



**HAL**  
open science

# Elaboration d'un outil de modification génétique au locus Dct dans les cellules souches embryonnaires de souris à l'aide de la méganucléase I-SceI

Myriam Fenina

► **To cite this version:**

Myriam Fenina. Elaboration d'un outil de modification génétique au locus Dct dans les cellules souches embryonnaires de souris à l'aide de la méganucléase I-SceI. Génétique. Université Pierre et Marie Curie - Paris VI, 2011. Français. NNT : 2011PA066491 . tel-00824944

**HAL Id: tel-00824944**

**<https://theses.hal.science/tel-00824944>**

Submitted on 22 May 2013

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

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

**THESE DE DOCTORAT DE L'UNIVERSITE PIERRE ET MARIE CURIE**

Ecole doctorale Complexité du Vivant

Présentée par

**Myriam FENINA**

Pour obtenir le grade de  
**DOCTEUR de l'UNIVERSITÉ PIERRE ET MARIE CURIE**

**Elaboration d'un outil de modification génétique au locus *Dct*  
dans les cellules souches embryonnaires de souris à l'aide de la  
méganucléase *I-SceI***

Directeur de thèse : Pr Jean-Jacques PANTHIER

Soutenue le 23 décembre 2011 devant le jury composé de :

Pr Dominique HIGUET	Président du jury
Dr Olivier DANOS	Rapporteur
Dr Jean-Stéphane JOLY	Rapporteur
Dr Bruno A. BERNARD	Examineur
Dr Jean-Pierre CABANIOLS	Examineur
Pr Jean-Jacques PANTHIER	Directeur de thèse



# Elaboration d'un outil de modification génétique au locus *Dct* dans les cellules souches embryonnaires de souris à l'aide de la méganucléase I-SceI

## Résumé

De nombreux gènes sont surexprimés au cours de la progression du mélanome malin cutané. Pour mieux comprendre le rôle de ces gènes, des études fonctionnelles sont nécessaires. Mon projet visait à développer une nouvelle stratégie pour produire des souris portant n'importe quel ADN d'intérêt inséré à la place du premier exon du gène *Dct*. *Dct* code la dopachrome tautomérase, une enzyme de la voie de biosynthèse de la mélanine, exprimée de façon spécifique dans les mélanocytes et leurs précurseurs. Notre stratégie visait à augmenter la fréquence de recombinaison homologue au locus *Dct* dans des cellules souches embryonnaires (ES) en induisant une cassure double-brin à ce locus grâce aux propriétés endonucléases de la méganucléase I-SceI de levure.

Nous avons créé une lignée de cellule ES qui portent un site unique de reconnaissance de la méganucléase I-SceI inséré dans l'intron 1 du gène *Dct*. Notre hypothèse était que l'induction d'une cassure double-brin au locus *Dct* serait recombinogène, et qu'en présence d'une matrice de réparation, la fréquence des recombinaisons homologues à ce locus serait augmentée de façon significative. Nous avons testé cette hypothèse et intégré successivement les gènes rapporteurs *Lago1* et *H<sub>2</sub>B-mCherry* au locus *Dct*. Nos résultats indiquent que, de façon surprenante, l'induction d'une cassure double-brin ne favorise pas la recombinaison homologue au locus *Dct*. Nous avons analysé l'expression des gènes rapporteurs *Lago1* et *H<sub>2</sub>B-mCherry* insérés au locus *Dct* dans des cellules en culture, l'embryon ou chez l'adulte.



## Remerciements

Je tiens tout d'abord à remercier le Pr Dominique Hiquet, président de mon jury de thèse, ainsi que l'ensemble des autres membres de ce jury d'avoir accepté de se réunir un vendredi 23 décembre, avant-veille de Noël. Je remercie le Dr Olivier Danos d'avoir accepté d'être rapporteur de ma thèse. Nos rencontres ont toujours été riches d'enseignements. Je remercie très sincèrement le Dr Jean-Stéphane Joly d'avoir accepté cette même fonction et d'avoir toujours été disponible pour des réunions de travail, comité de thèse ou autre. J'ai beaucoup apprécié nos échanges sur le sujet. Merci également au Dr Jean-Pierre Cabaniols d'être l'un des examinateurs de ce jury. Votre expertise, issue du secteur privé, permettra d'enrichir le débat.

Je souhaite ici remercier le responsable des ressources humaines de l'Oréal Recherche, M. Demarquette, pour avoir permis le financement de ce doctorat réalisé en convention CIFRE au sein du Département des Sciences du Vivant de l'Oréal et pour avoir aidé à la prolongation de ce travail.

Je remercie très sincèrement le Dr Bruno Bernard, mon responsable chez l'Oréal. Votre collaboration active a été d'une aide précieuse pour l'avancée de mon projet de thèse. Vous étiez toujours force de propositions. Nos réunions régulières m'auront beaucoup appris dans l'art de synthétiser le travail accompli.

Un grand merci au Pr Jean-Jacques Panthier qui a encadré ce travail de thèse. Comme tu le disais si bien, un étudiant qui arrive au laboratoire en Master II est encore un bébé. J'ai maintenant grandi et je suis certaine que tous les enseignements de rédaction et de présentation orale que tu m'as inculqués pendant ces années me seront utiles pour mes futurs projets professionnels.

Je remercie chaleureusement les membres du laboratoire « Soin, Couleur et Qualité du cheveu » de l'Oréal qui m'ont accueilli au début de ma thèse, en particulier Stéphane Commo et Séverine Escaich qui sont encore au rendez-vous lorsque je viens à Clichy. Vous avez toujours été attentifs aux avancées de mon projet et à mes déboires d'expérimentatrice. J'aurais aimé collaborer plus étroitement avec vous, mais mes résultats en avaient décidé autrement.

Je souhaite évidemment associer à ce travail toute l'équipe de l'Unité de Génétique fonctionnelle de la Souris, sans qui cette thèse n'aurait pu aboutir. Je remercie particulièrement Dominique Simon qui a démarré ce projet et qui m'aide à le finaliser maintenant. Je te remercie pour ta gentillesse et pour nos discussions animées, scientifiques et extra-scientifiques. Je tiens à remercier spécialement Sandrine Vandormael-Pournin

*pour sa disponibilité et sa patience face à toutes mes questions. Tu m'as initiée à l'art de la culture des cellules ES. Cet enseignement restera très précieux pour moi. Je souhaite remercier également Michel Cohen-Tannoudji dont les remarques pertinentes ont contribué à la bonne avancée de ce projet. Je remercie Teddy, mon compagnon de galère. Nous aurons fait un bon bout de chemin ensemble. Je te souhaite bon courage pour la reprise des études de pharmacie. Merci beaucoup à Aline, Marie et Satoko pour votre aide et votre écoute. Je souhaite remercier tout spécialement Tânia qui est repartie au Brésil. Tu as été mon pilier au laboratoire pendant mes trois premières années de thèse. Enfin je remercie bien sûr Ziza, notre fée pour résoudre toutes les questions administratives.*

*Je remercie tous mes amis qui ont suivi de près ou de loin mes avancées en tant qu'apprentie chercheuse. La thèse est une étape particulière de la vie et vous m'avez tous aidé dans cette épreuve. Vos questions régulières : « est-ce que tes souris sont bleues finalement ? » ou « as-tu été nourrir tes cellules ce week-end ? » m'ont toujours aidé à prendre du recul par rapport à mon travail. Un merci tout particulier à Laurence pour ton soutien et ta compréhension. Je fais un clin d'œil à Delphine qui soutient sa thèse une semaine avant moi et avec qui les discussions sur nos projets respectifs ont été animées.*

*Je remercie avec beaucoup d'émotions mes parents et mon frère Yacine qui m'ont toujours soutenue et aidée à aller plus loin. Grâce à vous, je suis devenue Myriam Fenina et j'espère vous avoir rendus fiers de moi.*

*Enfin, je dédicace ce manuscrit à Christophe, ou Tof pour les intimes, mon amoureux. Tu as été présent du début à la fin et tu m'as aidée à avoir la force nécessaire pour finir ce travail.*

# SOMMAIRE

Sommaire .....	5
Table des illustrations .....	9
Liste des tableaux.....	11
Glossaire .....	13
Préambule .....	15
Introduction Bibliographique.....	17
Chapitre 1 : Méganucléases et nucléases à doigts de zinc, de nouveaux outils de transgénèse chez les Mammifères.....	17
1 Les techniques classiques de transgénèse chez les Mammifères.....	17
1.1 Transgénèse par micro-injection.....	17
1.2 Transgénèse par recombinaison homologue (RH) dans les cellules souches embryonnaires (cellules ES) .....	21
1.3 Comment augmenter l'efficacité de la RH ?.....	26
2 Les méganucléases (MN) induisent des cassures double-brin (DSB) de l'ADN .....	28
2.1 Historique et nomenclature .....	29
2.2 Diversité, mécanisme d'action et spécificité des MN.....	31
2.3 Premiers ciblage de gènes dans les cellules de mammifères favorisés par I-SceI .....	33
2.4 Conception de MN à façon .....	38
3 Les nucléases à doigts de zinc (ZFN), un outil créé pour induire des DSB .....	40
3.1 Couplage de protéines à doigts de zinc à l'endonucléase <i>FokI</i> .....	40
3.2 Spécificité des protéines à doigts de zinc et mécanisme d'action des ZFN.....	41
3.3 Premiers ciblage de gènes à l'aide de ZFN .....	43
3.4 Méthodes de développement et de délivrance des ZFN .....	45
4 Réparation de DSB dans les cellules de mammifères.....	46
4.1 Les différents mécanismes de réparation induits par un DSB .....	46
4.2 Une singularité des cellules souches embryonnaires (ES) : la recombinaison homologue est le mécanisme de réparation majoritaire.....	53
5 Utilisation de MN et de ZFN pour induire des modifications génétiques chez les Mammifères .....	57
5.1 Utilisation de MN .....	57

5.2 Utilisation de ZFN .....	63
Chapitre 2 : Le gène <i>Dct</i> , de sa spécificité au lignage mélanocytaire à son utilisation en transgénèse.....	71
1 Le mélanocyte, origine embryonnaire et développement du follicule pileux.....	71
1.1 Origine embryonnaire et migration.....	71
1.2 Développement du follicule pileux et cycle de vie.....	83
2 Expression et régulation du gène <i>Dct</i> .....	91
2.1 La DOPAchrome tautomérase, une enzyme de la famille des Tyrosinases .....	91
2.2 Le profil d'expression du gène <i>Dct</i> en fait un bon marqueur du lignage mélanocytaire.....	93
2.3 Régulation du gène <i>Dct</i> .....	97
3 Utilisation du gène <i>Dct</i> en transgénèse.....	99
3.1 La souris transgénique <i>Dct-LacZ</i> .....	99
3.2 La souris <i>Dct<sup>Cre/+</sup></i> : <i>knock in</i> et <i>knock out</i> au locus <i>Dct</i> .....	102
3.3 La lignée de souris avec un chromosome artificiel bactérien (BAC) <i>Dct-LacZ</i> ..	105
3.4 Les souris transgéniques inductibles <i>iDct-LacZ</i> et <i>iDct-GFP</i> .....	107
Objectifs de la thèse.....	113
Chapitre 1 : Elaboration d'un nouvel outil d'ingénierie génétique.....	113
Chapitre 2 : Etude du patron d'expression des gènes rapporteurs <i>Lago1</i> et <i>H2BmCherry</i> intégrés au locus <i>Dct</i> .....	116
Résultats et discussions.....	119
Chapitre 1 : Elaboration d'un nouvel outil d'ingénierie génétique.....	119
1 Article : Une cassure double-brin induite par I- <i>SceI</i> ne favorise pas la recombinaison homologue au locus <i>Dct</i> dans les cellules embryonnaires souches de souris.....	119
2 Discussion et perspectives .....	123
2.1 Le site I- <i>SceI</i> intégré au locus <i>Dct</i> : peut-il être coupé par la méganucléase <i>in vivo</i> ?.....	123
2.2 Le site I- <i>SceI</i> inséré dans le gène <i>Dct</i> est-il placé de façon optimale ? .....	129
2.3 La méganucléase I- <i>SceI</i> est-elle aussi spécifique que nous le pensions ? .....	130
2.4 Les cellules ES se prêtent-elles à une telle stratégie ? .....	131
2.5 Conclusion .....	132
Chapitre 2 : Etude du patron d'expression des gènes rapporteurs <i>Lago1</i> et <i>H2BmCherry</i> inséré au locus <i>Dct</i> .....	135

1 Etude du profil d'expression des gènes rapporteurs <i>Lago1</i> et <i>H2B-mCherry</i> dans les mélanocytes en culture.....	135
1.1 Objectif et stratégie.....	135
1.2 Matériels et méthodes.....	136
1.3 Résultats.....	137
1.4 Conclusion.....	143
2 Etude du phénotype des souris <i>Dct<sup>Lago1-Néo/Lago1-Néo</sup></i> et <i>Dct<sup>Lago1/+</sup></i> : KO et KI de <i>Dct</i> ...	145
2.1 Objectif et stratégie.....	145
2.2 Matériels et méthodes.....	145
2.3 Résultats.....	149
2.4 Conclusion.....	153
3 Discussion et perspectives.....	153
3.1 Le promoteur <i>Dct</i> est-il un promoteur trop faible pour exprimer des gènes d'intérêt dans les mélanocytes et leurs précurseurs ?.....	154
3.2 La construction utilisée est-elle responsable de la faible expression des gènes insérés au locus <i>Dct</i> ?.....	157
References bibliographiques.....	161
Annexe.....	175



# TABLE DES ILLUSTRATIONS

Figure 1 : Obtention de souris transgéniques.....	18
Figure 2 : Micro-injection d'une solution d'ADN dans le pronoyau mâle d'un œuf fécondé de souris. ....	18
Figure 3 : Photo de micro-injection de cellules ES dans la cavité d'un blastocyste de souris. ....	20
Figure 4 : Création de souris génétiquement modifiées via les cellules ES .....	20
Figure 5 : Principe général de la recombinaison homologue.....	23
Figure 6 : Le système Cre/ <i>loxP</i> .....	24
Figure 7 : Principe de propagation d'un intron mobile .....	29
Figure 8 : Comparaison du système I- <i>SceI</i> et du système VDE .....	30
Figure 9 : Mécanisme de clivage par I- <i>SceI</i> .....	32
Figure 10 : Structure du transgène <i>tk</i> non fonctionnel et des vecteurs de réparation .....	35
Figure 11 : Mécanisme de RH induite par I- <i>SceI</i> .....	36
Figure 12 : Schéma d'un motif ZF .....	41
Figure 13 : Schéma d'une paire de ZFN liée à l'ADN .....	42
Figure 14 : Schéma d'une paire de ZFN constituée de domaines de clivage FokI modifiés ...	42
Figure 15 : Les principales étapes du mécanisme de NHEJ .....	47
Figure 16 : Modèles de réparation par RH.....	50
Figure 17 : Les différents mécanismes de réparation d'un DSB.....	52
Figure 18 : Mesure de la fréquence de réparation par RH.....	54
Figure 19 : Ciblage de gène à un locus naturel favorisé par un DSB .....	58
Figure 20 : Différentes stratégies de ciblage de gène favorisés par I- <i>SceI</i> .....	60
Figure 21 : Principe du cGPS.....	62
Figure 22 : Ciblage du gène <i>Oct4</i> à l'aide de ZFN.....	64
Figure 23 : Ciblage de gène au locus ROSA favorisés par des ZFN .....	66
Figure 24 : Représentation schématique de la voie de migration des cellules de crête neurale .....	73
Figure 25 : Voies de migration des cellules de crête neurale .....	75
Figure 26 : Illustration schématique des signaux et interactions dans le système nerveux périphérique .....	76

Figure 27 : Illustration conceptuelle des origines cellulaires et des voies de migration des mélanocytes.....	80
Figure 28 : Etapes de signalisations précoces pour la spécification de la peau de l'embryon .	83
Figure 29 : La morphogenèse du follicule pileux .....	84
Figure 30 : Schéma d'une coupe de bulbe de follicule pileux.....	86
Figure 31 : Cycle de vie du follicule pileux.....	88
Figure 32 : Voie de biosynthèse des mélanines .....	90
Figure 33 : Représentation schématique du gène <i>Dct</i> et de ses transcrits .....	92
Figure 34 : Hybridation <i>in situ</i> sur un embryon de souris entier à E11,5 .....	96
Figure 35 : Coloration X-gal de souris transgéniques <i>Dct-LacZ</i> .....	100
Figure 36 : Comparaison du profil d'expression de <i>LacZ</i> .....	103
Figure 37 : Profils d'expression de <i>LacZ</i> chez les embryons de souris BAC <i>Dct-LacZ</i> .....	106
Figure 38 : Génération de la souris <i>iDct-GFP</i> .....	108
Figure 39 : Principe du projet d'élaboration d'un nouvel outil d'ingénierie génétique .....	114
Figure 40 : Schéma de RH favorisée par I- <i>SceI</i> au locus <i>Dct</i> .....	128
Figure 41 : Fréquence d'insertion de mutations silencieuses selon de la distance à la cassure .....	128
Figure 42 : Coloration X-gal au stade J25 du processus de différenciation .....	139
Figure 43 : Coloration X-gal au stade J41 du processus de différenciation .....	140
Figure 44 : Révélation de l'expression de mCherry par immunofluorescence .....	142
Figure 45 : Principe d'obtention des souris hétérozygotes <i>Dct<sup>Lago1-Néo/+</sup></i> .....	144
Figure 46 : Comparaison de souris homozygotes <i>Dct<sup>Lago1-Néo/Lago1-Néo</sup></i> et de souris sauvages <i>Dct<sup>+/+</sup></i> .....	148
Figure 47 : Comparaison des profils d'expression de la $\beta$ -galactosidase chez un embryon <i>Dct<sup>Lago1/+</sup></i> et chez un embryon <i>Dct-LacZ</i> de stade E11,5 .....	151
Figure 48 : Comparaison des profils d'expression de la $\beta$ -galactosidase sur des coupes de peaux de souris <i>Dct-LacZ</i> et <i>Dct<sup>Lago1/+</sup></i> de 34 jours.....	152
Figure 49 : Comparaison des constructions des souris <i>knock in Dct<sup>Lago1/+</sup></i> , <i>Dct<sup>Cre/+</sup></i> et de la souris transgénique <i>Dct-LacZ</i> .....	156

## **LISTE DES TABLEAUX**

Tableau 1 : Exemples de ciblage de gène réussis via l'induction d'un DSB par des ZFN .....	63
Tableau 2 : Stades d'apparition des marqueurs des mélanocytes dans l'épithélium pigmentaire de la rétine, les mélanoblastes en migration et le télencéphale.....	96



## GLOSSAIRE

AA	acides aminés
ADN	acide désoxyribonucléique
ARNm	acide ribonucléique messenger
BAC	<i>bacterial artificial chromosome</i>
CMV	cytomégalovirus
DCT	DOPAchrome tautomérase
DHI	5,6-dihydroxyindole
DHICA	5,6-dihydroxyindole-2-carboxylique
DOPA	dihydroxyphenylalanine
DSBR	<i>double-strand break repair</i>
DSB	<i>double-strand break</i>
ES	<i>embryonic stem</i>
GFP	<i>green fluorescent protein</i>
kb	kilobases
LM-PCR	<i>ligation-mediated polymerase chain reaction</i>
MN	méganucléase
Néo	Néomycine
NHEJ	<i>non-homologous end joining</i>
pb	paire de bases
RH	recombinaison homologue
RPE	épithélium pigmentaire de la rétine
SCP	<i>schwann cell precursor</i>
SDSA	<i>synthesis-dependant strand annealing</i>
SSA	<i>single-strand annealing</i>
TRE	<i>tet-response element</i>
TRP	<i>tyrosinase related protein</i>
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
YAC	<i>yeast artificial chromosome</i>
ZFN	<i>zinc-finger nuclease</i>



# PREAMBULE

L'objectif de ma thèse était de développer un outil original d'ingénierie génétique qui permette d'exprimer toute séquence d'intérêt dans les mélanocytes et leurs précurseurs chez la souris. Nous avons choisi comme cible le gène *Dct* et comme méthode d'intégration des séquences d'intérêt le *knock in*. L'avantage du *knock in* est que la séquence d'intérêt est exprimée en lieu et place du gène ciblé. Son inconvénient est que la fréquence de recombinaison homologue qui conduit à une intégration de la séquence exogène au locus ciblé est faible et qu'elle varie avec la cible. Nous avons donc choisi de tirer profit des découvertes faites sur la méganucléase I-SceI. I-SceI induit une cassure double-brin à l'endroit où son site de reconnaissance est intégré et favorise un événement de recombinaison homologue au lieu de coupure. Dans ce cadre, notre stratégie a consisté à (1) intégrer un site I-SceI au locus *Dct* dans les cellules souches embryonnaires de souris, (2) tester l'hypothèse selon laquelle l'induction d'une cassure double-brin au locus *Dct* favorise la recombinaison homologue à ce site, (3) étudier le profil d'expression d'un gène rapporteur intégré au locus *Dct* avec notre outil.

