410. Goodwin AJ, McInerney JM, Glander MA, Pomerantz O, Lowrey CH: In vivo formation of a human β -globin locus control region core element requires binding sites for multiple factors including GATA-1, NF-E2, erythroid Krüppel-like factor, and Sp1. *J.Biol.Chem.* 2001; 276: 26883-92
411. Gong QH, McDowell JC, Dean A: Essential role of NF-E2 in remodeling of chromatin structure and transcriptional activation of the ϵ -globin gene in vivo by 5' hypersensitive site 2 of the β -globin locus control region. *Mol.Cell Biol.* 1996; 16: 6055-64
412. Kiekhäfer CM, Boyer ME, Johnson KD, Bresnick EH: A WW domain-binding motif within the activation domain of the hematopoietic transcription factor NF-E2 is essential for establishment of a tissue-specific histone modification pattern. *J. Biol.Chem.* 2004; 279: 7456-61
413. Kooren J, Palstra RJ, Klous P, Splinter E, von LM, Grosveld F, de LW: β -globin active chromatin Hub formation in differentiating erythroid cells and in p45 NF-E2 knock-out mice. *J.Biol.Chem.* 2007; 282: 16544-52

414. Shivdasani RA, Orkin SH: Erythropoiesis and globin gene expression in mice lacking the transcription factor NF-E2. *Proc.Natl.Acad.Sci.U.S.A* 1995; 92: 8690-4
415. Lopez RA, Schoetz S, DeAngelis K, O'Neill D, Bank A: Multiple hematopoietic defects and delayed globin switching in Ikaros null mice. *Proc.Natl.Acad.Sci.U.S.A* 2002; 99: 602-7
416. Key JR, Papathanasiou P, Goodnow CC, Perkins AC: The transcription factor, Ikaros, plays a key role in γ - β globin gene switching. Résumé présenté à: the 15th conference on hemoglobin switching. St. John's College, Oxford, UK, 2006.
417. Tanabe O, Katsuoka F, Campbell AD, Song W, Yamamoto M, Tanimoto K, Engel JD: An embryonic/fetal β -type globin gene repressor contains a nuclear receptor TR2/TR4 heterodimer. *EMBO J.* 2002; 21: 3434-42
418. Nemeth MJ, Lowrey CH: An Erythroid-Specific Chromatin Opening Element Increases β -Globin Gene Expression from Integrated Retroviral Gene Transfer Vectors. *Gene Ther.Mol.Biol.* 2004; 8: 475-86
419. Shyu YC, Wen SC, Lee TL, Chen X, Hsu CT, Chen H, Chen RL, Hwang JL, Shen CK: Chromatin-binding in vivo of the erythroid Krüppel-like factor, EKLF, in the murine globin loci. *Cell Res.* 2006; 16: 347-55
420. Layon ME, Ackley CJ, West RJ, Lowrey CH: Expression of GATA-1 in a non-hematopoietic cell line induces β -globin locus control region chromatin structure remodeling and an erythroid pattern of gene expression. *J.Mol.Biol.* 2007; 366: 737-44
421. Ashe HL, Monks J, Wijgerde M, Fraser P, Proudfoot NJ: Intergenic transcription and transinduction of the human β -globin locus. *Genes Dev.* 1997; 11: 2494-509
422. Haussecker D, Proudfoot NJ: Dicer-dependent turnover of intergenic transcripts from the human β -globin gene cluster. *Mol.Cell Biol.* 2005; 25: 9724-33
423. Kim A, Zhao H, Ifrim I, Dean A: β -globin intergenic transcription and histone acetylation dependent on an enhancer. *Mol.Cell Biol.* 2007; 27: 2980-6
424. Bender MA, Bulger M, Close J, Groudine M: β -globin gene switching and DNase I sensitivity of the endogenous β -globin locus in mice do not require the locus control region. *Mol.Cell* 2000; 5: 387-93

425. Bottardi S, Aumont A, Grosveld F, Milot E: Developmental stage-specific epigenetic control of human β -globin gene expression is potentiated in hematopoietic progenitor cells prior to their transcriptional activation. *Blood* 2003; 102: 3989-97
426. Felsenfeld G: Chromatin unfolds. *Cell* 1996; 86: 13-9
427. Forsberg EC, Downs KM, Christensen HM, Im H, Nuzzi PA, Bresnick EH: Developmentally dynamic histone acetylation pattern of a tissue-specific chromatin domain. *Proc.Natl.Acad.Sci.U.S.A* 2000; 97: 14494-9
428. Miles J, Mitchell JA, Chakalova L, Goyenechea B, Osborne CS, O'Neill L, Tanimoto K, Engel JD, Fraser P: Intergenic transcription, cell-cycle and the developmentally regulated epigenetic profile of the human β -globin locus. *PLoS ONE*. 2007; 2: e630
429. Bellard M, Kuo MT, Dretzen G, Chambon P: Differential nuclease sensitivity of the ovalbumin and β -globin chromatin regions in erythrocytes and oviduct cells of laying hen. *Nucleic Acids Res.* 1980; 8: 2737-50
430. Gazit B, Cedar H: Nuclease sensitivity of active chromatin. *Nucleic Acids Res.* 1980; 8: 5143-55
431. Groudine M, Eisenman R, Gelinis R, Weintraub H: Developmental aspects of chromatin structure and gene expression. *Prog.Clin.Biol.Res.* 1983; 134: 159-82
432. Nicolas RH, Wright CA, Cockerill PN, Wyke JA, Goodwin GH: The nuclease sensitivity of active genes. *Nucleic Acids Res.* 1983; 11: 753-72
433. Davies P, Thomas P, Manning DL: Correlations between prostate chromatin structure and transcriptional activity and acceptor site distribution. *Prostate* 1986; 8: 151-66
434. Gilbert N, Boyle S, Fiegler H, Woodfine K, Carter NP, Bickmore WA: Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. *Cell* 2004; 118: 555-66
435. Jenuwein T, Allis CD: Translating the histone code. *Science* 2001; 293: 1074-80
436. Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J, Kouzarides T: Active genes are tri-methylated at K4 of histone H3. *Nature* 2002; 419: 407-11

437. Schneider R, Bannister AJ, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T: Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. *Nat.Cell Biol.* 2004; 6: 73-7
438. Calestagne-Morelli A, Ausio J: Long-range histone acetylation: biological significance, structural implications, and mechanisms. *Biochem.Cell Biol.* 2006; 84: 518-27
439. Rougeulle C, Navarro P, Avner P: Promoter-restricted H3 Lys 4 di-methylation is an epigenetic mark for monoallelic expression. *Hum.Mol.Genet.* 2003; 12: 3343-8
440. Alvarez-Venegas R, Avramova Z: Methylation patterns of histone H3 Lys 4, Lys 9 and Lys 27 in transcriptionally active and inactive *Arabidopsis* genes and in *atx1* mutants. *Nucleic Acids Res.* 2005; 33: 5199-207
441. Litt MD, Simpson M, Gaszner M, Allis CD, Felsenfeld G: Correlation between histone lysine methylation and developmental changes at the chicken β -globin locus. *Science* 2001; 293: 2453-5
442. Kiekhäfer CM, Grass JA, Johnson KD, Boyer ME, Bresnick EH: Hematopoietic-specific activators establish an overlapping pattern of histone acetylation and methylation within a mammalian chromatin domain. *Proc.Natl.Acad.Sci.U.S.A* 2002; 99: 14309-14
443. Kim A, Dean A: Developmental stage differences in chromatin subdomains of the β -globin locus. *Proc.Natl.Acad.Sci.U.S.A* 2004; 101: 7028-33
444. Schubeler D, Groudine M, Bender MA: The murine β -globin locus control region regulates the rate of transcription but not the hyperacetylation of histones at the active genes. *Proc.Natl.Acad.Sci.U.S.A* 2001; 98: 11432-7
445. Gui CY, Dean A: A major role for the TATA box in recruitment of chromatin modifying complexes to a globin gene promoter. *Proc.Natl.Acad.Sci.U.S.A* 2003; 100: 7009-14
446. Keshet I, Lieman-Hurwitz J, Cedar H: DNA methylation affects the formation of active chromatin. *Cell* 1986; 44: 535-43
447. Rupon JW, Wang SZ, Gaensler K, Lloyd J, Ginder GD: Methyl binding domain protein 2 mediates γ -globin gene silencing in adult human β YAC transgenic mice. *Proc.Natl.Acad.Sci.U.S.A* 2006; 103: 6617-22