L'introduction bibliographique de ce manuscrit se divise en deux chapitres. Le premier chapitre présente les nouveaux outils de transgénèse que sont les méganucléases, dont I-SceI fait partie, et les nucléases à doigts de zinc. Cette étude sera faite dans le cadre des cellules de Mammifères et l'accent sera mis sur le mécanisme de recombinaison homologue. Le deuxième chapitre traite du lignage mélanocytaire et de l'expression de *Dct* dans ce lignage. Nous développerons spécialement l'utilisation du promoteur de ce gène en transgénèse.

Les résultats obtenus au cours de ma thèse sont présentés en deux chapitres. Chaque chapitre est suivi d'une discussion qui replace notre travail dans le contexte bibliographique présenté dans l'introduction. Le premier chapitre expose l'outil d'ingénierie génétique que nous avons développé. La stratégie que nous avons élaborée est ensuite discutée à la lumière de nos connaissances sur les méganucléases. Le deuxième chapitre décrit le profil d'expression des gènes rapporteurs introduits au locus *Dct*, *Lago1* et *H2B-mCherry*. Ces résultats sont ensuite discutés et comparés aux outils de transgénèse existants qui utilisent le promoteur du gène *Dct*.



# INTRODUCTION BIBLIOGRAPHIQUE

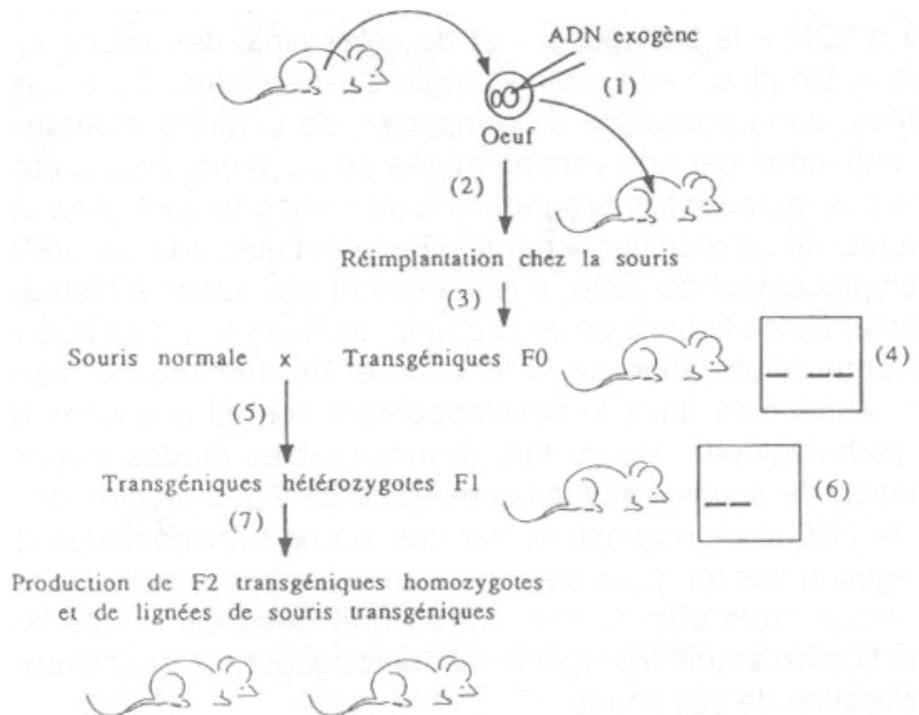
## *Chapitre 1 : Méganucléases et nucléases à doigts de zinc, de nouveaux outils de transgénèse chez les Mammifères*

En trente ans, les techniques de manipulation du génome des Mammifères se sont considérablement améliorées. Les chercheurs ont affiné les outils génétiques et on est loin de la simple injection d'un ADN nu qui a permis l'obtention de la première souris transgénique. L'utilisation des cellules souches embryonnaires, des méganucléases et des nucléases à doigt de zinc permet aujourd'hui des constructions « à façon ». Le mécanisme de modification génétique qui fait l'objet de ce chapitre est la recombinaison homologue. Celle-ci permet de modifier le génome avec une très grande précision et d'intégrer dans le génome hôte des séquences de quelques nucléotides à plusieurs kilobases. Dans ce chapitre, nous passerons en revue les différentes techniques de transgénèse classique utilisées chez les Mammifères. Nous discuterons ensuite de la découverte de nouveaux outils utilisables en transgénèse, les méganucléases et les nucléases à doigts de zinc. Nous nous intéresserons aux mécanismes de réparation dont disposent les cellules de mammifères pour s'affranchir des effets délétères des cassures double-brin induites par ces nucléases. Enfin, nous présenterons quelques exemples d'utilisation de méganucléases et de nucléases à doigts de zinc en transgénèse chez les Mammifères.

## **1 Les techniques classiques de transgénèse chez les Mammifères**

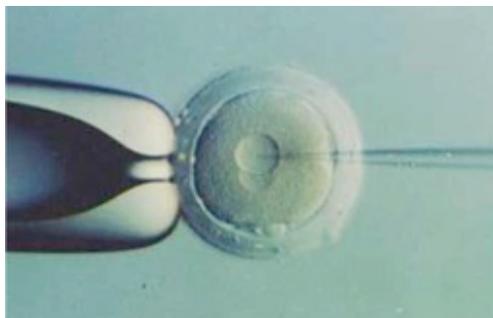
### 1.1 Transgénèse par micro-injection

La transgénèse par micro-injection, aussi appelée transgénèse par « addition » de gène, consiste à injecter le gène cloné – appelé transgène – directement dans l'un des deux pronucléi de l'œuf tout juste fécondé. La première souris transgénique par micro-injection est obtenue au tout début des années 1980. Les expérimentateurs ont introduit le gène codant la



**Figure 1 : Obtention de souris transgéniques**  
(Babinet, 2003)

La solution d'ADN est injectée (1) dans un pronucléus de l'œuf tout juste fécondé, à l'aide d'un micromanipulateur qui commande deux micropipettes, l'une pour la micro-injection de l'œuf, l'autre pour son maintien (voir **B**). L'injection entraîne le gonflement du noyau (voir **B**). Les œufs injectés sont ensuite réimplantés dans une mère porteuse (2, 3). Dans un petit nombre de cas, l'ADN injecté est intégré dans le génome de l'hôte et ceci avant la première division. Au fil des divisions, l'ADN exogène ou transgène est transmis aux cellules filles et une souris, dite transgénique, va naître, contenant le transgène dans toutes ses cellules, y compris les cellules de la lignée germinale. Les souris transgéniques sont repérées par la présence du transgène dans leur ADN total (4). Le croisement des souris transgéniques avec une souris non transgénique donne naissance à des individus F1 dont la moitié a hérité du transgène (5, 6). Le croisement des deux souris F1 produit des souris dont un quart est homozygote pour le transgène (7). La lignée transgénique est ainsi créée.



**Figure 2 : Micro-injection d'une solution d'ADN dans le pronoyau mâle d'un œuf fécondé de souris.**

thymidine kinase du virus herpès simplex, sous le contrôle du promoteur du gène murin *Mt1* (*Metallothionein-1*), dans le pronucléus d'un œuf fécondé. Puis cet œuf a été réimplanté dans l'oviducte d'une femelle pseudo-gestante qui a donné naissance à des souris transgéniques. Les expérimentateurs ont alors étudié l'expression du transgène dans les cellules somatiques des animaux transgéniques (Brinster, *et al.*, 1981). Depuis, cette méthode est la plus largement utilisée en transgénèse chez les Mammifères. L'animal transgénique contient une à plusieurs dizaines de copies du transgène intégrées au hasard dans le génome. Ni le nombre de copies du transgène, ni le site d'insertion ne sont contrôlés. La transgénèse par micro-injection est couramment utilisée car le transgène, s'il contient les séquences de régulation appropriées, est généralement exprimé de manière spécifique. Les **Figure 1** et **Figure 2** présente la méthode d'obtention de lignées de souris transgéniques par micro-injection dans l'œuf fécondé.

Avec l'avènement de la transgénèse, le « ciblage » de l'expression d'un gène exogène a été rendu possible. Cette technique permet de construire des transgènes qui contiennent les séquences régulatrices d'un gène A et les séquences codantes d'un gène B. Les souris transgéniques fabriquées avec un tel transgène exprimeront en principe la protéine codée par le gène B dans les tissus ou types cellulaires où est normalement exprimé le gène A. Le gène B est alors un gène « rapporteur » de l'activité des séquences régulatrices du gène A. D'autres applications possibles de la transgénèse par micro-injection sont l'étude fonctionnelle ou physiopathologique d'un gène et la production de protéines d'intérêt biologique ou médical (pour revue voir (Babinet, 2003)).

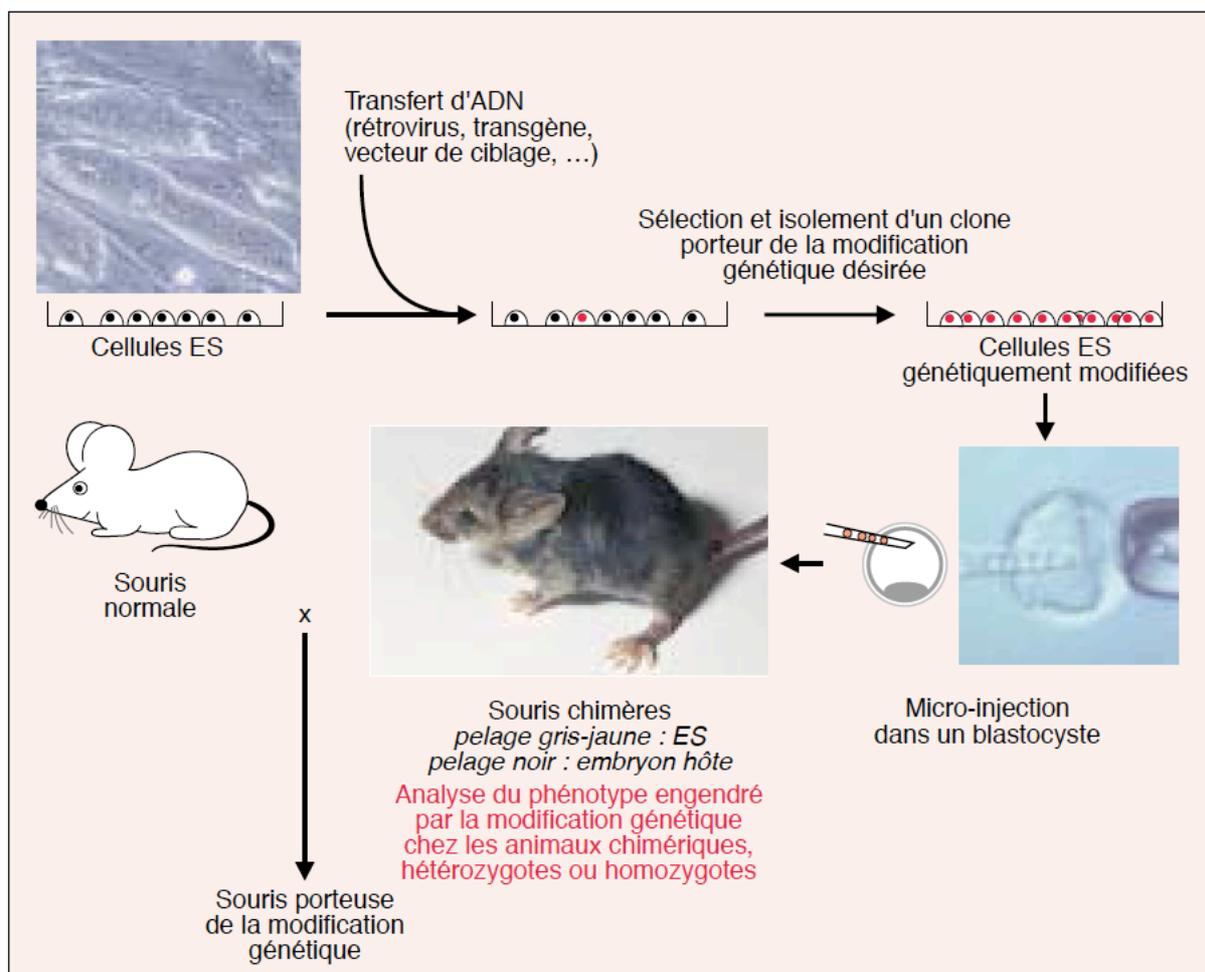
La limitation de cette approche est liée au mode d'obtention des souris transgéniques qui fait que le transgène est intégré en une ou plusieurs copies au hasard dans le génome. Les séquences voisines du site d'intégration peuvent interférer avec l'expression du transgène ; ce phénomène est appelé « effet de position ». Il se manifeste de deux façons :

- le niveau d'expression d'un transgène varie d'une lignée transgénique à l'autre, sans que le nombre de copies du transgène en soit la cause ;
- le transgène peut être exprimé de manière ectopique.

Le risque d'observer un « effet de position » oblige l'expérimentateur à produire plusieurs lignées indépendantes de souris transgéniques. L'étude de l'ensemble des lignées permet de distinguer, selon les phénotypes observés, la part de l'expression qui relève de la structure du transgène *per se* de l'expression ectopique due aux effets de position. Il a été découvert que des séquences régulatrices essentielles à l'expression correcte d'un gène peuvent être situées à



**Figure 3 : Photo de micro-injection de cellules ES dans la cavité d'un blastocyste de souris**  
 Les cellules ES sont injectées dans la cavité d'un blastocyste à l'aide d'un micromanipulateur qui commande deux micropipettes, l'une pour la micro-injection des cellules ES, l'autre pour le maintien du blastocyste.



**Figure 4 : Création de souris génétiquement modifiées via les cellules ES**  
 (Babinet and Cohen-Tannoudji, 2000)

plusieurs dizaines de kilobases de ses séquences codantes. Cependant, la taille d'un transgène a longtemps été limitée à 40-50 kb de séquences pour des raisons techniques. La mise au point de chromosomes artificiels de levure (YAC pour *Yeast Artificial Chromosome*) et de bactérie (BAC pour *Bacterial Artificial Chromosome*) a permis de fabriquer des transgènes qui atteignent deux mégabases de longueur. A l'heure actuelle, la technique de micro-injection est toujours limitée par l'absence de maîtrise du niveau d'expression du transgène et le risque des effets de position.

## 1.2 Transgénèse par recombinaison homologue (RH) dans les cellules souches embryonnaires (cellules ES)

En 1981, deux équipes, en particulier celle de Martin Evans – prix Nobel de médecine en 2007 –, rapportent l'isolement de lignées de cellules à partir de la culture directe de blastocystes de souris *in vitro* (Evans and Kaufman, 1981; Martin, 1981). Ces chercheurs ont réussi à dériver des lignées de cellules appelées « cellules souches embryonnaires » ou cellules ES (pour *embryonic stem cells*). Ces cellules ont la propriété de conserver leur pluripotence au cours de leur croissance *in vitro*. En effet, une fois réintroduites par micro-injection dans un blastocyste (**Figure 3**), elles sont capables de participer à la formation de tous les tissus de l'embryon, y compris la lignée germinale. L'individu ainsi obtenu est une chimère constituée d'un ensemble de cellules issu pour partie des cellules ES et pour partie des cellules de l'embryon hôte. Le croisement de cette chimère avec une souris de génotype sauvage donnera naissance, avec une certaine proportion, à des individus dont le génome provient pour moitié des cellules ES. Cette proportion reflète le degré de colonisation de la lignée germinale de la chimère par les descendants des cellules ES injectées. Il a ensuite été démontré que les cellules ES peuvent être modifiées *in vitro*, par exemple par l'introduction d'un transgène, et que des souris transgéniques correspondantes peuvent être obtenues (Gossler, *et al.*, 1986; Robertson, *et al.*, 1986) (**Figure 4**). Il est aussi intéressant de noter que, dans des conditions particulières de culture *in vitro*, les cellules ES sont capables de donner naissance à différents types de cellules différenciées (neurones, cellules musculaires, mélanocytes, etc.).

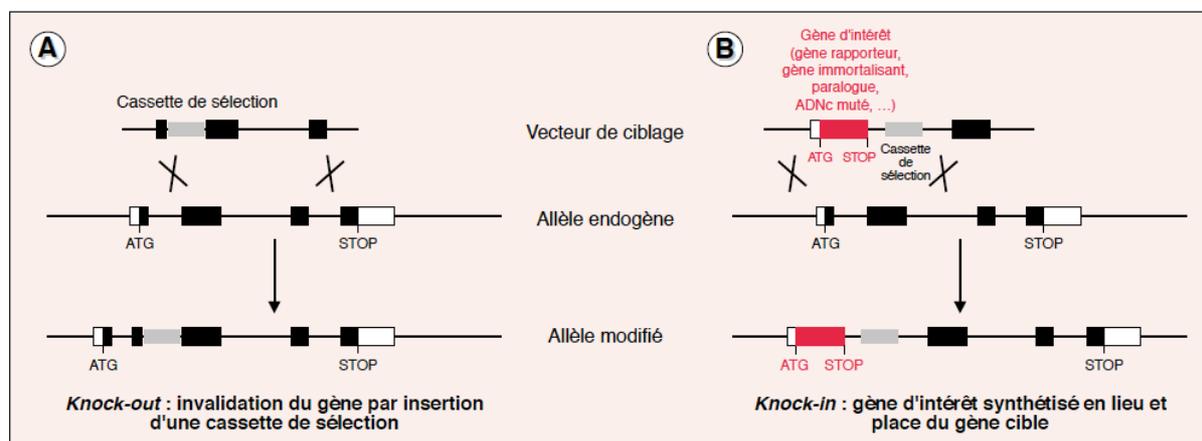
### 1.2.1 Création de mutations nulles (souris knock-out)

Au milieu des années 1980, des études réalisées par les équipes de Mario Capecchi et Oliver Smithies – tous deux prix Nobel de médecine en 2007 – ont démontré que les cellules de Mammifères sont capables d'effectuer une recombinaison entre une séquence d'ADN exogène et la séquence homologue présente *in situ* dans les chromosomes (Smithies, *et al.*, 1985; Wong and Capecchi, 1986). Cet événement reste relativement rare en comparaison de l'intégration au hasard de ce même ADN. À partir de ces premiers travaux, ils ont montré qu'il était possible de cibler le gène *Hprt* (*Hypoxanthine-guanine phosphoribosyl transferase*) des cellules ES, pour créer, ou au contraire pour corriger, des mutations nulles dans celui-ci (Doetschman, *et al.*, 1987; Thomas and Capecchi, 1987). La particularité du gène *Hprt* est que des mutations, aussi bien perte-de-fonction (LOF) que gain-de-fonction (GOF), peuvent être directement sélectionnées en culture cellulaire par l'ajout de drogues dans le milieu. Avec la mise au point un an plus tard de la stratégie de sélection « positive-négative », le groupe de Mario Capecchi a démontré que n'importe quel gène de souris peut être ciblé dans les cellules ES (Mansour, *et al.*, 1988). La publication des premières souris mutantes obtenues par recombinaison homologue dans les cellules ES date de 1989 (Koller, *et al.*, 1989; Thompson, *et al.*, 1989; Zijlstra, *et al.*, 1989).

Pour générer une souris dite *knock out* (KO) pour un gène, un transgène est construit avec la séquence du gène ciblé, accompagnée d'une cassette de sélection. Cette cassette de sélection code pour un gène de résistance à un antibiotique. Elle est insérée à l'intérieur d'un exon de manière à empêcher la formation d'un ARN messager fonctionnel, ce qui aboutit à la création d'un allèle nul. Ce transgène, alors appelé « vecteur de ciblage », est transfecté dans les cellules ES. La culture des cellules ES en présence de l'antibiotique permet de sélectionner des clones qui ont intégré le transgène. La proportion des clones qui ont intégré le transgène par recombinaison avec les séquences homologues du gène endogène – on parle alors de « recombinaison homologue » (RH) – est variable d'un gène à l'autre mais elle est généralement faible (**Figure 5A**). La méthode de sélection « positive-négative », développée par l'équipe de Capecchi, permet de faciliter l'obtention des clones ayant effectivement subi une recombinaison homologue, mais il est certain que la fréquence de RH n'est pas modifiée par le processus de sélection. Après sélection des clones de cellules ES recombinants homologues, les souris KO, porteuses de la mutation nulle, peuvent être générées via l'obtention de chimères. L'analyse des phénotypes engendrés par ces mutations a élargi la connaissance sur la fonction des gènes concernés.

### 1.2.2 Une variante du knock out, le knock in

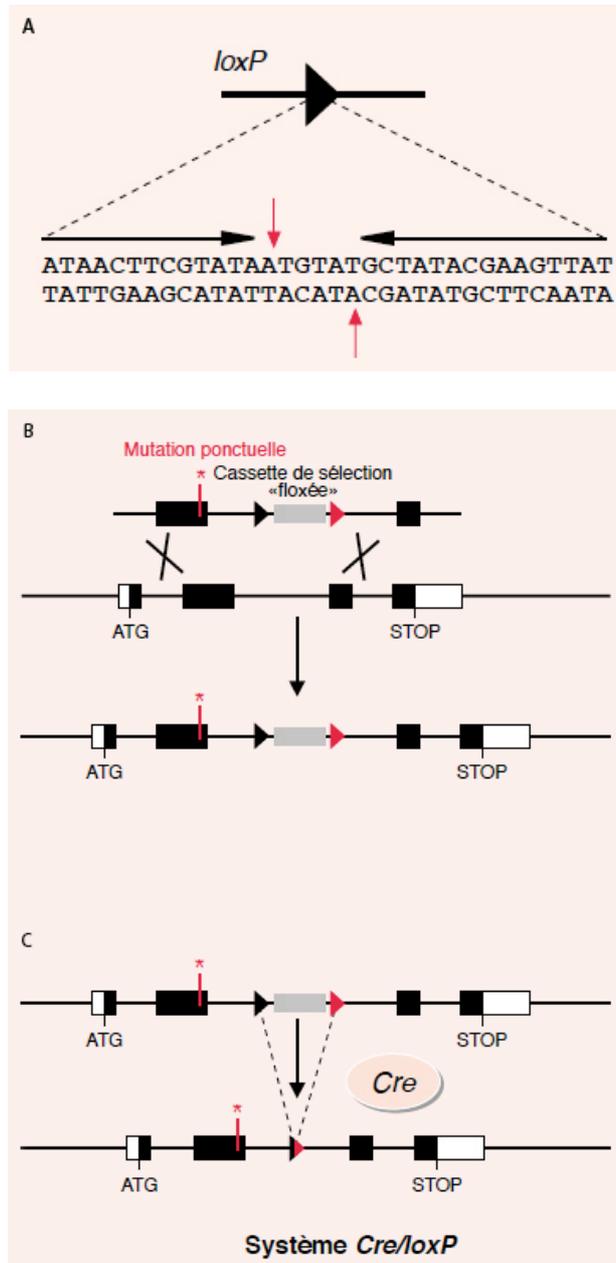
Une variante du *knock out* consiste à introduire, au sein de la séquence d'un gène endogène ciblé, une séquence d'ADNc codant une protéine d'intérêt. Après recombinaison homologue, l'allèle modifié exprime la protéine d'intérêt en lieu et place de celle normalement codée par le gène endogène, car l'ADNc est placé sous le contrôle du promoteur et des séquences régulatrices de ce gène. Le *knock in* entraîne généralement une invalidation complète du gène ciblé (**Figure 5B**). L'utilisation d'un gène rapporteur, comme *LacZ* codant la  $\beta$ -galactosidase, permet d'une part, de mimer l'expression du gène ciblé, ce qui facilite la détermination de son profil d'expression, d'autre part de suivre le destin des cellules qui expriment normalement le gène ciblé.



**Figure 5 : Principe général de la recombinaison homologue**  
(Babinet and Cohen-Tannoudji, 2000)

- A-** *Knock out*. Le vecteur de ciblage contient une cassette de sélection insérée dans un exon et entourée de régions d'homologie avec le gène cible. La recombinaison s'effectue au niveau des séquences homologues pour créer un allèle nul, où l'invalidation du gène résulte de l'insertion de la cassette de sélection.
- B-** *Knock in*. Dans cette variante, outre l'invalidation du gène cible, un gène d'intérêt est introduit dans le locus. Après recombinaison homologue, ce gène d'intérêt est placé sous le contrôle du promoteur et des séquences régulatrices du gène cible et est donc synthétisé en lieu et place du gène cible.

Il est intéressant de noter que les utilisations d'une souris *knock in* sont assez proches des utilisations d'une souris transgéniques par micro-injection. La technique du *knock in* est généralement plus intéressante car le transgène est intégré en une seule copie et est contrôlé par tous les éléments de régulation du gène ciblé. En outre, aucun effet de position n'est



**Figure 6 : Le système Cre/loxP**  
(Babinet and Cohen-Tannoudji, 2000)

- A- Le site *loxP*, symbolisé ici par un triangle, est une séquence de 34 pb composée de deux séquences palindromiques de 13 pb, séparées par une séquence de 8 pb. La recombinase Cre reconnaît spécifiquement cette séquence, provoque la coupure de l'ADN (flèches rouges) et induit la recombinaison de l'ADN de part et d'autre des deux sites *loxP*.
- B- Dans un premier temps, le gène cible est modifié à l'aide d'un vecteur de ciblage comprenant une mutation subtile et une cassette de sélection « floxée », c'est-à-dire entourée de deux sites *loxP* dans la même orientation.
- C- Dans un deuxième temps, l'expression transitoire de la recombinase Cre dans les cellules recombinantes induit la délétion de la cassette de sélection. En dehors de la modification subtile souhaitée, seul un site *loxP* de 34 pb persiste dans l'allèle modifié final. La position de ce site *loxP* est choisie de manière à ne pas interférer avec l'expression du gène cible (en général dans un intron).

observé. Comme expliqué ci-dessus, il est nécessaire de générer plusieurs lignées indépendantes de souris transgéniques par micro-injection pour étudier le phénotype dû au site d'intégration de l'ADNc d'intérêt. La technique du *knock in* permet de s'affranchir de ce problème. La difficulté majeure à résoudre tient alors au fait que la proportion des clones ayant intégré le transgène par RH est généralement très faible – de l'ordre de  $10^{-5}$  à  $10^{-8}$  – chez les cellules d'eucaryotes supérieurs. Cela les différencie des organismes plus simples comme la levure, les trypanosomes ou les *Leishmania*.

### 1.2.3 Le système Cre/loxP

Les mutations nulles sont un instrument génétique puissant. Cependant, les maladies génétiques humaines sont rarement engendrées par des mutations nulles. Par exemple, la mucoviscidose est une maladie génétique humaine dont environ 70% des mutations correspondent à la délétion d'un seul acide aminé (Kerem, *et al.*, 1989). La création de mutations ponctuelles (un seul acide nucléique muté) ou de courtes délétions permet d'étudier des maladies génétiques ou de raffiner l'analyse de la fonction de protéines. Particulièrement dans ces cas-là, il est nécessaire de se débarrasser de la cassette de sélection qui peut interférer avec la régulation de l'expression du gène ciblé ou de l'ADNc d'intérêt (pour revue voir (Fiering, *et al.*, 1999)). La suppression de la cassette de sélection permet de créer une mutation dite « propre », c'est-à-dire dépourvue de séquences étrangères. Pour cela, le système le plus utilisé est le système Cre/loxP. La protéine Cre est une recombinase, identifiée chez le bactériophage P1. Elle agit lorsqu'elle reconnaît, dans un segment d'ADN, une séquence de 34 paires de bases (pb), appelée *loxP* (pour revue voir (Kilby, *et al.*, 1993)). Lorsque deux sites *loxP* sont orientés dans le même sens, la Cre recombinase induit par recombinaison intrachromosomique – dite « site-spécifique » – la délétion de la séquence d'ADN située entre les deux sites (**Figure 6**). A l'issue de la recombinaison site-spécifique, l'allèle muté porte un site *loxP* a priori neutre.

Dans l'ensemble de ces techniques, la mutation est portée par toutes les cellules de la souris. Cette caractéristique limite bon nombre d'études, en particulier, lorsqu'une mutation dans un gène donné entraîne une létalité. La fonction éventuelle de ce gène ne peut être analysée au-delà du moment où intervient la létalité. L'utilisation du système Cre/loxP a permis de s'affranchir de cette limitation et de mettre au point des stratégies qui déclenchent la survenue conditionnelle d'une mutation chez l'animal en cours de développement, ou chez l'adulte (Gu, *et al.*, 1994). Le principe est de développer dans un premier temps des souris porteuses d'un allèle dans lequel deux sites *loxP* encadrent une partie essentielle du gène

d'intérêt sans pour autant perturber son fonctionnement. Ces souris sont ensuite croisées avec une souris transgénique qui exprime la Cre dans un type cellulaire, grâce à l'emploi de séquences régulatrices appropriées. Chez les souris double-transgéniques issues de ce croisement, la recombinase entraîne l'excision des séquences situées entre les sites *loxP* et donc induit une mutation LOF dans le type cellulaire où elle s'exprime. Cette stratégie permet de contourner le problème de la létalité embryonnaire, mais aussi d'examiner les effets de cette mutation dans n'importe quel tissu.

Au contrôle de la mutation dans l'espace peut s'ajouter un raffinement supplémentaire qui est le contrôle de la mutation dans le temps, par un « système inductible ». Dans ce cas, la recombinase Cre est exprimée sous la forme d'une protéine de fusion avec le domaine de fixation au ligand d'un récepteur aux stéroïdes ou LBD (pour *ligand binding protein*) (Feil, *et al.*, 1996; Kellendonk, *et al.*, 1996). Par défaut, cette protéine de fusion ne présente pas d'activité Cre. Cependant, en présence du ligand approprié, un changement de conformation de la protéine s'opère, entraînant la restauration de l'activité Cre. Cette technique requiert la production de souris transgéniques qui portent deux allèles du gène d'intérêt « floxés » et le transgène d'expression LBD/Cre sous la dépendance d'un promoteur spécifique d'un tissu donné. L'induction de la mutation LOF dans ce tissu peut être obtenue par l'injection, au moment voulu, du ligand approprié.

Ces différents niveaux de raffinement dans le contrôle des mutations sont très précieux pour élargir le domaine des connaissances sur la fonction des gènes. Ces méthodes restent néanmoins dépendantes de la fréquence du mécanisme de RH, variable pour chaque locus.

### 1.3 Comment augmenter l'efficacité de la RH ?

Comme remarqué précédemment, la principale limite du mécanisme de recombinaison homologue (RH) dans les cellules ES est sa faible fréquence ; l'événement de RH a lieu dans une cellule sur  $10^5$  à  $10^8$  environ. Le développement de systèmes de sélection positive et négative facilite l'obtention des clones recombinants homologues mais n'en modifie pas la fréquence. La sélection des cellules recombinantes homologues reste longue, coûteuse et traumatisante pour les cellules. De plus, cette méthode n'a pas pu être étendue à l'ensemble des types cellulaires chez les Mammifères car la RH semble surtout efficace dans les cellules ES. Pour augmenter la fréquence de RH, l'expérimentateur peut jouer sur trois facteurs : (1) le vecteur de ciblage, (2) les activités de recombinaison de la cellule et (3) la cible génomique.

Les recombinaisons extrachromosomiques ont été étudiées de façon intensive chez les Mammifères (pour revue voir (Bollag, *et al.*, 1989)). Les laboratoires travaillant sur ce sujet ont découvert que :

- l'induction, dans un des plasmides transfectés, d'une cassure double brin (ou DSB, pour double strand break) proche d'une séquence de RH augmente d'environ 10 fois le taux de RH,
- lorsqu'un seul des deux plasmides subit une DSB, ce plasmide joue le rôle de « receveur » de l'échange d'information non réciproque.

Suite à ces découvertes chez les Mammifères, les chercheurs étudiant la RH entre une molécule exogène et un locus endogène, ont utilisé des vecteurs de ciblage linéarisés. Pour cela, ils induisaient une DSB dans le plasmide, à proximité des régions d'homologie (pour revue voir (Bollag, *et al.*, 1989)). Par ailleurs, la quantité de vecteur de ciblage introduit dans les cellules n'a pas ou peu d'effet sur la fréquence de la RH. Il semble qu'une quantité importante de vecteur de ciblage favorise même d'avantage l'intégration non-homologue (Vasquez, *et al.*, 2001). La taille et le degré d'homologie que le vecteur de ciblage partage avec sa cible sont quant à eux des facteurs déterminants. Plus la taille de l'homologie est grande, plus l'efficacité de RH est importante (Deng and Capecchi, 1992). Au-delà d'une certaine taille, l'expérimentateur sera limité par la difficulté de produire le vecteur de ciblage dans une bactérie. Inversement, la fréquence des événements de recombinaison infidèle augmente lorsque l'un des deux bras d'homologie est d'une longueur inférieure à 1 kb (Scheerer and Adair, 1994). Ceci est très probablement lié à la dégradation qui se produit naturellement aux extrémités du vecteur linéaire. Concernant le degré d'homologie, l'existence de petites divergences entre les régions d'homologie du vecteur de ciblage et sa cible contribue également à la faible fréquence des événements de RH (Deng and Capecchi, 1992). Il est donc essentiel d'utiliser des séquences isogéniques, c'est-à-dire de la même lignée que les cellules à modifier. Accroître la taille et le degré d'homologie ne suffit pourtant pas à augmenter significativement la fréquence de RH.

La modulation des activités de recombinaison de la cellule est un autre moyen pour augmenter la fréquence de la RH. Par exemple, il a été montré que la surexpression de RAD51 dans des cellules humaines, à deux loci différents, augmente la fréquence de la RH de deux à trois fois (Yanez and Porter, 1999). Ce résultat suggère que la protéine RAD51 est impliquée dans les mécanismes de RH et que son niveau d'expression est limitant. Je ne développerai pas davantage ma revue sur ces travaux, car leur intérêt pour augmenter la

fréquence de RH est, à ce jour, limitée. De plus, il a été montré que plusieurs lignées de cellules cancéreuses sur-expriment la protéine RAD51 (Raderschall, *et al.*, 2002). Ainsi, la modification de l'expression du gène *Rad51* pourrait entraîner, parallèlement à une amélioration de l'efficacité de ciblage du gène, la transformation des cellules manipulées (Maacke, *et al.*, 2000).

Le principal obstacle à une meilleure efficacité de la RH est le fait qu'une cible génomique intacte est peu réactive. Il a été démontré que des dommages réalisés sur le gène cible entraînent une augmentation de la fréquence de recombinaison. Les premières expériences dans ce domaine ont montré que l'introduction d'agents responsables de dommages à l'ADN stimule la RH entre les chromatides sœurs (Latt, 1981).

Tout au long de cette première partie, nous avons vu que la recombinaison homologue est le mécanisme privilégié pour obtenir différents types de mutations dans les cellules ES de souris, de la mutation ponctuelle à l'introduction d'un ADNc d'intérêt. Cependant, la fréquence faible de cet événement dans les cellules ES et son caractère pratiquement inexistant dans les autres types cellulaires chez les Mammifères limite son utilisation. Une avancée majeure a été faite lorsque des études ont montré qu'une cassure double-brin (DSB) unique au niveau du gène cible augmente de façon très importante la fréquence de recombinaison à cet endroit précis. Ces études ont été inspirées par la découverte des événements naturels de recombinaison initiés par des DSB chez la levure.

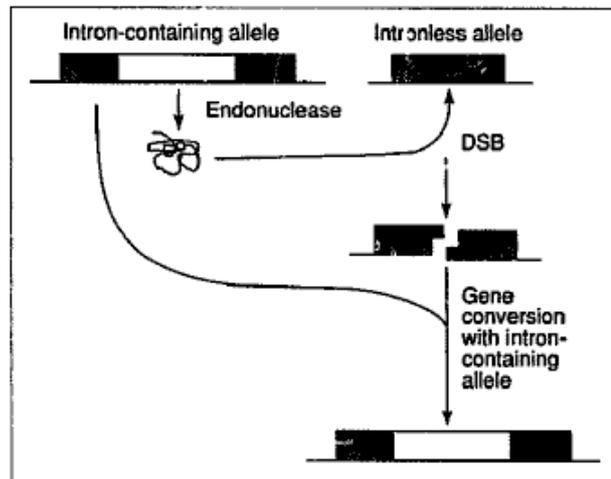
## **2 Les méganucléases (MN) induisent des cassures double-brin (DSB) de l'ADN**

Les méganucléases (MN), aussi appelées endonucléases de homing (*homing endonucleases*), sont des enzymes de restriction qui reconnaissent sur l'ADN un site de clivage asymétrique et de grande taille et y induisent une cassure double-brin (DSB).