448. Enver T, Zhang JW, Papayannopoulou T, Stamatoyannopoulos G: DNA methylation: a secondary event in globin gene switching? *Genes Dev.* 1988; 2: 698-706
449. Yisraeli J, Frank D, Razin A, Cedar H: Effect of in vitro DNA methylation on β -globin gene expression. *Proc.Natl.Acad.Sci.U.S.A* 1988; 85: 4638-42
450. Singal R, Ferris R, Little JA, Wang SZ, Ginder GD: Methylation of the minimal promoter of an embryonic globin gene silences transcription in primary erythroid cells. *Proc.Natl.Acad.Sci.U.S.A* 1997; 94: 13724-9
451. Lavelle D, Vaitkus K, Hankewych M, Singh M, DeSimone J: Developmental changes in DNA methylation and covalent histone modifications of chromatin associated with the ϵ -, γ -, and β -globin gene promoters in *Papio anubis*. *Blood Cells Mol.Dis.* 2006; 36: 269-78
452. Goren A, Simchen G, Fibach E, Szabo PE, Tanimoto K, Chakalova L, Pfeifer GP, Fraser PJ, Engel JD, Cedar H: Fine tuning of globin gene expression by DNA methylation. *PLoS ONE.* 2006; 1: e46
453. Mabaera RH, Richardson CA, Johnson KJ, Hsu M, Fiering SH, Lowrey CH: Developmental and differentiation-specific patterns of human γ and β -globin promoter DNA methylation. *Blood* 2007; 110: 1343-52
454. Hsu M, Mabaera R, Lowrey CH, Martin DI, Fiering S: CpG Hypomethylation in a large domain encompassing the embryonic β -like globin genes in primitive erythrocytes. *Mol.Cell Biol.* 2007; 27: 5047-54
455. Carter D, Chakalova L, Osborne CS, Dai YF, Fraser P: Long-range chromatin regulatory interactions in vivo. *Nat.Genet.* 2002; 32: 623-6
456. Dekker J, Rippe K, Dekker M, Kleckner N: Capturing chromosome conformation. *Science* 2002; 295: 1306-11
457. Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, de WE, van SB, de LW: Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat.Genet.* 2006; 38: 1348-54
458. Dostie J, Richmond TA, Arnaout RA, Selzer RR, Lee WL, Honan TA, Rubio ED, Krumm A, Lamb J, Nusbaum C, Green RD, Dekker J: Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* 2006; 16: 1299-309

459. Mueller-Storm HP, Sogo JM, Schaffner W: An enhancer stimulates transcription in *trans* when attached to the promoter via a protein bridge. *Cell* 1989; 58: 767-77
460. Heuchel R, Matthias P, Schaffner W: Two closely spaced promoters are equally activated by a remote enhancer: evidence against a scanning model for enhancer action. *Nucleic Acids Res.* 1989; 17: 8931-47
461. Muller HP, Matthias P, Schaffner W: A transcriptional terminator between enhancer and promoter does not affect remote transcriptional control. *Somat.Cell.Mol.Genet.* 1990; 16: 351-60
462. Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W: Looping and interaction between hypersensitive sites in the active β -globin locus. *Mol.Cell* 2002; 10: 1453-65
463. Patrinos GP, de Krom M, de Boer E, Langeveld A, Imam AM, Strouboulis J, de Laat W, Grosveld FG: Multiple interactions between regulatory regions are required to stabilize an active chromatin hub. *Genes Dev.* 2004; 18: 1495-509
464. Wijgerde M, Grosveld F, Fraser P: Transcription complex stability and chromatin dynamics in vivo. *Nature* 1995; 377: 209-13
465. Tuan D, Kong S, Hu K: Transcription of the hypersensitive site HS2 enhancer in erythroid cells. *Proc.Natl.Acad.Sci.U.S.A* 1992; 89: 11219-23
466. Blackwood EM, Kadonaga JT: Going the distance: a current view of enhancer action. *Science* 1998; 281: 60-3
467. Lozzio CB, Lozzio BB: Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. *Blood* 1975; 45: 321-34
468. Martin P, Papayannopoulou T: HEL cells: a new human erythroleukemia cell line with spontaneous and induced globin expression. *Science* 1982; 216: 1233-5
469. Rutherford TR, Clegg JB, Weatherall DJ: K562 human leukaemic cells synthesise embryonic haemoglobin in response to haemin. *Nature* 1979; 280: 164-5
470. Benz EJ, Jr., Murnane MJ, Tonkonow BL, Berman BW, Mazur EM, Cavallero C, Jenko T, Snyder EL, Forget BG, Hoffman R: Embryonic-fetal erythroid characteristics of a human leukemic cell line. *Proc.Natl.Acad.Sci.U.S.A* 1980; 77: 3509-13
471. Groudine M, Kohwi-Shigematsu T, Gelinas R, Stamatoyannopoulos G, Papayannopoulou T: Human fetal to adult hemoglobin switching: changes in chromatin structure of the β -globin gene locus. *Proc.Natl.Acad.Sci.U.S.A* 1983; 80: 7551-5

472. Hsiung N, Roginski RS, Henthorn P, Smithies O, Kucherlapati R, Skoultchi AI: Introduction and expression of a fetal human globin gene in mouse fibroblasts. *Mol.Cell Biol.* 1982; 2: 401-11
473. Rutherford T, Nienhuis AW: Human globin gene promoter sequences are sufficient for specific expression of a hybrid gene transfected into tissue culture cells. *Mol. Cell Biol.* 1987; 7: 398-402
474. Cone RD, Weber-Benarous A, Baorto D, Mulligan RC: Regulated expression of a complete human β -globin gene encoded by a transmissible retrovirus vector. *Mol. Cell Biol.* 1987; 7: 887-97
475. Acuto S, Donovan-Peluso M, Giambona N, Bank A: The role of human globin gene promoters in the expression of hybrid genes in erythroid and non-erythroid cells. *Biochem.Biophys.Res.Comm.* 1987; 143: 1099-106
476. Alter BP, Goff SC: Hemoglobin switching: a cellular model. *Science* 1980; 207: 647-9
477. Papayannopoulou TH, Lindsley D, Kurachi S, Lewison K, Hemenway T, Melis M, Anagnou NP, Najfeld V: Adult and fetal human globin genes are expressed following chromosomal transfer into MEL cells. *Proc.Natl.Acad.Sci.U.S.A* 1985; 82: 780-4
478. Takegawa S, Brice M, Stamatoyannopoulos G, Papayannopoulou T: Only adult hemoglobin is produced in fetal nonerythroid x MEL cell hybrids. *Blood* 1986; 68: 1384-8
479. Papayannopoulou T, Stamatoyannopoulos G: Human hemoglobin switching: insights from studies of erythroid cultures. *J.Cell Physiol.Suppl.* 1982; 1: 145-9
480. Fibach E, Burke LP, Schechter AN, Noguchi CT, Rodgers GP: Hydroxyurea increases fetal hemoglobin in cultured erythroid cells derived from normal individuals and patients with sickle cell anemia or β -thalassemia. *Blood* 1993; 81: 1630-5
481. Itskovitz-Eldor J, Schuldiner M, Karsenti D, Eden A, Yanuka O, Amit M, Soreq H, Benvenisty N: Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers. *Mol.Med.* 2000; 6: 88-95
482. Bobbert M: Ethical questions concerning research on human embryos, embryonic stem cells and chimeras. *Biotechnol.J.* 2006; 1: 1352-69