## 2.1 Historique et nomenclature

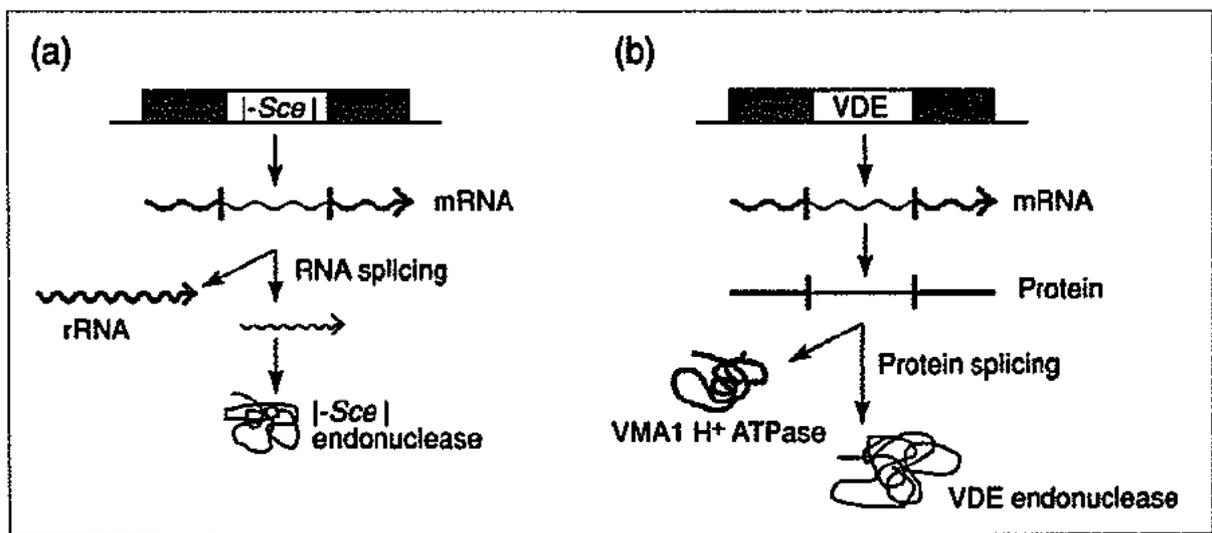
La nomenclature utilisée pour les enzymes MN suit le principe suivant : leur nom commence par un I (pour intron), suivi de l'abréviation de l'espèce d'origine et d'un nombre en chiffre romain qui indique l'ordre de découverte. Par exemple, I-SceI est la première endonucléase intronique découverte chez la levure *Saccharomyces cerevisiae* (Belfort and Roberts, 1997).

Une catégorie de MN, les endonucléases introniques, participe au processus de transfert d'intron, un phénomène commun aux Eucaryotes inférieurs, les bactéries et les bactériophages. Un intron dit intron mobile présent sur l'allèle d'un gène donné est copié puis transféré sur l'autre allèle du même gène dépourvu de cet intron. L'intron mobile des introns de groupe I code une MN qui induit un DSB à son site d'insertion. Ainsi, les introns mobiles favorisent leur propre propagation au sein d'une population cellulaire (**Figure 7**). Les introns de groupe I sont localisés dans l'ADN des mitochondries de levures, d'algues, le noyau des protistes et dans les bactériophages.



**Figure 7 : Principe de propagation d'un intron mobile**  
(Jasin, 1996)

Une endonucléase spécifique d'un site est produite à partir de l'intron d'un allèle d'un gène. Elle reconnaît et coupe une version du même gène sans intron, au niveau du site d'insertion de l'intron. Cela déclenche le phénomène de conversion génique avec l'allèle qui contient l'intron. Le site de reconnaissance se situe aux extrémités des deux exons du gène ; il est non-palindromique et souvent complexe. Ainsi, les endonucléases introniques ont beaucoup moins de sites de reconnaissance chez les génomes complexes que n'en ont les enzymes de restriction des bactéries.



**Figure 8 : Comparaison du système I-SceI et du système VDE**  
(Jasin, 1996)

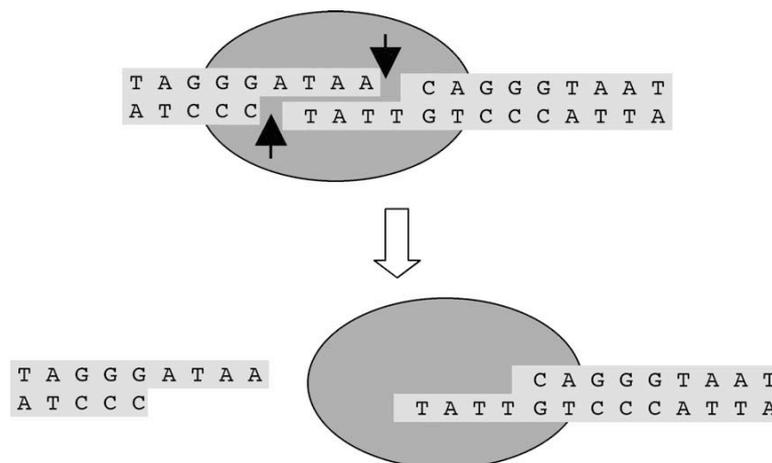
L'expression des méganucléases introniques résulte de (a) l'épissage de l'ARN (ex : I-SceI) ou (b) l'épissage de la protéine à partir d'un polypeptide primaire (ex : VDE).

La MN I-*SceI* est codée par un intron de groupe I. I-*SceI* est probablement la MN la mieux caractérisée (Colleaux, *et al.*, 1986; Jacquier and Dujon, 1985; Monteilhet, *et al.*, 1990). Elle a été découverte dans les années 1970 en étudiant la recombinaison de l'ADN mitochondrial chez les levures (Dujon, *et al.*, 1976; Heyting, *et al.*, 1979; Heyting and Menke, 1979). Au niveau moléculaire, les souches  $\omega^+$  ont un intron dans le gène codant l'ARNr 21S, tandis que dans les souches  $\omega^-$ , le gène de l'ARNr 21S est dépourvu d'intron. Lorsque des souches  $\omega^+$  et  $\omega^-$  sont croisées, le phénomène de propagation d'intron a lieu, avec une efficacité tellement importante que la conversion du locus  $\omega^-$  en locus  $\omega^+$  est égale à 95%. L'intron présent dans les souches  $\omega^+$  code la MN I-*SceI*. Celle-ci coupe l'allèle dépourvu d'intron des souches  $\omega^-$ , au niveau d'une séquence non-palindromique située à la jonction entre les deux exons. Cette DSB très spécifique permet d'initier le processus de transfert d'intron.

## 2.2 Diversité, mécanisme d'action et spécificité des MN

Comme nous venons de le voir, I-*SceI* appartient à la famille des endonucléases introniques. Il existe deux autres familles de MN. La deuxième correspond aux MN codées par des gènes indépendants ; ceux-ci possèdent leur propre unité transcriptionnelle. Cette famille contient deux endonucléases : HO et Endo-*SceI*, toutes deux issues de la levure *Saccharomyces cerevisiae*. La troisième famille est constituée des MN intercalaires ou intéines (*intein* en anglais) qui dérivent de l'épissage d'un polypeptide primaire. Les séquences codantes des MN intercalaires sont insérées dans les phases ouvertes de lecture de gènes de bactéries, d'Archaeobactéries et de levures. La MN VDE (pour *Vma1-derived endonuclease*) appartient à cette dernière famille. Le gène *Vma1* est transcrit en ARN messager qui est traduit en un polypeptide primaire. Les domaines N-terminal et C-terminal du polypeptide primaire sont épissés et soudés pour former une sous-unité ATPasique (Kane, *et al.*, 1990). De cette façon, un second polypeptide est libéré : la MN intercalaire VDE. Le système I-*SceI* est mitochondrial et basé sur l'épissage d'un ARN messager. Le système VDE est nucléaire et basé sur l'épissage d'un polypeptide primaire (**Figure 8**). La MN VDE clive un allèle *Vma1* qui ne contient pas la séquence intercalaire ; cela permet l'insertion de la séquence codante de VDE au sein du gène *Vma1*.

Les MN sont généralement de petite taille : I-CreI, la plus petite, est composée de 163 acides aminés (AA) et HO, la plus grande, comprend 586 AA. Les MN agissent sous la forme de monomères ou d'homodimères. Elles sont réparties en quatre familles caractérisées par la présence d'un motif d'acides aminés : 1) LAGLIDADG – le groupe majoritaire –, 2) GIY-YIG, 3) H-N-H, et 4) à boîte His-Cys (Belfort and Roberts, 1997). La MN I-SceI est une protéine globulaire de 235 AA (Monteilhet, *et al.*, 1990). Elle appartient à la famille LAGLIDADG et est formée de deux exemplaires de ce motif d'acides aminés. I-SceI agit sous forme de monomère. Elle clive son site de reconnaissance de façon asymétrique en présence d'un cation divalent, Mg<sup>2+</sup> ou Mn<sup>2+</sup>. À la suite de cette coupure, I-SceI libère une extrémité 3'OH sortante de 4 nucléotides (nt) (Monteilhet, *et al.*, 1990). I-SceI a une forte affinité pour la partie longue de son demi-site de reconnaissance (**Figure 9**) (Perrin, *et al.*, 1993). Par conséquent, elle se sépare lentement de son substrat.



**Figure 9 : Mécanisme de clivage par I-SceI**  
(Grabher, *et al.*, 2004)

I-SceI agit sous forme d'un monomère ; cette méganucléase reconnaît et coupe sa séquence de reconnaissance de 18 pb de façon asymétrique. Elle a un faible taux de renouvellement du fait de sa forte affinité pour la partie longue de son demi-site de reconnaissance.

La taille du site de reconnaissance des MN varie de 12 à 40 nt et leurs séquences ne présentent aucune symétrie apparente (Belfort and Roberts, 1997). Certaines MN sont peu spécifiques et donc capables de reconnaître leur séquence de reconnaissance modifiée de quelques nucléotides, comme Endo.SceI. D'autres sont décrites comme spécifiques, comme

I-SceI. En outre, I-SceI reconnaît une séquence de 18 pb. En supposant une nature aléatoire des séquences nucléotidiques, ce site de reconnaissance est présent une fois toutes les  $7 \times 10^{10}$  pb ( $1/4^{18}$ ), ce qui correspond à 20 fois la taille du génome des Mammifères (Jasin, 1996). D'un point de vue structural, la spécificité de I-SceI pour son site de reconnaissance est due à de nombreuses liaisons anhydrides phosphoriques directes, des contacts spécifiques entre bases et des interactions entre la MN et son site via des molécules d'eau (Moure, *et al.*, 2003).

### 2.3 Premiers ciblage de gènes dans les cellules de mammifères favorisés par I-SceI

Le terme de ciblage de gène désigne l'événement d'intégration d'un ADN exogène par RH dans le génome de l'organisme hôte. L'insertion de l'ADN exogène nécessite la construction d'un vecteur de ciblage, ou vecteur de réparation.

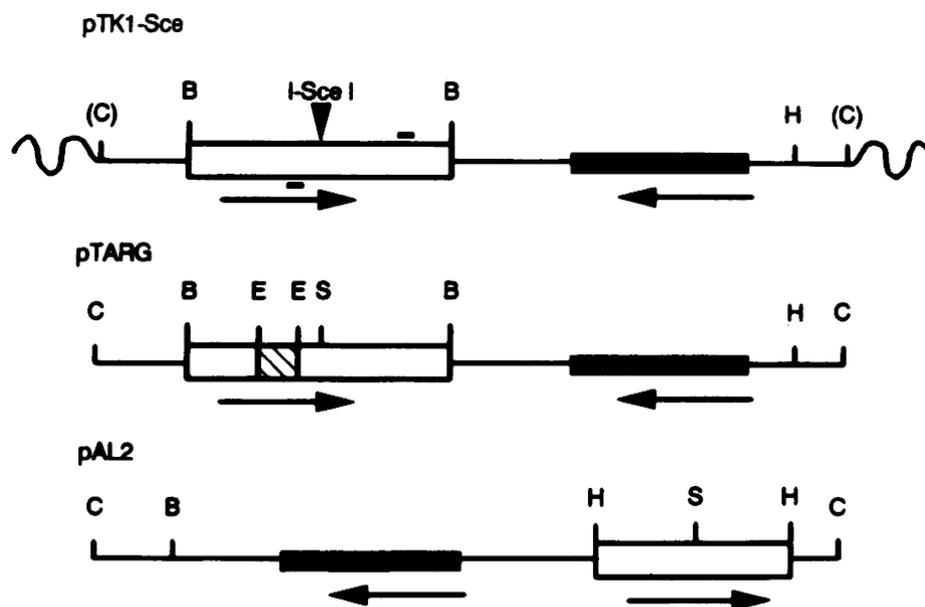
La MN I-SceI a d'abord été utilisée chez la levure pour mieux comprendre les mécanismes de réparation-recombinaison. Par la suite, des chercheurs ont voulu utiliser I-SceI pour augmenter la fréquence de la RH chez les Mammifères. Les premières expériences ont été réalisées sur des cellules fibroblastiques murines. Rouet et collaborateurs ont publié deux articles sur ce sujet en 1994. Dans leur première étude, ils ont construit un vecteur d'expression de la MN I-SceI portant une séquence de localisation nucléaire (*nls*), induit l'expression d'I-SceI dans des cellules *in vivo* et observé l'effet des DSB induits sur du matériel génétique extrachromosomique. Ils ont également testé l'hypothèse selon laquelle une expression constitutive d'I-SceI avait un effet délétère sur les cellules murines (Rouet, *et al.*, 1994b). Ils ont conclu des résultats obtenus que l'expression d'I-SceI induisait un DSB dans les cellules *in vivo* et favorisait des recombinaisons homologues extrachromosomiques. Leurs données montraient que l'expression constitutive d'I-SceI n'était pas toxique pour des cellules fibroblastiques murines. La forte spécificité d'I-SceI et l'absence de son site de reconnaissance dans le génome de la souris en étaient très probablement les raisons.

Dans une seconde étude, grâce au vecteur précédemment construit, l'équipe de Maria Jasin a cherché à induire des DSB directement dans le génome de la souris et à étudier les phénomènes de réparation qui suivaient les cassures (Rouet, *et al.*, 1994a). Pour cela, ils ont introduit dans le génome de fibroblastes 3T3 un gène de résistance à la néomycine (*néo*)

modifié. Ce gène de résistance portait un site I-SceI inséré dans sa phase de lecture, ce qui rendait le gène *neo* non fonctionnel. Ils ont sélectionné un clone (S1E) caractérisé par l'intégration à un site unique de deux copies du gène *neo* en position tête-queue. Ce clone contenait donc deux sites I-SceI insérés dans son génome. Ils ont étudié les mécanismes de réparation qui résultaient des DSB induits à ces deux sites, en présence ou non d'un vecteur de réparation de la cassure. Environ 12% des cellules transfectées avec le vecteur d'expression d'I-SceI portaient au moins une coupure sur l'un des deux sites et 70% d'entre elles portaient des coupures aux deux sites. L'analyse des clones obtenus après transfection a démontré l'existence de plusieurs mécanismes de réparation d'un DSB dans les cellules de mammifères. Le premier mécanisme décrit était une réparation par religation des extrémités, aussi appelée NHEJ (*non-homologous end joining*). Les clones réparés par NHEJ étaient caractérisés par la délétion de 1 à 22 nt et la religation des bouts francs ainsi générés. Le deuxième mécanisme observé était une réparation du DSB grâce à de très courtes séquences d'homologie intrachromosomique. Ce mécanisme était appelé SSA (pour *Single-Strand Annealing*). Selon leur hypothèse, cet événement était précédé d'une activité exonucléasique qui démasquait les séquences homologues. Enfin, le dernier mécanisme mis en jeu était le ciblage de gène par RH avec un vecteur de réparation, électroporé sous forme linéaire en même temps que le vecteur d'expression d'I-SceI. Le modèle de réparation d'un DSB par RH, aussi appelé *gap-repair*, a été proposé par J. W. Szostak (Szostak, *et al.*, 1983). En conclusion, Rouet et collaborateurs ont démontré que l'induction d'un DSB stimulait le ciblage de gène d'un facteur 100 dans leur système. En l'absence de DSB, le ciblage n'était pas détectable.

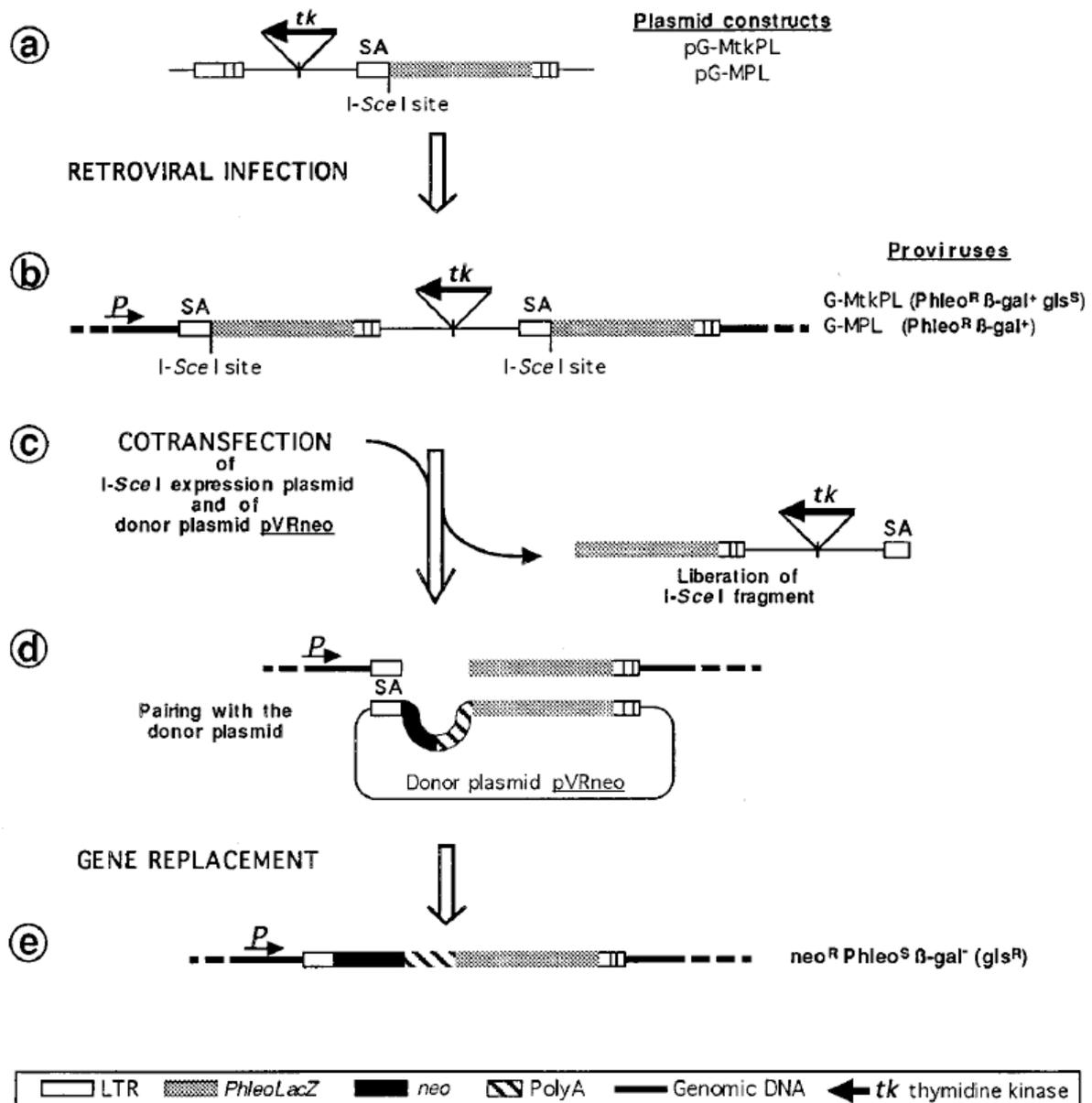
Une étude sur la même problématique a été publiée en 1994 par une autre équipe. Ce groupe a cherché à déterminer quel mécanisme utilisait préférentiellement une cellule mammifère lorsqu'elle était confrontée à un DSB. Les deux mécanismes qui ont été comparés étaient la réparation par RH avec une molécule exogène et le NHEJ (Lukacsovich, *et al.*, 1994). Pour cela, Lukacsovich et collaborateurs ont introduit, dans le génome de fibroblastes murins *Ltk<sup>-</sup>* (déficients en thymidine kinase), le gène de la thymidine kinase (*tk*) interrompu par un site I-SceI inséré dans sa phase de lecture, ce qui rendait le transgène non fonctionnel. Ils ont sélectionné un clone portant une copie, intégrée de façon stable, du transgène *tk* non fonctionnel. Ils ont ensuite introduit dans ces cellules la protéine I-SceI en présence ou non d'un vecteur de réparation. Deux vecteurs de réparation différents ont été testés. Chacun

d'eux possédait des régions d'homologie avec le gène *tk*, qui permettaient de restaurer la fonctionnalité du transgène par RH (**Figure 10**). Leurs résultats montraient que la MN induisait effectivement un DSB au niveau de son site de reconnaissance. Contrairement aux résultats obtenus par Rouet et collaborateurs, la réparation du DSB se faisait très majoritairement par NHEJ plutôt que par RH, même en présence d'un vecteur de réparation.