483. Olivier EN, Qiu C, Velho M, Hirsch RE, Bouhassira EE: Large-scale production of embryonic red blood cells from human embryonic stem cells. *Exp.Hematol.* 2006; 34: 1635-42
484. Porcu S, Kitamura M, Witkowska E, Zhang Z, Mutero A, Lin C, Chang J, Gaensler KM: The human β globin locus introduced by YAC transfer exhibits a specific and reproducible pattern of developmental regulation in transgenic mice. *Blood* 1997; 90: 4602-9
485. Peterson KR, Navas PA, Li Q, Stamatoyannopoulos G: LCR-dependent gene expression in β -globin YAC transgenics: detailed structural studies validate functional analysis even in the presence of fragmented YACs. *Hum.Mol.Genet.* 1998; 7: 2079-88
486. Liu Q, Bungert J, Engel JD: Mutation of gene-proximal regulatory elements disrupts human ϵ -, γ -, and β -globin expression in yeast artificial chromosome transgenic mice. *Proc.Natl.Acad.Sci.U.S.A* 1997; 94: 169-74
487. Strouboulis J, Dillon N, Grosveld F: Developmental regulation of a complete 70-kb human β -globin locus in transgenic mice. *Genes Dev.* 1992; 6: 1857-64
488. Dillon N, Trimborn T, Strouboulis J, Fraser P, Grosveld F: The effect of distance on long-range chromatin interactions. *Mol.Cell* 1997; 1: 131-9
489. McMorrow T, van den WA, Wollenschlaeger A, van de CM, Monkhorst K, Trimborn T, Fraser P, van Lohuizen M, Jenuwein T, Djabali M, Philipsen S, Grosveld F, Milot E: Activation of the β globin locus by transcription factors and chromatin modifiers. *EMBO J.* 2000; 19: 4986-96
490. Kaufman RM, Pham CT, Ley TJ: Transgenic analysis of a 100-kb human β -globin cluster-containing DNA fragment propagated as a bacterial artificial chromosome. *Blood* 1999; 94: 3178-84
491. Huang Y, Liu DP, Wu L, Li TC, Wu M, Feng DX, Liang CC: Proper developmental control of human globin genes reproduced by transgenic mice containing a 160-kb BAC carrying the human β -globin locus. *Blood Cells Mol.Dis.* 2000; 26: 598-610
492. Li Q, Han H, Ye X, Stafford M, Barkess G, Stamatoyannopoulos G: Transcriptional potentials of the β -like globin genes at different developmental stages in transgenic mice and hemoglobin switching. *Blood Cells Mol.Dis.* 2004; 33: 318-25
493. Vadolas J, Wardan H, Bosmans M, Zaibak F, Jamsai D, Voullaire L, Williamson R, Ioannou PA: Transgene copy number-dependent rescue of murine β -globin

- knockout mice carrying a 183 kb human β -globin BAC genomic fragment. *Biochim.Biophys.Acta* 2005; 1728: 150-62
494. Vadolas J, Nefedov M, Wardan H, Mansooriderakshan S, Voullaire L, Jamsai D, Williamson R, Ioannou PA: Humanized β -thalassemia mouse model containing the common IVSI-110 splicing mutation. *J.Biol.Chem.* 2006; 281: 7399-405
 495. Feng DX, Liu DP, Huang Y, Wu L, Li TC, Wu M, Tang XB, Liang CC: The expression of human α -like globin genes in transgenic mice mediated by bacterial artificial chromosome. *Proc.Natl.Acad.Sci.U.S.A* 2001; 98: 15073-7
 496. Al-Hasani K, Vadolas J, Voullaire L, Williamson R, Ioannou PA: Complementation of alpha-thalassaemia in α -globin knockout mice with a 191 kb transgene containing the human α -globin locus. *Transgenic Res.* 2004; 13: 235-43
 497. Tang Y, Wang Z, Huang Y, Liu DP, Liu G, Shen W, Tang X, Feng D, Liang CC: Gene order in human α -globin locus is required for their temporal specific expressions. *Genes Cells* 2006; 11: 123-31
 498. Katsantoni EZ, de KM, San J, Imam AM, Grosveld F, Anagnou NP, Strouboulis J: An embryonic-specific repressor element located 3' to the γ -globin gene influences transcription of the human β -globin locus in transgenic mice. *Exp.Hematol.* 2004; 32: 224-33
 499. Wallace HA, Marques-Kranc F, Richardson M, Luna-Crespo F, Sharpe JA, Hughes J, Wood WG, Higgs DR, Smith AJ: Manipulating the mouse genome to engineer precise functional syntenic replacements with human sequence. *Cell* 2007; 128: 197-209
 500. Suzuki N, Nishii K, Okazaki T, Ikeno M: Human artificial chromosomes constructed using the bottom-up strategy are stably maintained in mitosis and efficiently transmissible to progeny mice. *J.Biol.Chem.* 2006; 281: 26615-23
 501. Lin SD, Cooper P, Fung J, Weier HU, Rubin EM: Genome scan identifies a locus affecting γ -globin level in human beta-cluster YAC transgenic mice. *Mamm. Genome* 2000; 11: 1024-9
 502. Jamsai D, Zaibak F, Khongnium W, Vadolas J, Voullaire L, Fowler KJ, Gazeas S, Fucharoen S, Williamson R, Ioannou PA: A humanized mouse model for a common β^0 -thalassemia mutation. *Genomics* 2005; 85: 453-61

503. Tanimoto K, Liu Q, Grosveld F, Bungert J, Engel JD: Context-dependent EKLF responsiveness defines the developmental specificity of the human ϵ -globin gene in erythroid cells of YAC transgenic mice. *Genes Dev.* 2000; 14: 2778-94
504. Navas PA, Li Q, Peterson KR, Stamatoyannopoulos G: Investigations of a human embryonic globin gene silencing element using YAC transgenic mice. *Exp.Biol. Med.(Maywood.)* 2006; 231: 328-34
505. Liu Q, Tanimoto K, Bungert J, Engel JD: The $^A\gamma$ -globin 3' element provides no unique function(s) for human β -globin locus gene regulation. *Proc.Natl.Acad.Sci. U.S.A* 1998; 95: 9944-9
506. Calzolari R, McMorrow T, Yannoutsos N, Langeveld A, Grosveld F: Deletion of a region that is a candidate for the difference between the deletion forms of hereditary persistence of fetal hemoglobin and deltabeta-thalassemia affects β - but not γ -globin gene expression. *EMBO J.* 1999; 18: 949-58
507. Bauchwitz R, Costantini F: Developmentally distinct effects on human ϵ -, γ - and δ -globin levels caused by the absence or altered position of the human β -globin gene in YAC transgenic mice. *Hum.Mol.Genet.* 2000; 9: 561-74
508. Navas PA, Peterson KR, Li Q, Skarpidi E, Rohde A, Shaw SE, Clegg CH, Asano H, Stamatoyannopoulos G: Developmental specificity of the interaction between the locus control region and embryonic or fetal globin genes in transgenic mice with an HS3 core deletion. *Mol.Cell Biol.* 1998; 18: 4188-96
509. Tanimoto K, Liu Q, Bungert J, Engel JD: The polyoma virus enhancer cannot substitute for DNase I core hypersensitive sites 2-4 in the human β -globin LCR. *Nucleic Acids Res.* 1999; 27: 3130-7
510. Perez-Stable C, Costantini F: Roles of fetal $^G\gamma$ -globin promoter elements and the adult β -globin 3' enhancer in the stage-specific expression of globin genes. *Mol.Cell Biol.* 1990; 10: 1116-25
511. Harvey MP, Crosbie J, Trent RJ: Human $^G\gamma$ and $^A\gamma$ globin gene constructs containing the 3' $^A\gamma$ enhancer show persistent fetal expression in transgenic mice. *Transgenic Res.* 1993; 2: 121-4
512. Sabatino DE, Cline AP, Gallagher PG, Garrett LJ, Stamatoyannopoulos G, Forget BG, Bodine DM: Substitution of the human β -spectrin promoter for the human $^A\gamma$ -globin promoter prevents silencing of a linked human β -globin gene in transgenic mice. *Mol.Cell Biol.* 1998; 18: 6634-40