**Figure 10 : Structure du transgène *tk* non fonctionnel et des vecteurs de réparation**  
(Lukacsovich, *et al.*, 1994)

La construction pTK1-Sce est représentée telle qu'intégrée dans le génome de cellules de fibroblastes murins Ltk<sup>-</sup>. Les séquences génomiques flanquant le transgène sont figurées par des lignes en vague. Les vecteurs pTARG et pAL2 sont des vecteurs de réparation du transgène. Les séquences contenant le gène *tk* sont représentées par des rectangles blancs et le gène *Neo* est illustré par un rectangle noir. Le site de reconnaissance d'I-SceI est indiqué par la pointe d'une flèche dans le gène *tk*. Les gènes *tk* des vecteurs pTARG et pAL2 sont également non fonctionnels, mais permettent la réparation du gène *tk* au niveau du site I-SceI. Dans le vecteur pTARG, une séquence de 104 pb du gène *tk*, représentée par un segment haché, est inversée. Dans le vecteur pAL2, le gène *tk* n'est pas entier. Les constructions sont illustrées sous forme linéaire, avec le vecteur coupé au site unique *Cla*I (indiqué par un C).



**Figure 11 : Mécanisme de RH induite par I-SceI**  
 (Choulika, *et al.*, 1995)

- A-** Carte du rétrovirus *tk-Phleo-lacZ* (G-MtkPL) de 7,5 kb et du rétrovirus *Phleo-lacZ* (G-MPL) de 6 kb. Le rétrovirus G-MtkPL est constitué d'un gène de fusion *Phleo-lacZ* qui permet la sélection positive des cellules infectées dans un milieu avec phléomycine (représenté par un rectangle gris) et un gène *tk* de sélection négative dans un milieu avec gancyclovir (représenté par une flèche noire). Le rétrovirus G-MPL est constitué uniquement d'un gène de fusion *Phleo-lacZ*.
- B-** Carte des provirus G-MtkPL et G-MPL intégrés par infection virale. Les séquences *I-SceI*, *Phleo-lacZ* et LTR sont en duplicata. Le vecteur viral est transcrit par un promoteur cellulaire *P*.
- C-** *I-SceI* crée deux DSB dans l'ADN hôte, ce qui libère le segment central de la construction.
- D-** Ainsi les extrémités chromosomiques cassées sont disponibles pour s'apparier avec le plasmide donneur pVRneo.

E- Structure attendue du locus après RH.

D'après une quatrième étude, l'induction d'un DSB par I-SceI favorisait la réparation par RH avec un vecteur de réparation sous forme circulaire (Choulika, *et al.*, 1995). Deux provirus, qui contenaient chacun un LTR 5', ont été introduit dans le génome de fibroblastes 3T3 et de cellules multipotentes de souris PCC7-S. Chaque LTR 5' portait un site I-SceI. Les cellules modifiées ont été électroporées avec un vecteur de réparation avec ou sans un vecteur d'expression de la MN. Le vecteur de réparation était constitué de régions homologues aux séquences qui entouraient la cassure et d'un gène de sélection à la néomycine dépourvu de promoteur (**Figure 11**). Les résultats obtenus confirmaient que I-SceI induisait un DSB et démontraient que le mode de réparation favorisé en présence du vecteur de ciblage était la RH. Choulika et collaborateurs ont conclu que la fréquence de RH induite par I-SceI était au moins 100 fois supérieure à celle d'une RH spontanée, dont le seuil n'avait pas été détecté. Ces auteurs ont également démontré l'efficacité du système de réparation par RH à différents locus et dans deux types cellulaires distincts : les fibroblastes murins 3T3 et les cellules embryonnaires de carcinomes de souris PCC7-S, cellules souches multipotentes.

De ces premières expériences, nous pouvons extraire plusieurs informations importantes :

- I-SceI reconnaît son site de coupure inséré dans le génome d'une cellule de souris et elle y induit un DSB ;
- l'expression de la MN ne semble pas avoir d'effet létal sur les cellules somatiques ;
- l'induction d'un DSB dans le génome d'une cellule de mammifère favorise la réparation à ce point de cassure ;
- trois mécanismes de réparation sont observés après induction d'un DSB : la recombinaison homologue, le NHEJ et le SSA.

Soulignons que les premières équipes travaillant sur le ciblage de gène favorisé par un DSB ont des résultats et des interprétations différentes quant au mécanisme majoritairement actif dans les cellules fibroblastiques murines. Bien que les résultats de T. Lukacsovich démontrent le contraire, on considère aujourd'hui qu'un DSB favorise principalement une réparation par RH en présence d'un vecteur de réparation. Ces travaux ont amené de nombreuses équipes à utiliser cette méthode pour favoriser le ciblage de gène dans les cellules de mammifères.

## 2.4 Conception de MN à façon

Pour favoriser le ciblage de gène à un locus particulier du génome, il est nécessaire d'insérer au préalable le site I-SceI à ce locus. Cette étape est limitante car l'expérimentateur doit réaliser une RH par une technique conventionnelle. Comme nous l'avons vu, la fréquence de la RH est faible par la technique conventionnelle et elle dépend du locus ciblé. Des laboratoires ont donc modifié la structure de certaines MN, pour qu'elles reconnaissent des séquences génomiques qui divergent du site de reconnaissance de la MN d'origine. Ces nouvelles MN sont appelées « MN à façon ».

L'enzyme I-CreI de *Chlamydomonas reinhardtii* est l'endonucléase la plus utilisée pour développer des MN à façon. Il s'agit d'une enzyme intronique de groupe I, comme I-SceI. Elle appartient aussi à la famille LAGLIDADG et est caractérisée par la présence de ce motif en un seul exemplaire. Cette MN agit sous forme d'homodimère. I-CreI reconnaît une séquence de 22 pb pseudo-palindromique. L'analyse de la structure de l'enzyme I-CreI liée à sa cible a montré que 8 résidus de chaque monomère établissent des interactions directes avec 7 bases du site de reconnaissance (Jurica, *et al.*, 1998). I-CreI a été choisie comme modèle en raison de la stabilité de son activité dans des conditions variables de température et de la bonne caractérisation de son interaction avec sa cible. I-DmoI est aussi une protéine de la famille LAGLIDADG et est caractérisé par la présence de ce motif en deux exemplaires, comme I-SceI (Dalgaard, *et al.*, 1993). Les premières MN chimères ont été créées par fusion de domaines de I-CreI et de I-DmoI. Ces MN chimères sont fonctionnelles. Elles clivent la séquence hybride avec une grande spécificité (Chevalier, *et al.*, 2002; Epinat, *et al.*, 2003). Ces résultats ont ouvert la voie vers la production des MN à façon.

Les MN à façon sont développées pour reconnaître des séquences cibles dans le génome des mammifères, dans le but de modifier des locus chromosomiques. L'entreprise Collectis a été pionnière dans la production et l'étude des MN à façon. Dans une première étude, Arnould et collaborateurs ont développé une approche semi-aléatoire qui permet de dériver des centaines de nouvelles protéines à partir de I-CreI. Leur activité de coupure était ensuite évaluée par un dosage à haut-débit sur le génome de cellules CHO (*Chinese Hamster Ovary*) (Arnould, *et al.*, 2006). La première MN à façon capable de cibler spécifiquement un gène de mammifère a été synthétisée pour cliver le gène *XPC* (*Xeroderma Pigmentosum C*) de

l'Homme (Arnould, *et al.*, 2007). Arnould et collaborateurs ont d'abord cherché des cibles potentielles dans le gène *XPC* et choisi trois séquences qui diffèrent totalement de la séquence de coupure de I-CreI. Ils ont utilisé l'approche semi-aléatoire pour modifier de façon localisée la spécificité de I-CreI, suivie d'une stratégie combinatoire pour assembler les différents fragments mutés. Ainsi, ils ont créé des MN à façon qui ont une bonne spécificité pour leur site de clivage prédit, sans génotoxicité apparente (Redondo, *et al.*, 2008).

Toujours à partir de I-CreI, une MN a été développée selon la même stratégie pour cibler le gène *RAG1* de l'Homme (Grizot, *et al.*, 2009). Cette MN à façon a permis de favoriser une recombinaison par ciblage de gène au locus *RAG1*, jusqu'à atteindre une fréquence de recombinaison de 6% des cellules humaines transfectées. Les différentes applications possibles de cette approche sont présentées dans une revue récente de Arnould et collaborateurs (Arnould, *et al.*, 2011). Très récemment, l'équipe de Ryo Takeuchi a cherché à utiliser d'autres MN que I-CreI. Ils ont réalisé une analyse phylogénétique sur plus de 200 MN naturelles de la famille LAGLIDADG (Takeuchi, *et al.*, 2011). Ils ont démontré que ses membres possèdent des séquences d'acides aminés très proches et qu'ils ciblent une grande diversité de séquences ADN. Ils considèrent que la découverte de cette importante collection de MN devrait faciliter la fabrication de MN à façon pour cibler de nombreux gènes. Pour illustrer l'intérêt de ces nouvelles MN, ils ont développé une MN à façon à partir de l'enzyme I-OnuI pour cibler le gène *MAO-B* humain.

Le développement de MN à façon a fait la preuve de son efficacité dans les études présentées ci-dessus. Ces méthodes restent peu utilisées. Une raison est que les domaines responsables de la reconnaissance de l'ADN et de la cassure double brin sont très liés, ce qui rend difficile la modification de l'un sans affecter l'autre. Nous verrons que ce n'est pas le cas des nucléases à doigts de zinc, qui comprennent deux domaines distincts responsables de la reconnaissance de la séquence cible et de l'activité endonucléase, respectivement.

## 3 Les nucléases à doigts de zinc (ZFN), un outil créé pour induire des DSB

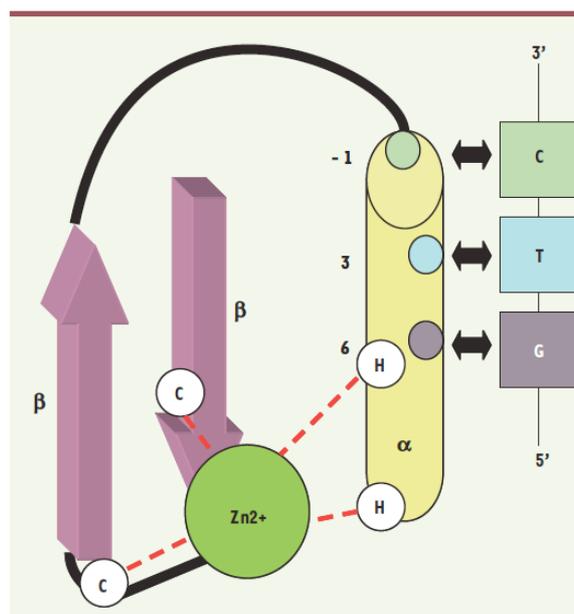
### 3.1 Couplage de protéines à doigts de zinc à l'endonucléase *FokI*

L'idée de créer des nucléases à doigt de zinc est issue des découvertes faites par l'équipe de S. Chandrasegaran sur les domaines fonctionnels de l'endonucléase *FokI* (Li, *et al.*, 1992). La protéine *FokI* de *Flavobacterium okeanoicoites* est une enzyme de restriction de classe IIS. Contrairement aux autres enzymes de type II dont le site de clivage est compris dans le site de reconnaissance ou est adjacent à ce site, les enzymes de classe IIS coupent l'ADN double-brin (db) à une certaine distance de leur site de reconnaissance. Li et collaborateurs ont séparé et identifié, par digestion à la trypsine, les domaines de reconnaissance et de clivage de l'endonucléase de restriction *FokI*. Ils ont ainsi montré qu'en présence d'un substrat ADN, *FokI* comprend deux domaines de 41 kDa et 25 kDa qui correspondent au domaine de liaison/reconnaissance de l'ADN et au domaine de clivage, respectivement. Cette importante découverte documente une séparation physique entre les activités de liaison et de clivage de cette enzyme.

L'intérêt principal de cette observation réside dans le fait que le domaine de clivage de *FokI* n'a aucune spécificité de séquence apparente. L'équipe de S. Chandrasegaran a publié, à la suite de ce travail, deux articles dans lesquels les auteurs décrivent la fabrication d'endonucléases « artificielles ». Le principe était de substituer le domaine naturel de reconnaissance de *FokI* par des domaines de reconnaissance alternatifs (Kim and Chandrasegaran, 1994; Kim, *et al.*, 1996). Ils ont d'abord couplé le domaine de clivage de *FokI* à l'homéodomaine Ubx de la drosophile. Ils ont démontré que l'endonucléase chimère a la même préférence de liaison à l'ADN que la protéine Ubx et qu'elle coupe l'ADN db à une distance fixe du site de liaison (Kim and Chandrasegaran, 1994). Ils ont ensuite rapporté la synthèse de deux endonucléases dites « hybrides ». Elles ont été élaborées par liaison de protéines à doigts de zinc (ZF pour *zinc fingers*) de type Cys<sub>2</sub>His<sub>2</sub> au domaine de clivage de *FokI* (Kim, *et al.*, 1996). Cette dernière tentative était la plus prometteuse car elle marquait l'avènement du développement des nucléases à doigts de zinc, ou ZFN (pour *zinc finger nucleases*).

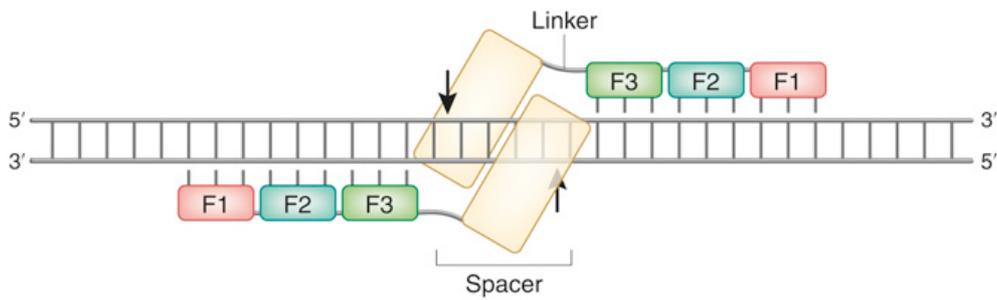
### 3.2 Spécificité des protéines à doigts de zinc et mécanisme d'action des ZFN

Le motif protéique ZF de type Cys<sub>2</sub>His<sub>2</sub> est constitué d'environ 30 acides aminés organisés en une structure compacte ββα qui est stabilisée par un ion zinc. L'atome de zinc est coordonné par les cystéines des feuilletts β et les histidines de l'hélice α, acides aminés qui sont conservés chez les protéines ZF. Un motif ZF reconnaît spécifiquement 3pb contiguës (Pavletich and Pabo, 1991). Les acides aminés en position -1, +3 et +6 de l'hélice α du motif ZF interagissent avec un des brins de l'ADN (**Figure 12**). L'acide aminé en position +2 de l'hélice α est important pour la spécificité de reconnaissance de l'ADN par le motif ZF, en permettant une interaction avec le brin d'ADN complémentaire (Isalan, *et al.*, 1998).



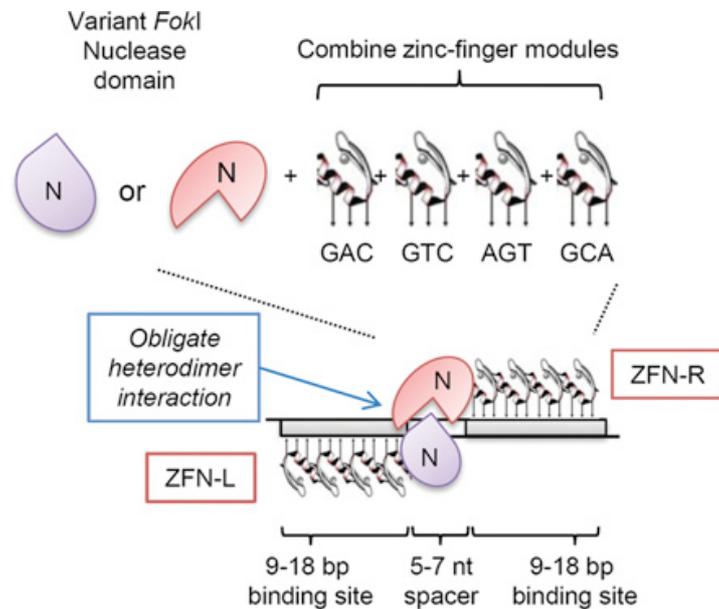
**Figure 12 : Schéma d'un motif ZF**  
(Dion, *et al.*, 2007)

Le motif est représenté avec les acides aminés jouant un rôle clé dans l'interaction avec le triplet de nucléotides. Le résidu en position +6 dans l'hélice α du motif ZF interagit avec la base en 5' du triplet d'ADN reconnu, le résidu en +3 est en contact avec la base centrale du triplet et le résidu en position -1 interagit avec la base en 3'.



**Figure 13 : Schéma d'une paire de ZFN liée à l'ADN**  
(Carroll, 2011)

Chaque protéine à doigts de zinc est représentée par un rectangle de couleur ; des petites lignes verticales indiquent les liaisons majeures avec les bases de l'ADN. Les domaines de clivage de *FokI* sont représentés par de grands rectangles semi-transparents, avec des sites de clivage non-spécifique, séparés par 4 pb et indiqués par des flèches verticales. La protéine F1 se trouve en N-terminal de l'ensemble de la ZFN. La séquence de liaison (« linker ») entre les domaines de liaison et de clivage de la protéine est indiquée. L'espace entre les domaines de liaison de chacune des ZFN (« spacer »), de 6 pb dans ce cas, est aussi indiqué.



**Figure 14 : Schéma d'une paire de ZFN constituée de domaines de clivage FokI modifiés**  
(Geurts and Moreno, 2010)

ZFN-R représente la ZFN de droite et ZFN-L, celle de gauche. Chaque ZFN peut être constituée de 3 jusqu'à 6 protéines à doigts de zinc ou motifs, qui se lient à 9 jusqu'à 18 pb respectivement. Ici 4 motifs sont représentés. L'espace entre les domaines de liaison fait entre 5 et 7 pb. Des variantes du domaine *FokI* sont utilisées pour obliger la formation d'un hétérodimère.

et le clivage. Un lot associé à un domaine de clivage *FokI* correspond à ce qu'on appelle une nucléase à doigts de zinc (ZFN). Plusieurs études ont montré que la configuration optimale nécessite de synthétiser une séquence de liaison (appelé « linker ») entre les deux domaines de la protéine ZFN (de clivage et de liaison à l'ADN) et de laisser un espace de 5 à 7 pb (appelé « spacer ») entre les sites de reconnaissance des deux lots de motifs ZF, qui sont dans des orientations inversées (**Figure 13**). Le domaine de clivage *FokI* ne fonctionne que sous forme de dimère. La cassure ne peut donc pas se produire aux sites de liaison d'un seul des deux lots de motifs ZF. Ainsi, l'ensemble de la protéine ZFN n'est assemblé qu'au niveau de la séquence cible, si les motifs ZF sont suffisamment spécifiques. La nécessité de combiner deux lots de liaison à l'ADN permet d'augmenter notablement la taille du site de reconnaissance. Par exemple, deux lots de 3 motifs ZF permettent de cibler une séquence de 18 pb, ce qui est normalement suffisant pour reconnaître une séquence unique dans le génome des Mammifères (pour revue voir (Carroll, 2011)). Chacun des motifs ne contribue pas de façon égale à l'affinité de la ZFN pour sa séquence. Les publications font état de ZFN contenant jusqu'à 6 motifs dans chaque lot. Des équipes ont également créé des variants du domaine de clivage *FokI* pour obliger la formation d'un hétérodimère, cela dans le but d'augmenter la spécificité du système ZFN (**Figure 14**) (Miller, *et al.*, 2007; Szczepek, *et al.*, 2007).

### 3.3 Premiers ciblage de gènes à l'aide de ZFN

Les capacités de clivage des premières ZFN ont été testées *in vitro* (Kim, *et al.*, 1996). Par la suite, une étude a été réalisée pour tester l'aptitude de l'enzyme *FokI* (d'origine procaryote) à agir *in vivo* sur une séquence d'ADN exogène, introduite dans la chromatine de l'hôte. Pour démontrer cela, une séquence cible exogène ainsi que la ZFN correspondante ont été introduites dans un oocyte de xénope (Bibikova, *et al.*, 2001). Ils ont injecté la séquence cible exogène dans l'oocyte et attendu 3 à 4 heures jusqu'à ce que celle-ci s'insère dans la chromatine. Ils ont ensuite introduit la ZFN directement dans le noyau de l'oocyte. Ils ont observé que la coupure de l'ADN db avait bien lieu et que cela nécessitait la formation d'un dimère du domaine de clivage de *FokI*. Ils ont également démontré que l'induction d'un DSB favorise la recombinaison homologue.

La même équipe a conçu la première ZFN capable de cibler un gène endogène chez la drosophile (Bibikova, *et al.*, 2002). Dans cette étude, ils ont développé une ZFN qui reconnaît

un site unique dans le gène *yellow* (*y*) de la drosophile. Après expression de la ZFN dans la larve, ils ont observé les mutations au locus *y* dans les cellules somatiques. Par séquençage de l'ADN, ils ont montré que les modifications correspondaient à de courtes délétions ou insertions au niveau de la séquence cible. Ce type de modification est typique d'un mécanisme de réparation par NHEJ à la suite d'un DSB. Il restait à montrer qu'une ZFN est capable de favoriser la RH lorsqu'elle induit un DSB à un locus endogène de la drosophile. L'année suivante, Bibikova et collaborateurs ont publié le premier ciblage de gène favorisé par une ZFN en présence d'un vecteur de réparation (Bibikova, *et al.*, 2003) Celui-ci possédait des régions d'homologie au gène interrompu par la cassure, ce qui favorisait la réparation par RH.

Le premier ciblage de gène favorisé par une ZFN dans des cellules de mammifères a été réalisé par M.H. Porteus et collaborateurs dans des cellules somatiques humaines (Porteus and Baltimore, 2003). Ces auteurs ont d'abord introduit le gène codant la GFP (*Green Fluorescent Protein*) sous une forme non fonctionnelle dans le génome de cellules somatiques humaines 293. La séquence codante de la GFP était en effet interrompue, soit par un codon stop et un site *I-SceI*, soit par un codon stop, un site *I-SceI* et une séquence de reconnaissance d'une ZFN. Ils ont ensuite mesuré la fréquence de réparation par recombinaison homologue dans différentes situations de DSB : co-transfection avec un vecteur d'expression d'*I-SceI* et un vecteur de réparation ou co-transfection avec une ZFN et un vecteur de réparation. Ils ont montré que dans des conditions optimales de ciblage de gène avec une ZFN, ils obtenaient la même efficacité de recombinaison homologue que lors d'un ciblage de gène favorisé par *I-SceI*. Ce travail a ouvert la voie vers les vastes potentialités des ZFN pour manipuler le génome des cellules somatiques de mammifère.

*IL2R $\gamma$*  est le premier gène endogène humain ciblé grâce à une ZFN (Urnov, *et al.*, 2005). F.D. Urnov et collaborateurs ont développé une ZFN ayant pour cible la mutation SCID au locus *IL2R $\gamma$*  sur le chromosome X. La fréquence de réparation par recombinaison homologue était égale à environ 18% dans des lignées cellulaires établies ou dans des cultures primaires. Cette étude illustre les possibilités offertes par les ZFN dans le domaine de la thérapie génique.

### 3.4 Méthodes de développement et de délivrance des ZFN

Les ZFN sont développées par Sangamo BioSciences. Le client leur fournit en premier lieu la séquence du gène d'intérêt. Sangamo BioSciences utilise des outils informatiques pour caractériser des séquences exoniques cibles. L'idéal est d'identifier des séquences cibles dans les premiers 2/3 de la séquence codante du gène pour abolir son expression de façon efficace. Sangamo BioSciences utilise des banques de 2 motifs ZFs dont la séquence cible est bien caractérisée. Les chercheurs de Sangamo BioSciences fabriquent alors jusqu'à 16 paires de ZFN et effectuent des tests dans des systèmes cellulaires homologues à celui ciblé par le client (dans le cas du rat, ils transfectent les ZFN dans des cellules de rat). Une paire de ZFN caractérisée et validée dans un système cellulaire est envoyée au client. Le coût pour le laboratoire est de 25.000 \$ (Geurts and Moreno, 2010).

Un laboratoire peut développer sa propre ZFN sans faire appel à Sangamo BioSciences. Pour cela, des laboratoires académiques ont mis à disposition des chercheurs des ressources informatiques qui répertorient les motifs ZF déjà caractérisés. Pour élaborer une ZFN, une des méthodes possibles est de traiter chaque motif ZF comme un module indépendant et de combiner les différents motifs pour cibler une séquence. Même si elle est rapide et accessible à tous, cette méthode ne prend pas en compte les interactions possibles entre les motifs ZF adjacents et l'effet du contexte chromatinien. Par ailleurs, les membres du « Zinc Finger Consortium », un groupe de laboratoires académiques, ont créé la méthode OPEN (*Oligomerized Pooled ENgineering*), disponible aux laboratoires publiques, qui permet le développement de ZFN personnalisées en combinant un système double-hybride en bactérie et des tests en culture cellulaire (Maeder, *et al.*, 2008). Cette méthode a permis d'augmenter le nombre de ZFN actives, capables de cibler une grande diversité de séquences (Foley, *et al.*, 2009).

Plusieurs méthodes ont été développées pour délivrer des ZFN. Dans les cellules en culture, les expérimentateurs introduisent une séquence codant pour la ZFN, accompagnée d'un promoteur spécifique du type cellulaire, soit par transfection d'ADN, soit par infection par un virus recombinant. La même méthode est utilisée pour introduire la séquence de réparation quand il s'agit du ciblage d'un gène. En 2008, l'équipe de D. Carroll a fait une avancée considérable dans ce domaine. Beumer et collaborateurs ont en effet montré qu'une injection, directement dans l'embryon de drosophile, de la ZFN sous la forme d'un ARN messager, accompagnée de la séquence de réparation si besoin, permet une bonne efficacité

de réparation homologue et non-homologue (Beumer, *et al.*, 2008). L'injection directe d'ARNm codant la ZFN s'est avérée un outil très intéressant dans bien d'autres organismes. Chez les Mammifères, cette méthode a été testée dans l'embryon de rat (Geurts, *et al.*, 2009; Mashimo, *et al.*, 2010) et de souris (Carbery, *et al.*, 2010; Meyer, *et al.*, 2010). L'ARNm de la ZFN est injecté dans le pronucléus mâle de l'œuf fécondé pour être traduit dans le cytoplasme en protéine ZFN. Celle-ci pénètre dans le noyau pour couper l'ADN db au niveau de sa cible. Les premières expériences ont montré que le ciblage a lieu au stade une cellule, ce qui aboutit à une modification portée par toutes les cellules de l'organisme. Chez le rat et la souris, le premier ciblage de gène favorisé par une ZFN, injectée directement dans l'embryon sous forme d'ARNm, a été réalisé avec succès récemment (Cui, *et al.*, 2011; Meyer, *et al.*, 2010). Les animaux adultes produits sont viables. Ils portent la mutation dans leur lignée germinale.

Les MN et ZFN sont les deux principaux outils qui ont été développés ces 15 dernières années pour induire des DSB dans les cellules de mammifères. L'intérêt de ces outils réside dans la maîtrise des réparations qui suivent le DSB. Nous allons maintenant étudier quels sont les différents mécanismes de réparation possibles après un DSB chez les Mammifères.

## **4 Réparation de DSB dans les cellules de mammifères**

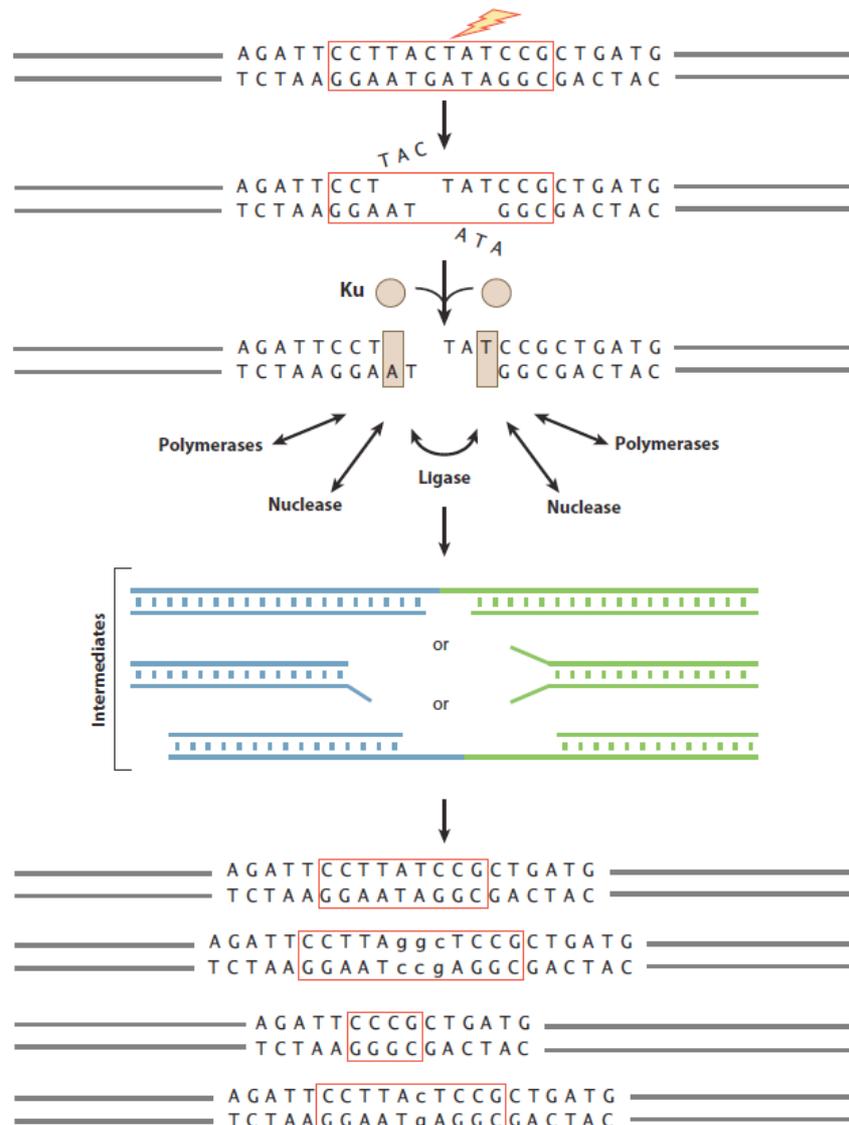
### **4.1 Les différents mécanismes de réparation induits par un DSB**

Suite à l'induction d'un DSB dans des cellules de mammifères, les trois principales voies de réparation possibles sont le NHEJ (*non-homologous end joining*), la recombinaison homologue et le SSA (*single-strand annealing*).

#### **4.1.1 *La réparation par NHEJ (Non-Homologous End Joining)***

Le mécanisme de NHEJ est fonctionnel pendant toutes les phases du cycle cellulaire. Il est cependant prédominant pendant la phase G1 et en début de phase S. Il favorise la religation des extrémités libres de l'ADN, créées par un DSB, et implique un traitement minimal de la cassure. La première protéine à intervenir dans la voie du NHEJ est l'hétérodimère Ku70/80 (Ku) qui se lie aux extrémités de l'ADN. La protéine Ku forme un complexe avec l'ADN, Ku:ADN, probablement à chacune des extrémités coupées. Ce complexe Ku:ADN peut être considéré comme un nœud sur lequel différentes protéines du

mécanisme de NHEJ s'accrochent pour favoriser la réparation du DSB. Les différentes protéines recrutées par le complexe Ku:ADN, comme les nucléases, polymérase et ligases, peuvent être recrutées dans n'importe quel ordre (**Figure 15**) (pour revue voir (Lieber, 2010)).



**Figure 15 : Les principales étapes du mécanisme de NHEJ**  
(Lieber, 2010)

La flèche en forme d'éclair illustre un événement de DSB, comme des radiations ionisantes ou une activité enzymatique. La liaison de la protéine Ku avec les extrémités du DSB favorise la fixation des autres composants du NHEJ : nucléases, polymérase, ligases. La fixation de Ku à l'ADN entraîne un changement de conformation de la protéine, figuré par une transformation d'une forme sphérique à une forme rectangulaire. La flexibilité du mécanisme de NHEJ est permise par une flexibilité dans l'ordre de recrutement des composants enzymatiques et par l'indépendance de chaque extrémité. Ainsi, le processus de NHEJ peut réparer diverses configurations de DSB et religuer les extrémités. Ce processus passe par la formation de plusieurs types d'intermédiaires (au centre de la figure), qui aboutissent à une diversité de séquences de réparation (en bas de la figure).

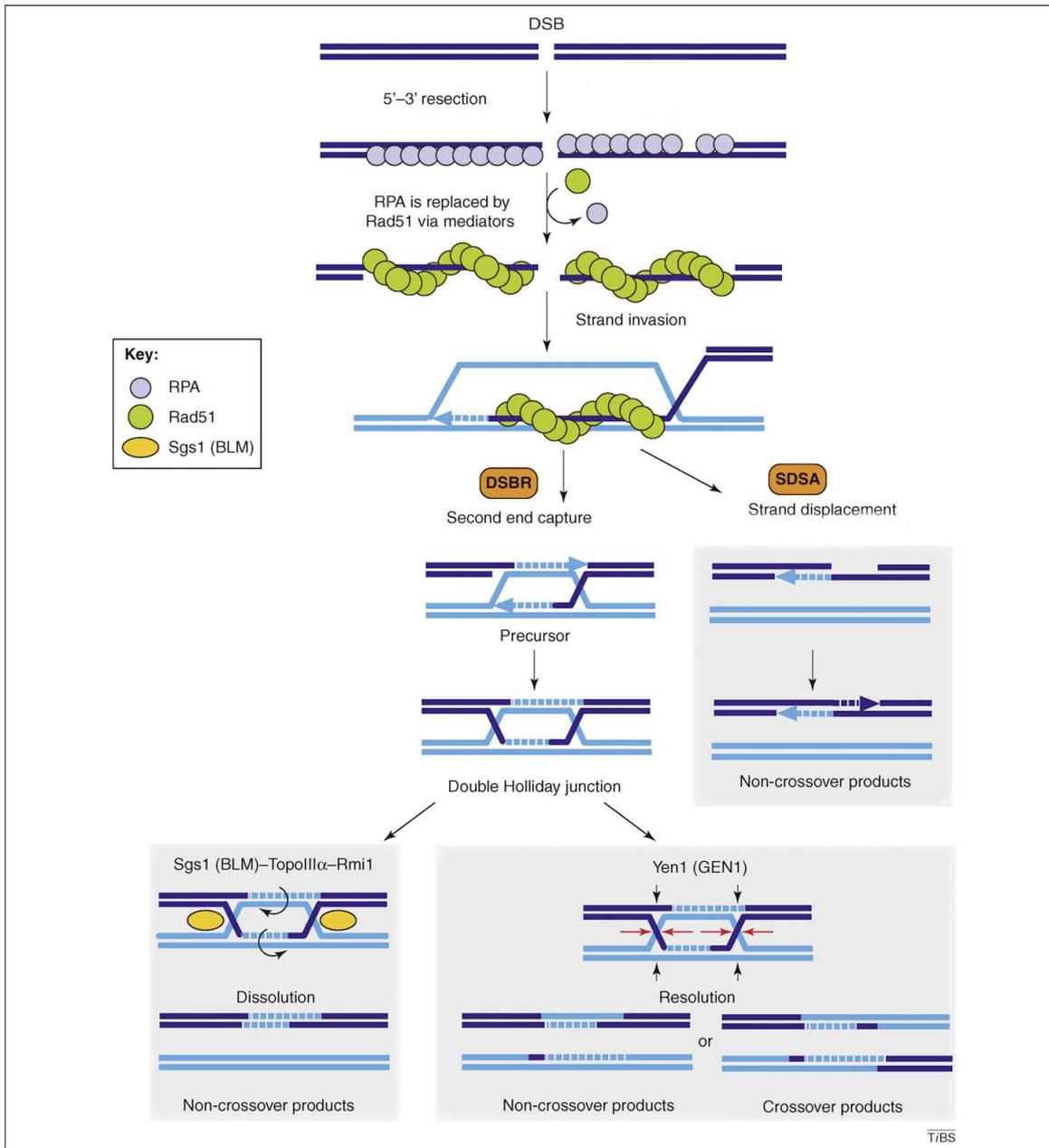


Un DSB peut créer différentes structures au niveau des extrémités de l'ADN. Pour cette raison, le NHEJ passe, si nécessaire, par des étapes intermédiaires, comme le clivage ou l'élongation de brin, qui précèdent la religation des extrémités. Dans les cellules de mammifères, Ku interagit avec la sous-unité catalytique de la protéine kinase dépendante de l'ADN (DNA-PKcs). La nucléase Artemis est recrutée au site de cassure via son interaction avec la protéine DNA-PKc. L'ensemble forme un complexe Artemis/DNA-PKcs capable de cliver une grande variété d'extrémités simple-brin endommagées. Ces clivages peuvent être à l'origine de brèches dans l'ADN, qu'il est alors nécessaire de compléter via des polymérase impliquées dans la voie de réparation par NHEJ. La famille PolX comprend les polymérase  $\mu$  et  $\lambda$  qui interagissent avec le complexe Ku:ADN. Les extrémités coupées sont enfin ressoudées grâce au complexe DNA ligase IV/XRCC4. La protéine XRCC4 n'a pas d'activité enzymatique connue. XRCC4 agit en tant que protéine de soutien, ce qui permet de favoriser les interactions avec la protéine Ku et l'ADN, mais aussi de stabiliser et de stimuler l'activité ligase de la protéine DNA ligase IV (ou Lig4) (pour revue voir (Kass and Jasin, 2010)). Les différentes modifications de l'ADN au lieu de cassure, qui précèdent la religation, donnent lieu à des délétions ou à des insertions. Cette caractéristique du NHEJ le rend plus sujet aux erreurs, comparé à la RH.