513. Li Q, Duan ZJ, Stamatoyannopoulos G: Analysis of the mechanism of action of non-deletion hereditary persistence of fetal hemoglobin mutants in transgenic mice. *EMBO J.* 2001; 20: 157-64
514. Sargent TG, Lloyd JA: The human γ -globin TATA and CACCC elements have key, distinct roles in suppressing β -globin gene expression in embryonic/fetal development. *J.Biol.Chem.* 2001; 276: 41817-24
515. Enver T, Raich N, Ebens AJ, Papayannopoulou T, Costantini F, Stamatoyannopoulos G: Developmental regulation of human fetal-to-adult globin gene switching in transgenic mice. *Nature* 1990; 344: 309-13
516. Anderson KP, Lloyd JA, Ponce E, Crable SC, Neumann JC, Lingrel JB: Regulated expression of the human β globin gene in transgenic mice requires an upstream globin or nonglobin promoter. *Mol.Biol.Cell* 1993; 4: 1077-85
517. Mignotte V, Eleouet JF, Raich N, Romeo PH: *Cis*- and *trans*-acting elements involved in the regulation of the erythroid promoter of the human porphobilinogen deaminase gene. *Proc.Natl.Acad.Sci.U.S.A* 1989; 86: 6548-52
518. Aries A, Trentesaux C, Ottolenghi S, Jardillier JC, Jeannesson P, Doubeikovski A: Activation of erythroid-specific promoters during anthracycline-induced differentiation of K562 cells. *Blood* 1996; 87: 2885-90
519. Eleouet JF, Romeo PH: CCACC-binding or simian-virus-40-protein-1-binding proteins cooperate with human GATA-1 to direct erythroid-specific transcription and to mediate 5' hypersensitive site 2 sensitivity of a TATA-less promoter. *Eur. J.Biochem.* 1993; 212: 763-70
520. Mignotte V, Wall L, deBoer E, Grosveld F, Romeo PH: Two tissue-specific factors bind the erythroid promoter of the human porphobilinogen deaminase gene. *Nucleic Acids Res.* 1989; 17: 37-54
521. Frampton J, Walker M, Plumb M, Harrison PR: Synergy between the NF-E1 erythroid-specific transcription factor and the CACCC factor in the erythroid-specific promoter of the human porphobilinogen deaminase gene. *Mol.Cell Biol.* 1990; 10: 3838-42
522. Porcher C, Pitiot G, Plumb M, Lowe S, de VH, Grandchamp B: Characterization of hypersensitive sites, protein-binding motifs, and regulatory elements in both promoters of the mouse porphobilinogen deaminase gene. *J.Biol.Chem.* 1991; 266: 10562-9

523. Walters M, Andrews NC, Magis W, Martin DI: Erythroid AP-1/NF-E2 elements vary in their response to NF-E2. *Exp.Hematol.* 1996; 24: 445-52
524. Beaupain D, Eleouet JF, Romeo PH: Initiation of transcription of the erythroid promoter of the porphobilinogen deaminase gene is regulated by a *cis*-acting sequence around the cap site. *Nucleic Acids Res.* 1990; 18: 6509-15
525. Peterson KR: Transgenic mice carrying yeast artificial chromosomes. *Expert.Rev. Mol.Med.* 2003; 2003: 1-25
526. Drabkin DL, Austin JH: Spectrophotometric studies. I. Spectrophotometric constants for common hemoglobin derivatives in human, dog, and rabbit blood. *J. Biol.Chem.* 1932; 98: 719-33
527. Thivierge C, Kurbegovic A, Couillard M, Guillaume R, Cote O, Trudel M: Over-expression of PKD1 causes polycystic kidney disease. *Mol.Cell Biol.* 2006; 26: 1538-48
528. Enver T, Ebens AJ, Forrester WC, Stamatoyannopoulos G: The human β -globin locus activation region alters the developmental fate of a human fetal globin gene in transgenic mice. *Proc.Natl.Acad.Sci.U.S.A* 1989; 86: 7033-7
529. Sargent TG, Buller AM, Teachey DT, McCanna KS, Lloyd JA: The γ -globin promoter has a major role in competitive inhibition of β -globin gene expression in early erythroid development. *DNA Cell Biol.* 1999; 18: 293-303
530. Fitch DH, Bailey WJ, Tagle DA, Goodman M, Sieu L, Slightom JL: Duplication of the γ -globin gene mediated by L1 long interspersed repetitive elements in an early ancestor of simian primates. *Proc.Natl.Acad.Sci.U.S.A* 1991; 88: 7396-400
531. Huber MC, Kruger G, Bonifer C: Genomic position effects lead to an inefficient reorganization of nucleosomes in the 5'-regulatory region of the chicken lysozyme locus in transgenic mice. *Nucleic Acids Res.* 1996; 24: 1443-52
532. Stunkel W, Kober I, Seifart KH: A nucleosome positioned in the distal promoter region activates transcription of the human U6 gene. *Mol.Cell Biol.* 1997; 17: 4397-405
533. Gilman JG, Mishima N, Wen XJ, Stoming TA, Lobel J, Huisman TH: Distal CCAAT box deletion in the γ globin gene of two black adolescents with elevated fetal γ globin. *Nucleic Acids Res.* 1988; 16: 10635-42

534. Pistidda P, Frogheri L, Oggiano L, Guiso L, Manca L, Dore F, Masala B, Gilman JG, Longinotti M: Fetal hemoglobin expression in compound heterozygotes for -117 (G→A)^Δγ HPFH and β⁰ 39 nonsense thalassemia. *Am.J.Hematol.* 1995; 49: 267-70
535. Coleman MB, Adams JG, III, Steinberg MH, Winter WP: A four base pair deletion 5' to the ^Δγ^T gene is associated not only with decreased expression of the ^Δγ^T-globin gene, but also of the ^Gγ-globin gene in *cis*. *Am.J.Hematol.* 1994; 47: 307-11
536. Huisman TH: Levels of Hb A2 in heterozygotes and homozygotes for β-thalassemia mutations: influence of mutations in the CACCC and ATAAA motifs of the β-globin gene promoter. *Acta Haematol.* 1997; 98: 187-94
537. Donze D, Jeancake PH, Townes TM: Activation of δ-globin gene expression by erythroid Krüppel-like factor: a potential approach for gene therapy of sickle cell disease. *Blood* 1996; 88: 4051-7
538. Zucker RM, Decal DL, Whittington KB: Embryonic hemoglobins in mouse erythro-leukemic cells. *Res.Commun.Chem.Pathol.Pharmacol.* 1983; 42: 513-6
539. Hu X, Bulger M, Roach JN, Eszterhas SK, Olivier E, Bouhassira EE, Groudine MT, Fiering S: Promoters of the murine embryonic β-like globin genes ε^y and β_{H1} do not compete for interaction with the β-globin locus control region. *Proc.Natl.Acad.Sci. U.S.A* 2003; 100: 1111-5
540. Hu X, Eszterhas S, Pallazzi N, Bouhassira EE, Fields J, Tanabe O, Gerber SA, Bulger M, Engel JD, Groudine M, Fiering S: Transcriptional interference among the murine β-like globin genes. *Blood* 2007; 109: 2210-6
541. Skow LC, Burkhart BA, Johnson FM, Popp RA, Popp DM, Goldberg SZ, Anderson WF, Barnett LB, Lewis SE: A mouse model for β-thalassemia. *Cell* 1983; 34: 1043-52
542. Curcio MJ, Kantoff P, Schafer MP, Anderson WF, Safer B: Compensatory increase in levels of β minor globin in murine β-thalassemia is under translational control. *J.Biol.Chem.* 1986; 261: 16126-32
543. Lingrel JB, Townes TM, Shapiro SG, Wernke SM, Liberator PA, Menon AG: Structural organization of the α and β globin loci of the goat. *Prog.Clin.Biol.Res.* 1985; 191: 67-79
544. Chada K, Magram J, Costantini F: An embryonic pattern of expression of a human fetal globin gene in transgenic mice. *Nature* 1986; 319: 685-9