La voie de réparation par NHEJ telle qu'elle est décrite ci-dessus correspond au NHEJ dit canonique. En l'absence des facteurs spécifiques du NHEJ canonique, il existe une voie supplémentaire de réparation de type NHEJ, dite NHEJ alternatif (alt-NHEJ). Plusieurs publications ont rapporté l'intervention de différents facteurs dans le mécanisme de alt-NHEJ, comme les ligases Lig1 et Lig3 (Simsek, *et al.*, 2011), ou encore le complexe Mre11 (Dinkelmann, *et al.*, 2009; Rass, *et al.*, 2009; Xie, *et al.*, 2009). Il semble que le choix entre les mécanismes NHEJ canonique et alternatif dépende de la proportion de micro-homologie au niveau du site de cassure.

#### **4.1.2 La réparation par recombinaison homologue**

Le mécanisme de recombinaison homologue (RH) a spécifiquement lieu en fin de phase S et pendant la phase G2 (Saleh-Gohari and Helleday, 2004). Il utilise des séquences homologues intactes comme modèle de réparation, préférentiellement situées sur la chromatide sœur, et est considéré comme étant plus précis que le NHEJ. Le processus de réparation par RH est aussi appelé mécanisme de conversion génique. L'information d'une séquence homologue donneuse est copiée au niveau du site endommagé, celui-ci est donc le



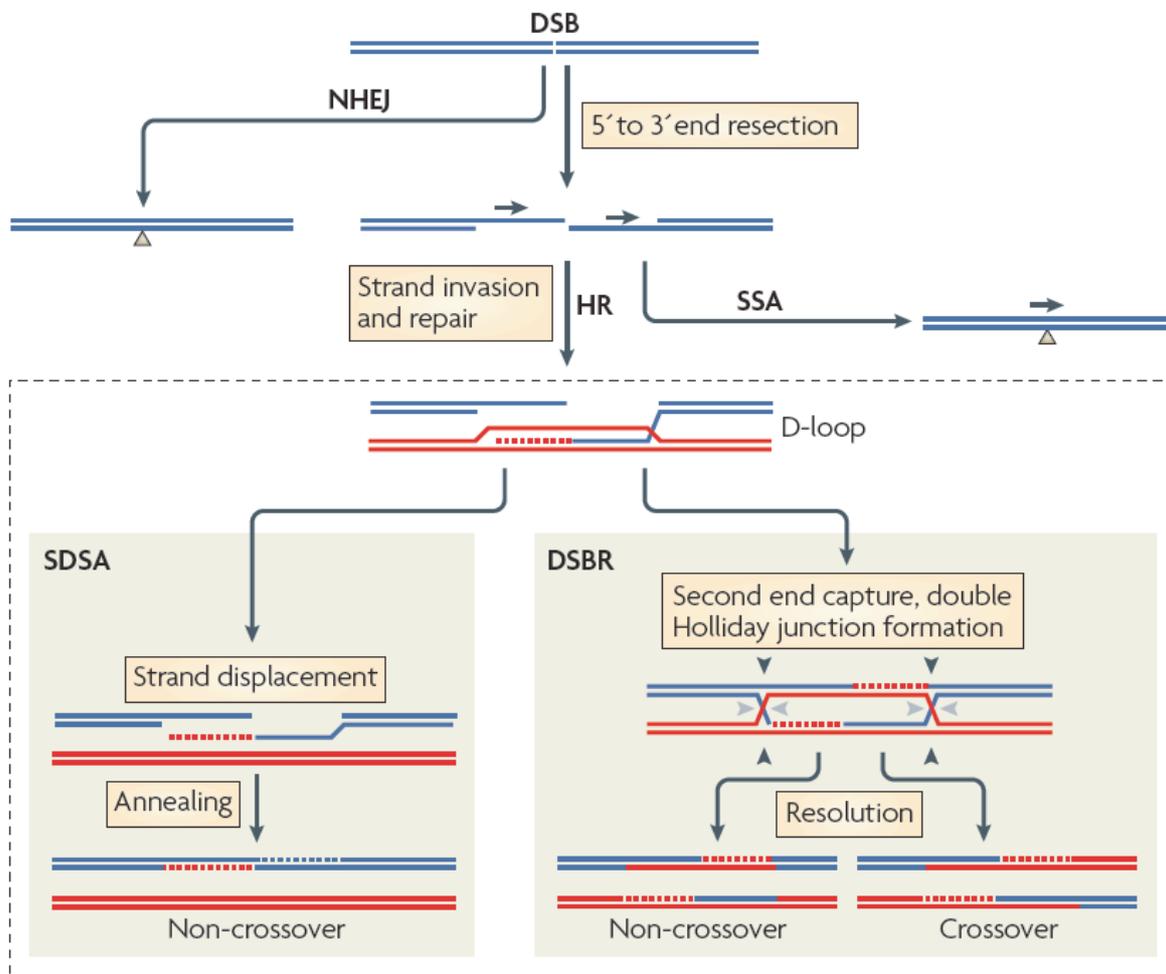
**Figure 16 : Modèles de réparation par RH**  
(d'après (Mimitou and Symington, 2009))

Les mécanismes de RH commencent par la digestion des brins cassés dans le sens 5'-3' pour donner des extrémités 3' sortantes simple-brin. Celles-ci sont rapidement recouvertes avec la protéine RPA (sphère bleue) qui est ensuite remplacée par RAD51 (sphère verte), pour former un filament de nucléoprotéines. Ce filament initie l'invasion par l'ADN simple-brin 3' sortant, de l'ADN double-brin homologue et favorise son appariement avec le brin complémentaire pour former la D-loop. L'extrémité 3' du brin invasif est prolongée par synthèse d'ADN grâce au brin homologue. Dans le modèle du SDSA, le brin 3' simple-brin prolongé se ré-apparie avec son complémentaire simple-brin, pour permettre la finalisation de la réparation par synthèse d'ADN. Dans le modèle de DSBR, l'extrémité 5' du brin cassé est capturée pour former une double jonction de Holliday. La résolution de la double jonction (via GEN1) aboutit à la formation de produits de crossover ou non-crossover et la dissolution de la double jonction (via BLM) crée un produit de non-crossover.

bénéficiaire de l'échange d'information génétique. Les deux principaux modèles de réparation par recombinaison homologue sont : la voie classique du DSBR (pour *double-strand break repair*), et la voie du SDSA (pour *synthesis-dependant strand annealing*) (pour revue voir (Moynahan and Jasin, 2010)). Le modèle de réparation par DSBR est dérivé de celui proposé par J.W. Szostak et collaborateurs en 1983, appelé mécanisme de gap-repair (Szostak, *et al.*, 1983).

Les mécanismes de DSBR et SDSA commencent tous les deux par la digestion des extrémités de l'ADN au niveau du DSB par des exonucléases, qui créent des extrémités 3' sortantes simple-brin. Il semble que la protéine BRCA1 soit impliquée dans la digestion des extrémités de l'ADN (Chen, *et al.*, 2008; Schlegel, *et al.*, 2006; Stark, *et al.*, 2004), mais son rôle est encore mal connu. Les extrémités 3' sortantes sont recouvertes par la protéine de réplication A (RPA) qui est nécessaire au recrutement des protéines du checkpoint et de la RH, comme RAD51 (pour revue voir (San Filippo, *et al.*, 2008)). RAD51 est une ATPase dépendante de l'ADN, homologue de la protéine bactérienne RecA. RAD51 forme des filaments de nucléoprotéines avec l'ADN. Dans les cellules de mammifères, la protéine BRCA2 permet le recrutement de RAD51 au niveau d'un DSB (Yuan, *et al.*, 1999). Les cellules mutantes pour le gène *BRCA2* sont déficientes pour la réparation par RH (Moynahan, *et al.*, 2001), ce qui démontre le caractère essentiel de cette protéine dans la RH. Une fois que la protéine RAD51 est recrutée au site de cassure, elle catalyse l'invasion par l'ADN simple-brin 3' sortant de l'ADN double-brin homologue, pour former la boucle de déplacement ou D-loop (*displacement loop*) (Figure 16).

La résolution de la D-loop peut se faire de différentes façons (pour revue voir (San Filippo, *et al.*, 2008)). Dans le cas du SDSA, l'extrémité 3' est prolongée par une polymérase grâce au brin homologue, puis le brin nouvellement formé se dissocie de l'ADN homologue pour se réapparier à l'extrémité 5' du brin cassé. Il ne se forme alors que des produits de type non-crossover. Dans le cas du DSBR, les extrémités 3' et 5' du brin cassé sont capturées pour former des jonctions de Holliday. La résolution des doubles jonctions de Holliday peut conduire à un produit de type crossover, ou à un produit de type non-crossover. Plusieurs protéines sont potentiellement impliquées dans la résolution des jonctions de Holliday, comme GEN1 (Ip, *et al.*, 2008). L'autre événement possible est la dissolution des jonctions de Holliday, grâce à des protéines qui interrompent la D-loop, comme BLM (Singh, *et al.*, 2008; Wu and Hickson, 2003; Xu, *et al.*, 2008) (Figure 16). Dans les cellules souches embryonnaires de souris et dans les lymphoblastes humains, le produit de non-crossover est



**Figure 17 : Les différents mécanismes de réparation d'un DSB**  
(Moynahan and Jasin, 2010)

Dans les cellules de mammifères, un DSB est efficacement réparée par le mécanisme de RH, NHEJ ou SSA. Le mécanisme de SSA se met en place lorsque la digestion des brins coupés révèle des séquences répétées (représentées par des flèches). Ces séquences s'apparient pour aboutir à un produit, contenant une seule copie de cette séquence et délété des séquences intermédiaires.

prédominant par rapport au produit de crossover (Stark and Jasin, 2003). Il semble donc que l'événement de crossing-over soit rare dans le cas d'une RH mitotique, d'autant plus qu'il peut être responsable d'effets délétères, dus à une perte d'hétérozygotie. Des protéines qui interviennent dans la dissolution des jonctions de Holliday, comme BLM, suppriment les crossover mitotiques. Cela permet de diminuer les risques d'instabilité génomique (Kass and Jasin, 2010). Le crossover joue naturellement un rôle très important dans la ségrégation des chromosomes lors de la recombinaison méiotique (Cole, *et al.*, 2010).

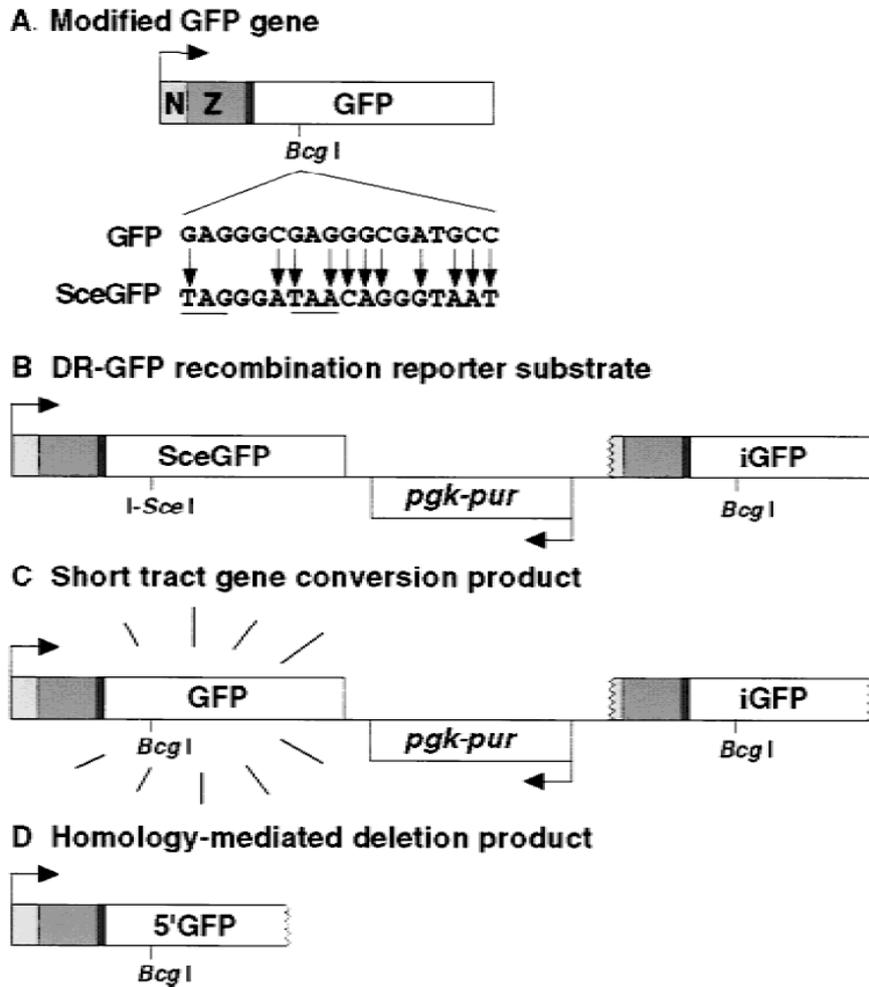
#### 4.1.3 La réparation par SSA (*single-strand annealing*)

Cette voie de réparation est distincte de la RH. Elle ne se met en place qu'en présence de séquences homologues de part et d'autre du DSB. Comme dans la RH, la réparation par SSA commence par une digestion des extrémités de l'ADN au niveau du DSB, ce qui crée des extrémités 3' sortantes simple-brin. Les séquences simple-brin complémentaires s'apparient, entraînant une perte des séquences qui se trouvaient entre les séquences répétées. Des protéines, comme RAD52, ERCC1 et Rad1/Rad10 ont été identifiées dans les cellules de mammifères comme étant impliquées dans le mécanisme de SSA (Fishman-Lobell and Haber, 1992; Mortensen, *et al.*, 1996; Stark, *et al.*, 2004). Le mécanisme de SSA est plus mutagène que la RH car il favorise une perte d'information génétique (Kass and Jasin, 2010). Cette voie de réparation est minoritaire quelle que soit la phase du cycle cellulaire au cours de laquelle le DSB se produit. Bien qu'ils soient désignés par deux termes distincts, le SSA n'est pas un mécanisme clairement différent du alt-NHEJ. Les deux processus nécessitent la présence de micro-homologies au niveau du DSB.

La **Figure 17** reflète les différentes voies qu'empruntent les systèmes de réparation d'une cellule mammifère à la suite d'une cassure double brin.

#### 4.2 Une singularité des cellules souches embryonnaires (ES) : la recombinaison homologue est le mécanisme de réparation majoritaire

La phase du cycle cellulaire pendant laquelle un DSB se produit est déterminante pour la voie de réparation empruntée. La RH est restreinte aux phases S et G2 du cycle cellulaire (Rothkamm, *et al.*, 2003). Une des raisons pouvant expliquer ce phénomène est que la chromatide sœur est présente pour servir de modèle de réparation, alors que ce n'est pas le cas



**Figure 18 : Mesure de la fréquence de réparation par RH**  
(Pierce, *et al.*, 1999)

- A-** Gène *GFP* modifié. Le gène modifié code pour la protéine EGFP fusionnée à un signal de localisation nucléaire (N) et à un domaine de doigt de zinc (Z). Ce gène est exprimé sous le contrôle de l'activateur hCMV associé au promoteur de la  $\beta$ -actine de poulet (représenté par une flèche). Le gène *GFP* est modifié en *SceGFP* de sorte qu'il contient le site de reconnaissance de la méganucléase I-SceI et une séquence de terminaison de lecture (soulignée).
- B-** Plasmide DR-GFP. Une séquence *iGFP* est située en aval du gène *SceGFP*. *iGFP* contient une séquence du gène *GFP* tronquée des parties 5' et 3'.
- C-** Produit de conversion génique. Un DSB au site I-SceI est réparé grâce à la séquence *iGFP* située sur la même chromatide ou sur la chromatide sœur. Le produit de la réparation par RH est un gène *GFP* fonctionnel.
- D-** Produit de réparation par SSA. Le gène issu de cette réparation n'est pas fonctionnel.

en phase G1. Le NHEJ, quant à lui, est efficace pendant toutes les phases du cycle cellulaire mais il est prédominant pendant la phase G1 et au début de la phase S (Rothkamm, *et al.*, 2003). Une propriété inhérente aux cellules ES, qui les différencie des cellules somatiques, est que la phase S représente 75% du cycle cellulaire, les phases G1 et G2 étant très courtes (Savatier, *et al.*, 2002). Le temps passé en phase S devrait donc favoriser dans les cellules ES le mécanisme de RH, plutôt que le NHEJ. Nous avons vu précédemment que RAD51 joue un rôle essentiel dans les mécanismes de RH. Les embryons déficients pour *Rad51* meurent très précocement au stade blastocyste, lors de la formation de la masse cellulaire interne. Les cellules ES qui sont dérivées à partir de la masse cellulaire interne ne peuvent être obtenues lorsque les deux allèles *Rad51* sont nuls. Ces résultats suggèrent que RAD51 joue un rôle clé dans la prolifération des cellules de la masse cellulaire interne ou dans les mécanismes de réparation des DSB générés par des processus endogènes (Lim and Hasty, 1996; Tsuzuki, *et al.*, 1996). Les protéines Ku70/80, DNA Ligase IV et XRCC4 interviennent, quant à elles, dans les mécanismes de NHEJ. Les souris déficientes pour *Ku70* et *Ku80* sont viables. Les embryons déficients pour *DNA Ligase IV* et *Xrcc4* meurent à un stade plus tardif, autour de E14,5 (Barnes, *et al.*, 1998; Frank, *et al.*, 2000; Gao, *et al.*, 1998). Ces résultats montrent que les protéines impliquées dans le NHEJ ne sont pas essentielles pour les cellules du stade blastocyste et donc pour les cellules ES.

E.D. Tichy et collaborateurs ont montré que le mécanisme de RH prédomine sur le NHEJ dans les cellules ES. En revanche le NHEJ est le mécanisme de réparation majoritaire dans les cellules somatiques (Tichy, *et al.*, 2010). Ils ont observé que les protéines impliquées dans le mécanisme de RH, comme RAD51, RAD52 ou RAD54, ont des niveaux d'expression plus élevés dans les cellules ES que dans les MEFs (*mouse embryonic fibroblasts*). Ils ont utilisé un système d'expression de la GFP qui permet de mesurer la fréquence de réparation par RH (**Figure 18**) (Pierce, *et al.*, 1999). Ils ont ainsi montré que la fréquence de la RH dans les cellules ES est quatre fois plus importante que dans les MEFs. Avec un système qui permet de mesurer le taux de réparation par NHEJ, ils ont montré que le NHEJ est cinq fois plus utilisé dans les MEFs que dans les cellules ES. Par ailleurs, la quantité de DNA ligase IV est limitante dans le NHEJ (Lee, *et al.*, 2000a; Riballo, *et al.*, 2009; Wang, *et al.*, 2007). Or, ils ont observé que le niveau d'expression de la DNA ligase IV est plus faible dans les cellules ES que dans les MEFs. Le faible niveau de réparation par NHEJ dans les cellules ES pourrait s'expliquer ainsi (Tichy, *et al.*, 2010). Enfin, ils ont induit la différenciation de cellules ES et ont observé que le taux de réparation par RH diminuait pour atteindre le même

niveau que dans les MEFs, tandis que le taux de réparation par NHEJ augmentait considérablement.

L'ensemble des données présentées jusqu'à présent indique que la technique de ciblage de gène est favorisée par l'induction d'un DSB, via une MN ou une ZFN. De plus, les différents types cellulaires des mammifères n'utilisent pas les mêmes mécanismes de réparation des DSB. Dans la suite de ce chapitre, nous allons étudier quelques exemples de ciblage de gène utilisant les propriétés des MN ou des ZFN.