545. Claus R, Fliegauf M, Stock M, Duque J, Kolanczyk M, Lubbert M: Inhibitors of DNA methylation and histone deacetylation independently relieve AML1/ETO-mediated lysozyme repression. *J.Leukoc.Biol.* 2006; 80: 1462-72
546. Feng YQ, Desprat R, Fu H, Olivier E, Lin CM, Lobell A, Gowda SN, Aladjem MI, Bouhassira EE: DNA methylation supports intrinsic epigenetic memory in mammalian cells. *PLoS Genet.* 2006; 2: e65
547. Grass JA, Jing H, Kim SI, Martowicz ML, Pal S, Blobel GA, Bresnick EH: Distinct functions of dispersed GATA factor complexes at an endogenous gene locus. *Mol.Cell Biol.* 2006; 26: 7056-67
548. Chang LS, Shi Y, Shenk T: Adeno-associated virus P5 promoter contains an adenovirus E1A-inducible element and a binding site for the major late transcription factor. *J.Virol.* 1989; 63: 3479-88
549. Shi Y, Lee JS, Galvin KM: Everything you have ever wanted to know about Yin Yang 1..... *Biochim.Biophys.Acta* 1997; 1332: F49-F66
550. Gordon S, Akopyan G, Garban H, Bonavida B: Transcription factor YY1: structure, function, and therapeutic implications in cancer biology. *Oncogene* 2006; 25: 1125-42
551. Rincon-Arano H, Valadez-Graham V, Guerrero G, Escamilla-Del-Arenal M, Recillas-Targa F: YY1 and GATA-1 interaction modulate the chicken 3'-side α -globin enhancer activity. *J.Mol.Biol.* 2005; 349: 961-75
552. McDonagh KT, Lin HJ, Lowrey CH, Bodine DM, Nienhuis AW: The upstream region of the human γ -globin gene promoter. Identification and functional analysis of nuclear protein binding sites. *J.Biol.Chem.* 1991; 266: 11965-74
553. Gallagher PG, Sabatino DE, Romana M, Cline AP, Garrett LJ, Bodine DM, Forget BG: A human β -spectrin gene promoter directs high level expression in erythroid but not muscle or neural cells. *J.Biol.Chem.* 1999; 274: 6062-73
554. Beaudoin M: BP1, un gène homéotique distal-less et son rôle dans l'érythropoïèse murine définitive. Mémoire de maîtrise, Université de Montréal. 2003; pp 1-135
555. Bresnick EH, Felsenfeld G: Dual promoter activation by the human β -globin locus control region. *Proc.Natl.Acad.Sci.U.S.A.* 1994; 91: 1314-7

556. Yin W, Barkess G, Fang X, Xiang P, Cao H, Stamatoyannopoulos G, Li Q: Histone acetylation at the human β -globin locus changes with developmental age. *Blood*. 2007; 110: 4101-7
557. Weatherall DJ, Clegg JB. *The thalassæmia syndromes*, 3rd Edition. Oxford, Blackwell Publications, 1981; pp. 1-875
558. Farrell CM, Grinberg A, Huang SP, Chen D, Pichel JG, Westphal H, Felsenfeld G. A large upstream region is not necessary for gene expression or hypersensitive site formation at the mouse β -globin locus. *Proc.Natl.Acad.Sci.U.S.A.* 2000; 97: 14554-9
559. Bender MA, Reik A, Close J, Telling A, Epner E, Fiering S, Hardison R, Groudine M. Description and targeted deletion of 5' hypersensitive site 5 and 6 of the mouse beta-globin locus control region. *Blood*. 1998; 92: 4394-403

Annexe I
Méthodes supplémentaires

Listes d'amorces

Tableau X: Clônage du promoteur érythroïde de PBGD (PCR niché)

Nom	Séquence	Description
00-25	5'-GCTGAGTCAGGAGAATCGCTTGAACC-3'	PBGD humain - sens
00-26	5'-AACCTTCCTCAGTACCTTCTGCCCTC-3'	PBGD humain - antisens
99-15	5'-GTAGTGAGCCTAGGTACGCCA-3'	PBGD humain - sens (2 ^e tour)*
99-16	5'-GGTCCCTTAACTTCCTCAGT-3'	PBGD humain - antisens (2 ^e tour)

(*) Génère un site AvrII

Tableau XI: Mutagenèse dirigée du promoteur de γ (-117: A→G) (PCR)

Nom	Séquence	Description
04-02	5'-GACAAGGCAAACCTTGACCAATAGTC-3'	Promoteur γ (-104 à -80) - sens
04-03	5'-AAGGCTATTGGTCAAGGCAAGGCTG-3'	Promoteur γ (-105 à -129) - antisens*

(*) Génère la mutation -117(A→G)

Tableau XII: Ammorces pour les Séquençages

Nom	Séquence	Description
04-17	5'-CCTCAGCCTCAGAGTTTTCCACATG-3'	5' promoteur γ (-1829) - sens
01-01	5'-AAATACTTCCTACCCAGTTAATGGTCC-3'	5' promoteur γ (-1116) - sens
00-52	5'-AAGATTCCCTGAGAAGTGAACCTAGC-3'	5' promoteur γ (-936) - sens
04-18	5'TCATAAAGACCACAGTTTCAGCGCAG-3'	5' promoteur γ (-603) - antisens
04-01	5'-TTGGTCAAGTTTGCTTGTCAGGC-3'	Promoteur γ (-109) - antisens
04-02	5'-GACAAGGCAAACCTTGACCAATAGTC-3'	Promoteur γ (-104) - sens
00-54	5'-TCTCCTCCAGCATCTTCCACATTCAC-3'	Exon 1 γ (+132) - antisens
04-16	5'-ACAAGCTGCATGTGGATCCTGAG-3'	Exon 2 γ (+458) - sens
00-05	5'-GCTGGACACATATAAAATGC-3'	Intron 2 γ (+927) - sens

Tableau XIII: Criblage et Géotypage (PCR)

Nom	Séquence	Description
00-52	5'-AAGATTCCCTGAGAAGTGAACCTAGC-3'	5' promoteur γ - sens*
FR- φ 3	5'-AGAGGCTTGTGATAGTAGCCTTGTC-3'	Exon 1 γ - antisens
98-07	5'-TCCGTAATCTTGAGATCGGGCGT-3'	Bras droit du YAC (sens)*
98-08	5'-GGTGATGTCGGCGATATAGGGGCCAGCAAC-3'	Bras droit du YAC (antisens)
98-09	5'-GTGATAAATTAAGTCTTGCGCCTTAAACC-3'	Bras gauche du YAC (sens)*
98-10	5'-GCTACTTGGAGCCACTATCGACTACGCGAT-3'	Bras gauche du YAC (antisens)
06-38	5'-TTCTCAGTCCTGTGGGAGGTG-3'	Promoteur EKLF (souris) - sens**
06-39	5'-CCCAAACAAGGCTCAGAGGC-3'	Intron 1 EKLF (souris) - antisens
06-40	5'-TGGCTACCCGTGATATTGCTG-3'	Neo - sens
06-41	5'-GGGCCTTTGGGGACATCGAAGGTC-3'	Ikaros (souris)***
06-42	5'-CATAGGGCATGTCTGACAGGCACTTGT-3'	Ikaros (souris)
06-43	5'-CCAGCCTCTGAGCCCAGAAAGCGA-3'	Neo

(*) Criblage des YIp, *Pop-in* et *Pop-out* et géotypage des souris pour les mutants β -YAC, PBGD^A γ -YAC, Δ -842/-384^A γ -YAC, Δ -384/-202^A γ -YAC et Δ -204/-135^A γ -YAC.

(**) Géotypage des souris EKLF-nulles.

(***) Géotypage des souris Ikaros-nulles.

Tableau XIV: Expression des gènes de globine (RT-PCR en temps réel)

Nom	Séquence	Description
05-22	5'-GCCTTTGCTAAGCTGAGTGAG-3'	ϵ -globine (humain) - sens
05-23	5'-TTGCCAAAGTGAGTAGCCAGAA-3'	ϵ -globine (humain) - antisens
05-08	5'-AATAAGCTCCTAGTCCAGACGC-3'	γ -globine - sens*
05-09	5'-TCCCTCTAGTCTCTGCTTCTCT-3'	PBGD ^A γ -globine - sens
05-07	5'-TAGACAACCAGGAGCCTTCC-3'	^A γ - et ^G γ -globine - antisens
05-05	5'-GCAAGGTGAACGTGGATGAAGT-3'	δ - et β -globine - sens
05-06	5'-TAACAGCATCAGGAGTGGACAGA-3'	δ - et β -globine - antisens
06-24	5'-CTCAAGGGCACCTTTGCCACA-3'	β -globine (humain) - sens
06-25	5'-TGGGGTGAATTCTTTGCCAAAGTGAT-3'	β -globine (humain) - antisens
06-26	5'-CAACCTCAAGGGCACTTTTTCTCA-3'	δ -globine - sens
06-27	5'-GGTGAATTCCTTGCCAAAGTTGC-3'	δ -globine - antisens
05-01	5'-GCTACCAGGGCCTTTGAGATG-3'	S16 (souris) - sens
05-02	5'-AGGAGCGATTTGCTGGTGTGG-3'	S16 (souris) - antisens
05-20	5'-CAGCCGAATGACAAAGAAAAGTTCA-3'	Glycophorine A (souris) - sens
05-21	5'-CAAACATTGGAGGACTCTTCATTAGGA-3'	Glycophorine A (souris) - antisens

(*) Reconnaît ^A γ et ^G γ chez les souris β -YAC et seulement ^G γ chez les PBGD^A γ -YAC.

Tableau XV: ChIP (PCR semi-quantitatif)

Nom	Séquence	Description
05-34	5'-CCTGGCTAAACTCCACCCAT-3'	Promoteur γ -globine (-158) - sens*
05-35	5'-GCGTCTGGACTAGGAGCTTATTG-3'	5'UTR γ -globine (+53) - antisens*
05-28	5'-TTTGCTGAGATGAAACAGGCG-3'	5' promoteur $G\gamma$ (-729) - sens
05-29	5'-CCAGGATTTTTGATGGGACACG-3'	5' promoteur $G\gamma$ (-458) - antisens
05-30	5'-ACGGCTGACAAAAGAAGTCCT-3'	5' promoteur γ (-351) - sens *
05-31	5'-TTTGCATTGAGATAGTGTGGGGAA-3'	Promoteur γ (-175) - antisens *
05-26	5'-AATCTATCCTCCAGTGAAGTCAAGCA-3'	Promoteur PBGD (-176) - sens
05-27	5'-TCCAGAGAAGCAGAGACTAGAGG-3'	5'UTR PBGD (+75) - antisens
05-24	5'-AGAAGAGCCAAGGACAGGTAC-3'	Promoteur β -globine (-146) - sens
05-25	5'-AGGTTGCTAGTGAACACAGTTGT-3'	5'UTR β -globine (+36) - antisens
05-32	5'-TACTTAGAAGGAATGGGAGCAGCT-3'	Promoteur Zfp37 - sens
05-33	5'-ACTTGAGAAAGAAGGGATTGCAGT-3'	Promoteur Zfp37 - antisens

(*) Reconnaît $A\gamma$ et $G\gamma$ chez les souris β -YAC et seulement $G\gamma$ chez les PBGD $A\gamma$ -YAC

Tableau XVI: ChIP (PCR en temps réel)

Nom	Séquence	T°C	Description
05-34	5'-CCTGGCTAAACTCCACCCAT-3'	64°C	Promoteur γ -globine (-158) - sens*
05-35	5'-GCGTCTGGACTAGGAGCTTATTG-3'		5'UTR γ -globine (+53) - antisens
05-37	5'-CCAGTGAAGTCAAGTCAAGTTC-3'	59°C	Promoteur PBGD (-169) - sens
05-38	5'-CGATAGTAGGACCTGCTGTTGC		5'UTR PBGD (+45) - antisens
05-24	5'-AGAAGAGCCAAGGACAGGTAC-3'	62°C	Promoteur β -globine (-146) - sens
05-25	5'-AGGTTGCTAGTGAACACAGTTGT-3'		5'UTR β -globine (+36) - antisens
05-32	5'-TACTTAGAAGGAATGGGAGCAGCT-3'	59°C	Promoteur Zfp37 - sens
05-33	5'-ACTTGAGAAAGAAGGGATTGCAGT-3'		Promoteur Zfp37 - antisens

T°C = température d'hybridation utilisée pour chacune de ces paires d'amorces.

(*) Reconnaît $A\gamma$ et $G\gamma$ chez les souris β -YAC et seulement $G\gamma$ chez les PBGD $A\gamma$ -YAC

Tableau XVII: Capture de conformation chromosomique (amplification par PCR)

Nom	Séquence	Description*
06-01	5'-GGGCCTAGAAATTATGTAGATGGTCCTG-3'	LCR humain, HS2-4 (#10305 - sens)
06-02	5'-AAGTGTTTCATGGGCTGAGGGATGGA-3'	ϵ -globine (#21132 - sens)
06-03	5'-GCTGCTTTGGCTCGCTTTTGGTGT-3'	Région entre ϵ et $G\gamma$ (#28536 - sens)
06-04	5'-TGGCCCTATCTTCAGCCTTCTTAAACA-3'	$G\gamma$ -globine (#29046 - antisens)
06-05	5'-CAATCAACCTGATAGCTTAGGGGATAAAC-3'	Région entre $G\gamma$ et $A\gamma$ (#38122 - sens)
06-06	5'-CTCTGAAAGTGATCCATGATCTCTAACCT-3'	$A\gamma$ -globine (#38246 - antisens)**
06-07	5'-CCAGAGCCCAAGTCAAGAACTACCG-3'	Promoteur PBGD (-617 - antisens)
06-09	5'-CAGAGAACTGAGGGAAGGATGAG-3'	$\psi\beta$ -globine (#50560 - sens)
06-10	5'-GGCTCAGTTTCTCAGAAGCCAGTCT-3'	δ -globine (#56056 - sens)
06-20	5'-ATCAGGAAACAGTCCAGGATCTCAATGG-3'	Région entre δ et β (#56280 - antisens)
06-15	5'-TTCATCTCTTGACCTCCTCATCTTCAATA-3'	β -globine (#58087 - antisens)
06-16	5'-GAAGACCATTTTCTCAAAGGCAATGGA-3'	3'HS (sens)
06-17	5'-CTTTACCAACCGTCTAGCCAGGTTTG-3'	Pkd1 (souris, exon 15 - sens)

(*) Les nombre marqués d'un (#) donnent la position du 1^{er} nucléotide de l'amorce dans le locus de la β -globine (accession U01317 - GeneBank) alors que la mention sens/antisens réfère à l'orientation de l'amorce par rapport à sa position au LCR (sens: s'éloigne du LCR, antisens: pointe vers le LCR). Cette description n'est valable que pour une digestion EcoRI.

(**) Emplifie le gène $A\gamma$ chez les β -YAC et la région en 5' du promoteur $A\gamma$ chez les PBGD $A\gamma$ -YAC.

Annexe II:
Résultats supplémentaires

Mutagenèse dirigée ^A γ : -117(A→G)

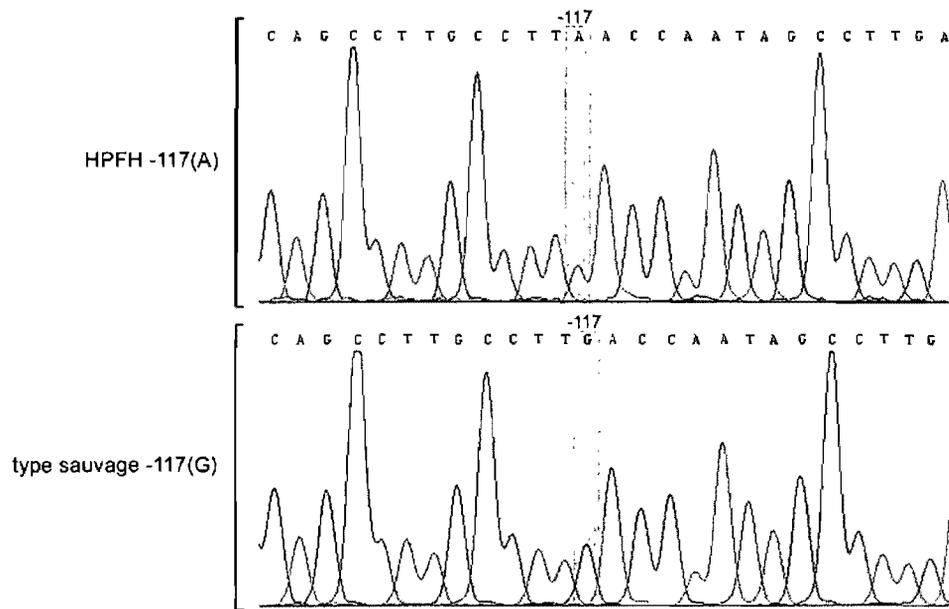


Figure 46: Séquençage des constructions mutées à la position -117(A→G). Chromatogrammes représentatifs de la séquence des constructions pRSΔ avant la mutagenèse dirigée [HPFH -117(A)] et après la mutagenèse [type sauvage -117(G)].

Caractérisation du PBGD^A γ -YAC par *PFGE*

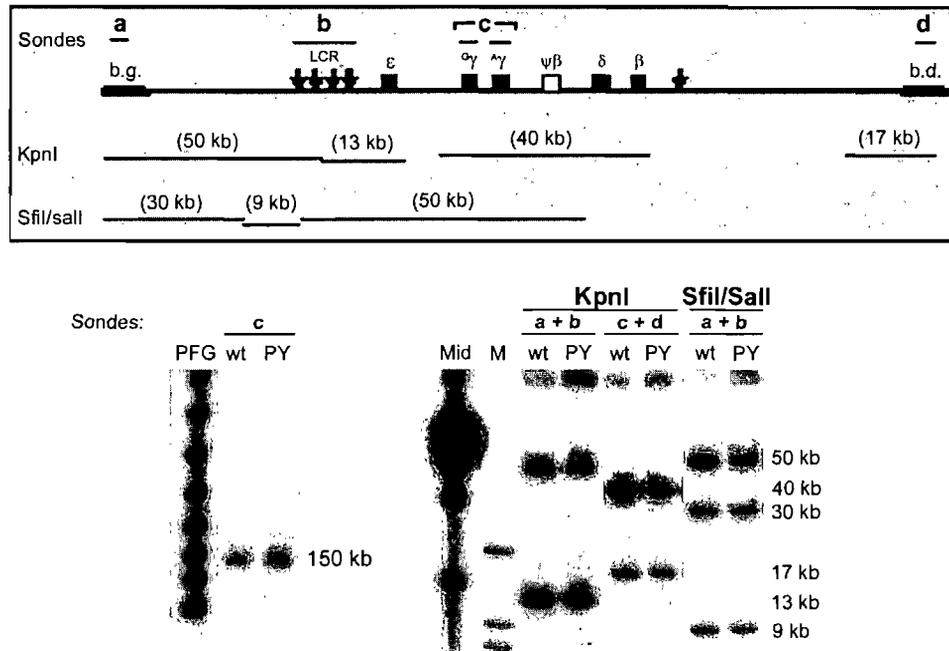


Figure 47: Caractérisation du mutant PBGD^A γ -YAC par *PFGE*. L'intégrité des clones *Po5-8* et *Po5-12* a été vérifiée par digestion enzymatique en utilisant des enzymes qui coupent rarement dans le locus. Les fragments ont été séparés par électrophorèse en champ pulsé puis transférés sur membrane pour être révélés par buvardage de *Southern* en utilisant quatre sondes réparties sur le YAC (a: bras gauche (b.g.) du YAC, b: LCR, c: γ -globine, d: bras droit (b.d.) du YAC). Le YAC entier a également été analysé de cette façon (à gauche). Ici sont présentés les autoradiogrammes du clone PBGD^A γ -YAC *Po5-12* (PY), en comparaison avec le locus sauvage (wt). PFG, marqueur de poids moléculaire pour électrophorèse en champs pulsé λ concaténé. Mid, marqueur de poids moléculaire *Mid-range*. M, marqueur de poids moléculaire λ HindIII

Caractérisation des fondatrice PBGD^A γ -YAC

Tableau XVIII: Caractérisation des fondatrices PBGD^A γ -YAC.

lignées	copies	Intégrité							Expression		
		B.G.	HS2	ϵ	γ	δ	β	B.D.	G γ	A γ	β
PY150	1		☑	☒	☑	☑	☑	x			
PY171	0,1			☑	☑	☑	☒				
PY175	<0,1			☒	☑	☒	☒				
PY186	0,5										
PY228	1	x	x	☑	☑	☑	☑	☐	-	+	+
PY367	2	☐	☑	☑	☑	☑	☑	☐	+	+	+
PY383	2	☐	☑	☑	☑	☑	☑	☐	+	+	+
PY388	<0,5	x	x	x	☐	x	x	x	-	-	-
PY397	1	☐	☑	☑	☑	☑	☑	☐	+	+	+
PY418	2	☐	☑	☑	☑	☑	☑	☐	+	+	+
PY422	1	☐	☑	☑	☑	x	x	x	-	±	
PY432	1	x	☑	☑	☑	☑	☑	☐	-	+	
PY435	0,5	☐	☑	☑	☑	x	x	x	-	±	
PY437	1	☐	☑	☑	☑	☑	☑	☐	+	+	+
PY452	5-10	☐	☑	☑	☑	☑	☑	☐	+	+	
PY457	0,5	x	x	☑	☑	☑	☑	☐	-	±	±
PY465	1	☐	☑	☑	☑	☑	☑	☐	+	+	+

Légende du tableau: Lignées, numéro de la fondatrice. L'intégrité des différentes régions du locus a été vérifiée par PCR pour le bras gauche (B.G.) et le bras droit (B.D.) du YAC et par buvardage de type *Southern* pour les autres régions: ☐: présent, ☑: présent et intègre, ☒: présent mais non-intègre (bandes de mauvaises longueurs), x: absent. Expression: expression protéique des chaînes de globine au stade adulte, +: expression aisément détectable, ±: expression faible, -: expression non-détectable.

Quantification postnatale des chaînes de globine

Les souris PBGD^{A γ} -YAC/Ikaros-nulles

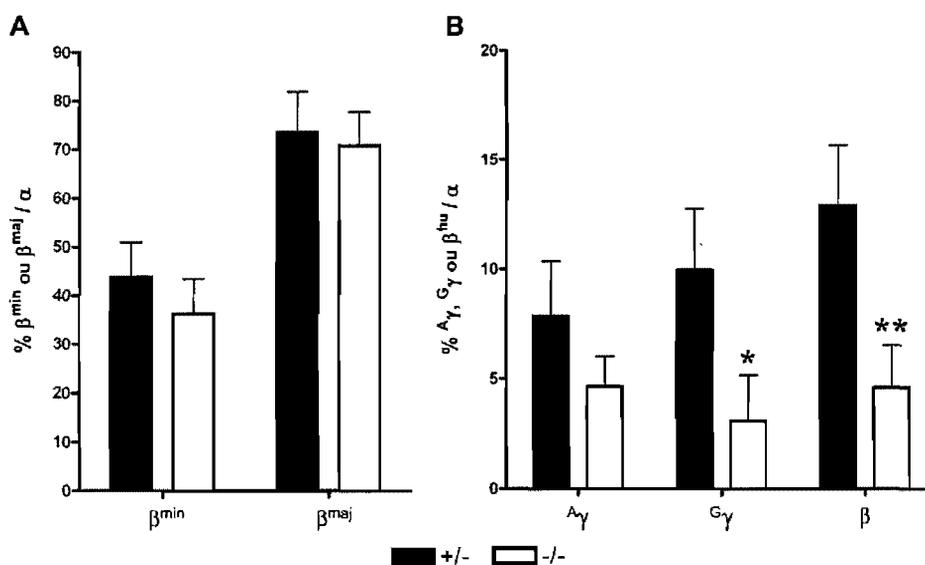


Figure 48: Expression des chaînes de globine chez les souris PBGD^{A γ} -YAC/Ikaros-nulles. Les chaînes de globine du sang périphérique ont été quantifiées par *UT-PAGE* chez des souris PBGD^{A γ} -YAC⁺ et Ikaros^{+/-} (+/-, $N = 11$) ou Ikaros^{-/-} (-/-, $N = 3$) de trois semaines. L'expression des chaînes murines β^{maj} et β^{min} (A) ainsi que l'expression des chaînes humaines A_{γ} , G_{γ} et β (B) ont été mesurées chez les deux lignées PBGD^{A γ} -YAC (PY367 et PY465) et sont rapportées en fonction du pourcentage de la chaîne α de la souris. Tous les histogrammes donnent la moyenne \pm l'écart type et le p a été déterminé à l'aide du t -test de Student: *, $p < 0,002$; **, $p < 0,0005$.

Annexe III
Résultats ajoutés en preuve

Expression chez les souris PBGD^A γ -YAC/ β BP1

Cette expérience a été effectuée par une étudiante à la maîtrise dans le laboratoire, Marthe-Sandrine Eiyemo Mwa Mpollo, qui s'intéressait à l'étude d'un répresseur potentiel du gène de la β -globine, le facteur transcriptionnel homéotique *Beta protein 1* (BP1) dans des souris transgéniques. Ces souris β BP1 de la lignée 43 expriment au stade adulte, le gène humain BP1 sous le contrôle du promoteur de la β -globine excisé de ses sites de liaison à BP1 de façon à éviter une autorégulation négative du transgène.⁵⁵⁴ Les souris β BP1-43 ont été croisées avec les souris PBGD^A γ -YAC hémizygotés pour les gènes endogènes β^{maj} et β^{min} , dans le but de forcer la répression du gène β après la naissance, et ainsi de pouvoir évaluer l'importance du promoteur β dans la régulation de la commutation tardive de l'hémoglobine fœtale à adulte observée dans le contexte du locus PBGD^A γ -YAC.

Les niveaux d'expressions des gènes γ^A , γ^G et β ont été mesurés au niveau de l'ARNm splénique par RT-PCR quantitatif en temps réel chez des nouveau-nés également hémizygotés pour les gènes endogènes de β -globine. Comme on peut le voir à la Figure 49, aucune différence significative n'a été mesurée pour chacun des trois gènes humains chez les nouveau-nés PBGD^A γ -YAC/ β BP1-43 lorsque comparés aux PBGD^A γ -YAC issus des mêmes portés ($p > 0,4$).

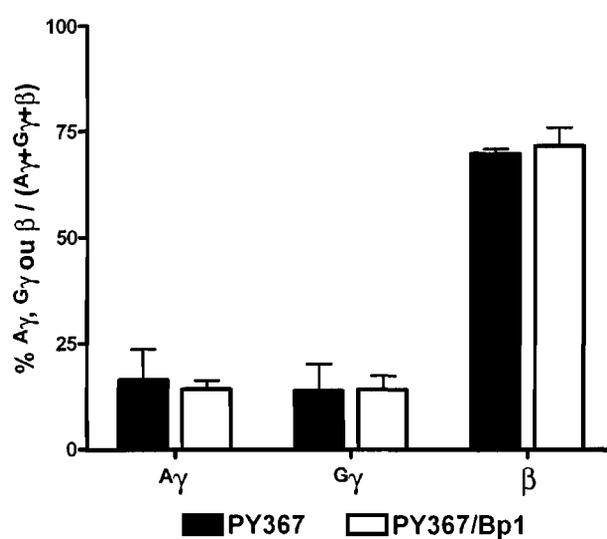


Figure 49: Expression des gènes de globine chez les fœtus PBGD^{A γ} -YAC/ β BP1-43. L'ARNm isolé de la rate a été quantifié par RT-PCR quantitatif en temps-réel pour l'expression des gènes G γ , A γ et β chez les nouveau-nés PBGD^{A γ} -YAC (YP367) et PBGD^{A γ} -YAC/ β BP1-43 (YP367/Bp1). Les histogrammes représentent la moyenne de la proportion des gènes de globine ($N = 4$ pour les deux lignées) \pm l'écart type.

Annexe IV
Cartes des plasmides

pRS406

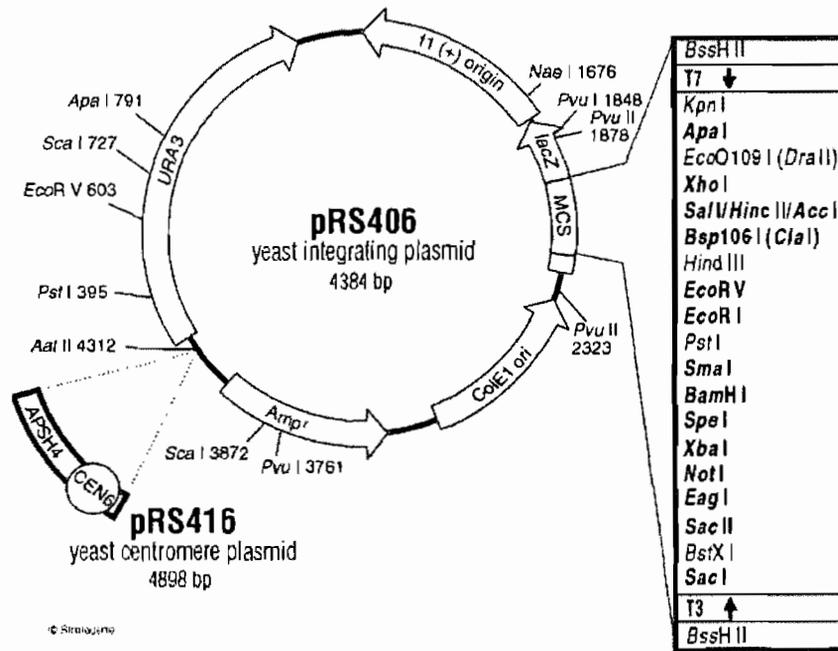


Figure 50: Plasmide d'intégration dans la levure pRS406 de Stratagene.

pBluescript

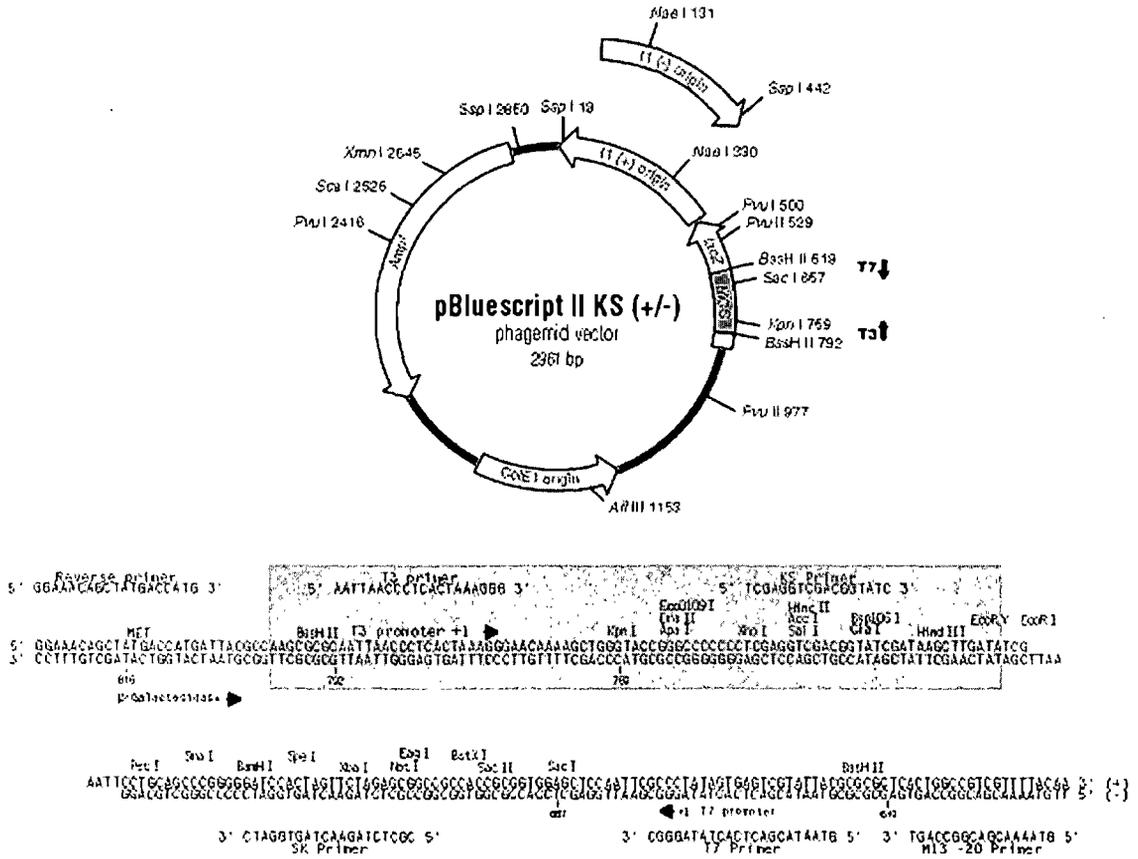


Figure 51: Plasmide pBluescript II. La zone ombragée du site de polyclonage indique la séquence éliminée pour le clonage du gène γ ayant servi à la génération des trois mutants de délétion.