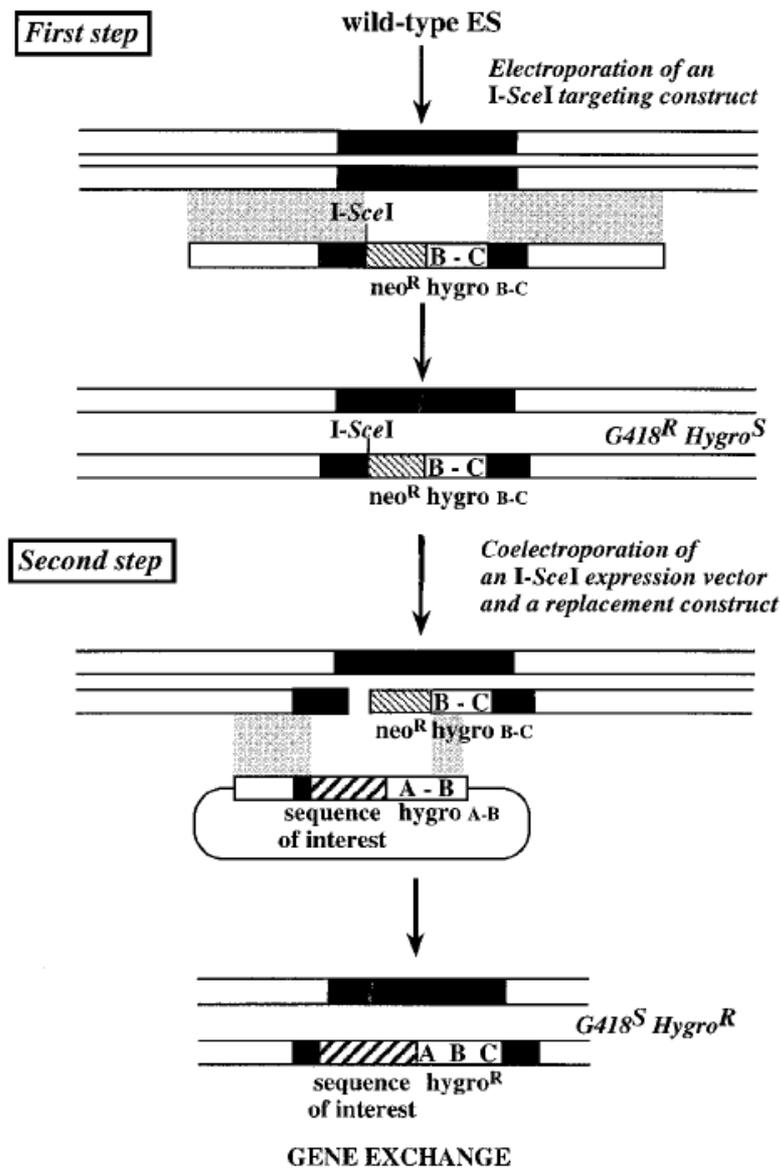
## 5 Utilisation de MN et de ZFN pour induire des modifications génétiques chez les Mammifères

### 5.1 Utilisation de MN

#### 5.1.1 Exemples de ciblage de gène dans les cellules ES murines

Les premiers ciblage de gènes, présentés dans ce chapitre, section 2.3, ont été réalisés majoritairement dans des lignées de cellules différenciées, comme les fibroblastes 3T3. A la suite de ces études, Maria Jasin et son équipe se sont intéressés à l'utilisation de I-SceI pour favoriser la RH dans les cellules ES murines. Ils ont montré que l'introduction d'un site I-SceI suivi d'un DSB à ce site permettait d'augmenter d'au moins 50 fois le ciblage de gène (Smih, *et al.*, 1995). Pour cela, ils ont introduit, dans le génome de cellules ES, le gène de résistance à la néomycine (*néo*) comprenant un site I-SceI interrompant sa phase ouverte de lecture. Ils ont ensuite transfecté ces cellules avec un plasmide d'expression de la protéine I-SceI et un fragment de séquence homologue qui permettait de réparer le gène *néo*. Ils ont obtenu des clones recombinants homologues uniquement lors d'une introduction simultanée des deux plasmides. Ils ont ainsi montré que l'induction d'un DSB via I-SceI stimule la réparation par RH dans le génome de cellules ES de souris. Cependant, cette étude n'a pas permis de montrer que I-SceI favorise l'intégration au lieu de cassure d'un transgène ou d'une cassette de sélection par RH.

Le premier ciblage de gène favorisé par I-SceI à un locus endogène dans les cellules ES murines a été réalisé par M. Cohen-Tannoudji et collaborateurs (Cohen-Tannoudji, *et al.*, 1998). Ils ont d'abord intégré le site I-SceI au locus *Villin* par un ciblage de gène conventionnel : ils ont construit un vecteur de réparation constitué de cassettes de sélection positive et négative, d'une cassette de résistance à l'hygromycine incomplète, de longues régions d'homologie (6,1 et 3,5 kb) et l'ont électroporé dans les cellules ES sous forme linéaire. Une fois le clone portant le site I-SceI sélectionné, ils ont testé la technique de ciblage de gène favorisée par un DSB sur ce clone. Pour cela, ils ont coélectroporé un vecteur d'expression de la MN I-SceI et un vecteur de réparation de la cassure. Le vecteur de réparation, transfecté sous une forme circulaire, est constitué d'une séquence d'intérêt, d'une cassette incomplète de sélection à l'hygromycine et de courtes régions d'homologie au locus *Villin* (2 kb et 0,8 kb, respectivement). Une RH entre le vecteur de réparation et le locus *Villin*



**Figure 19 : Ciblage de gène à un locus naturel favorisé par un DSB**  
(Cohen-Tannoudji, *et al.*, 1998)

La première étape de cette stratégie consiste en l'insertion du site de reconnaissance de I-SceI au locus naturel ciblé par une méthode de RH conventionnelle. Dans la deuxième étape, l'expérimentateur co-transfecte un vecteur d'expression de la méganucléase I-SceI et un vecteur de réparation. L'expression de I-SceI induit un DSB dans le locus ciblé au niveau du site de reconnaissance. Le DSB est ensuite réparé par RH avec le vecteur de réparation. Cette stratégie permet d'introduire dans le locus ciblé n'importe quel gène d'intérêt. La sélection des cellules ES recombinées par l'hygromycine est permise par la restauration d'un gène Hygro<sup>R</sup> fonctionnel lors de la RH (hygro A-B-C).

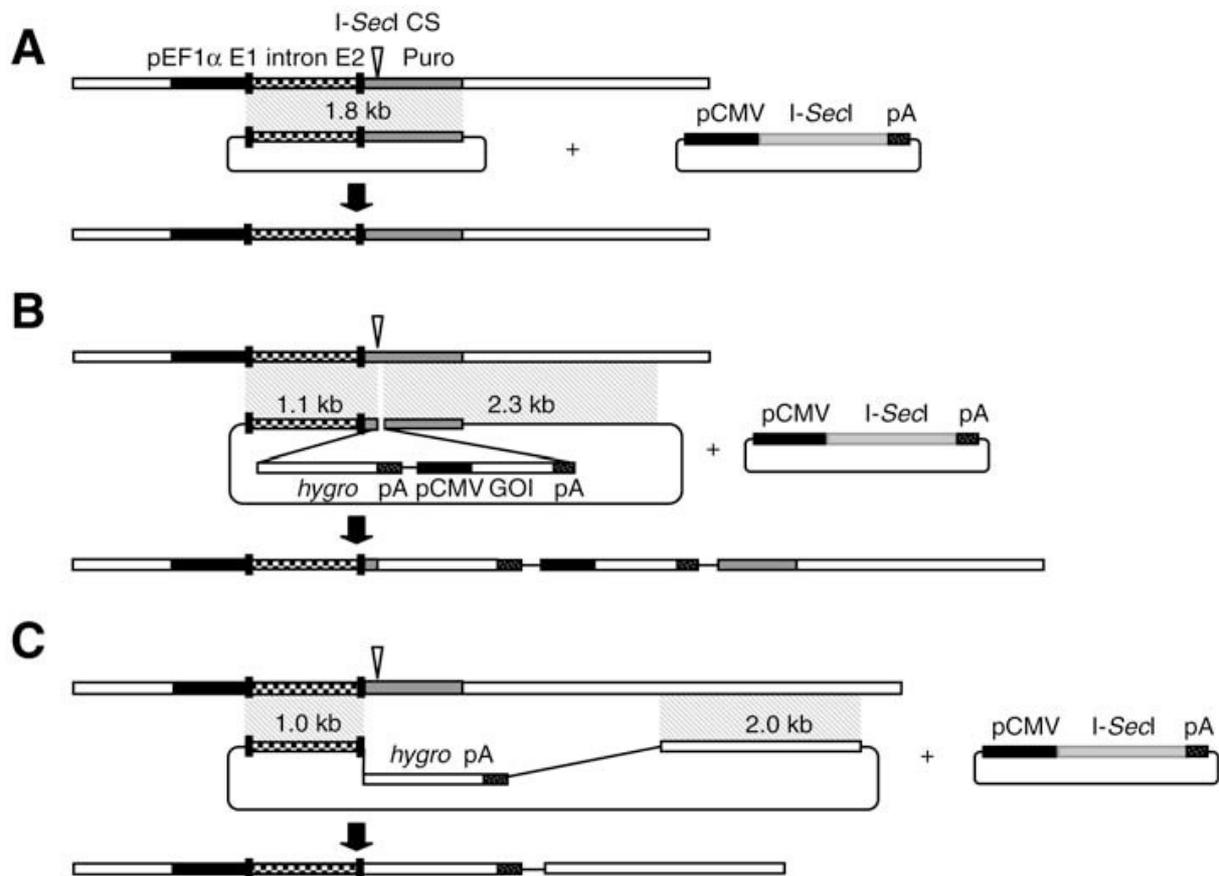
permet d'intégrer la séquence d'intérêt et de compléter la cassette de résistance à l'hygromycine, grâce à laquelle les clones recombinants homologues sont sélectionnés. La stratégie employée pour ce ciblage de gène est illustrée sur la **Figure 19**. Pour tester l'efficacité du ciblage de gène favorisé par un DSB, les expérimentateurs ont électroporé le vecteur de réparation avec ou sans le vecteur d'expression d'I-SceI. Aucun clone Hygro<sup>R</sup> n'a été obtenu en l'absence de MN. Une centaine de clones ont été obtenus lorsque I-SceI était présente. Une dizaine de clones Hygro<sup>R</sup> ont été analysés. Ils étaient tous le produit d'une RH. Ces résultats montrent qu'un DSB au locus *Villin* stimule le ciblage de gène d'un facteur 100 : l'efficacité du ciblage par la méthode conventionnelle au locus *Villin* est de  $5 \times 10^{-8}$ , tandis que l'efficacité du ciblage en présence de I-SceI à ce même locus est de  $6 \times 10^{-6}$ .

En conclusion, I-SceI semble bien favoriser le ciblage de gène dans les cellules ES de souris à un locus naturel. Pour cela, de courtes régions d'homologie sont suffisantes et le vecteur de réparation peut être transfecté sous forme circulaire. A la suite de ce travail, d'autres expériences de réparation par RH favorisée par I-SceI ont été réalisées dans les cellules ES, essentiellement par l'équipe de M. Jasin. Cependant à ma connaissance, aucune autre expérience de ciblage de gène permettant l'intégration d'une séquence d'intérêt à un locus endogène n'a été publiée.

En 1998, G. Donoho et collaborateurs ont également publié un article sur ce sujet. Ils se sont intéressés à l'effet d'un DSB sur les mécanismes de réparation (Donoho, *et al.*, 1998). En particulier, ils se sont demandés si le mode d'introduction du vecteur de réparation - sous forme linéaire ou sous forme circulaire - modifie la fréquence du ciblage de gène. Leurs résultats indiquent que la linéarisation du vecteur de réparation n'est pas nécessaire dans le cas d'un ciblage de gène favorisé par un DSB.

### 5.1.2 Exemples de ciblage de gène dans les cellules CHO

Les laboratoires Collectis Bioresearch ont développé des systèmes de ciblage de gène dans les cellules CHO-K1 (*Chinese hamster ovary-K1*). Dans une première étude, ils ont intégré de façon aléatoire une construction comprenant le gène de résistance à la puromycine (*Puro<sup>R</sup>*) précédé du promoteur, des exons 1 et 2 et du premier intron du gène *EF1 $\alpha$*  humain dans le génome des cellules CHO-K1. Le gène *Puro<sup>R</sup>* est inactivé par un site I-SceI inséré dans son cadre de lecture (Cabaniols and Paques, 2008). Ils ont ensuite construit plusieurs vecteurs de réparation et observé les fréquences de ciblage de gène obtenues avec chacun



**Figure 20 : Différentes stratégies de ciblage de gène favorisées par I-SceI**  
(Cabaniols and Paques, 2008)

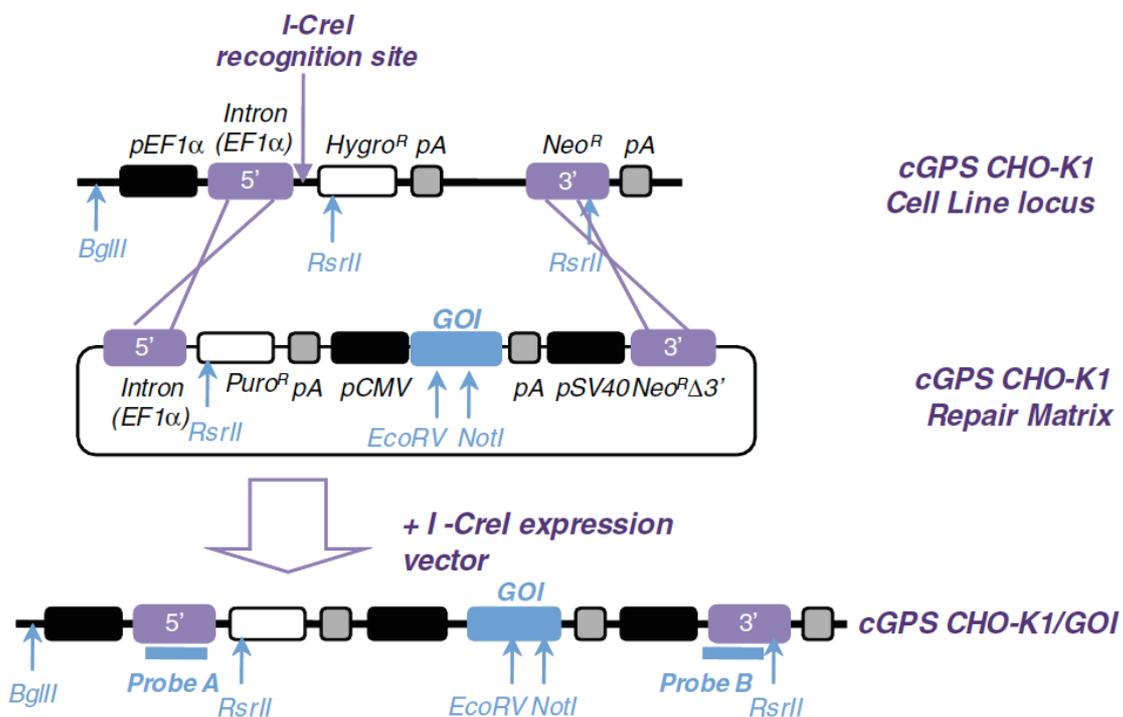
Dans l'allèle ciblé, le gène de la puromycine est interrompu par un site de clivage (CS pour cleavage site) de la méganucléase I-SceI. Les régions d'homologie sont représentées par des zones grisées semi-transparentes. GOI (pour gene of interest) correspond au gène d'intérêt, la  $\beta$ -galactosidase ou le gène CD4 dans cette étude.

- A- Elaboration d'un système de correction de gène. Le vecteur de réparation contient 1,8 kb d'homologie de séquence avec l'allèle ciblé correspondant aux exons non-traduits E1 et E2, l'intron 1 et un gène puromycine entier.
- B- Elaboration d'un système d'insertion de gène. Le vecteur de réparation contient un gène de résistance à l'hygromycine, dépourvu de promoteur, seul ou suivi d'une cassette d'expression d'un gène d'intérêt, l'ensemble étant encadré de séquences d'homologie de 1,1 et 2,3 kb.
- C- Elaboration d'un système de remplacement de gène. Dans ce système, le vecteur de réparation est très similaire à celui utilisé pour le système d'insertion de gène excepté pour : (1) la séquence d'homologie 3' commence après la séquence polyA du gène puromycine et (2) toutes les séquences du gène de résistance à la puromycine sont supprimées de la région d'homologie 5'.

d'entre eux. Les différentes stratégies de ciblage sont illustrées dans la **Figure 20**. Dans un premier cas (Figure 20A), les cellules ont été transfectées avec un vecteur de réparation qui permet la correction du gène *Puro<sup>R</sup>*, car il contient une copie intacte du gène *Puro<sup>R</sup>* et 1,8 kb d'homologie avec le locus ciblé. Des clones *Puro<sup>R</sup>* ont été obtenus avec une fréquence de  $3 \times 10^{-3}$ . Ces clones sont considérés comme des recombinants homologues. Soulignons qu'il ne peut pas être exclu que ces clones soient le résultat d'une NHEJ combinant une délétion du site *I-SceI* et une religation des bouts francs, rétablissant la phase de lecture du gène *Puro<sup>R</sup>*. Dans un deuxième cas (Figure 20B), les cellules ont été transfectées avec un vecteur qui contient une cassette de résistance à l'hygromycine (*hygro<sup>R</sup>*) seule ou avec une séquence d'intérêt. Deux séquences d'intérêt de tailles différentes, les gènes *CD4* ou *LacZ*, ont été testées. Le vecteur de réparation contenait deux régions d'homologies avec le locus cible jouxtant le site de coupure, de 1,1 kb et 2,3 kb respectivement. Les fréquences de RH mesurées étaient les suivantes :  $3,9 \times 10^{-4}$  lors d'une insertion d'un segment de 1,3 kb comprenant la cassette *hygro<sup>R</sup>* seulement,  $1,6 \times 10^{-4}$  lors de l'insertion d'un segment de 4,4 kb comprenant les gènes *hygro<sup>R</sup> + CD4*, et  $0,7 \times 10^{-4}$  lors de l'insertion d'un segment de 5,5 kb comprenant les gènes *hygro<sup>R</sup> + LacZ*. Dans un troisième cas (Figure 20C), les cellules ont été transfectées avec un vecteur de réparation qui contient uniquement la cassette de sélection *hygro<sup>R</sup>*. Dans ce dernier cas, les régions d'homologie de 1 kb et 2 kb étaient séparées de la cassure par 132 pb et environ 2,3 kb, respectivement. La fréquence de RH observée était égale à  $1,4 \times 10^{-4}$ . En conclusion, ces expériences montrent que le ciblage de gène est favorisé par *I-SceI* dans les cellules CHO-K1, même lorsque les séquences à insérer sont de grande taille ou que la recombinaison implique d'importantes délétions au locus ciblé. Néanmoins, la fréquence de ciblage de gène semble diminuer lorsque la taille de l'insertion augmente et lorsque les séquences homologues sont distantes du DSB.

Dans une étude ultérieure l'équipe de Cellectis bioresearch a construit une lignée cellulaire CHO-K1 similaire, portant le gène *hygro<sup>R</sup>* intégré de façon aléatoire, sa phase de lecture étant interrompue par le site de reconnaissance de la MN *I-CreI* (Cabaniols, *et al.*, 2010). Dans cette construction, le gène *hygro<sup>R</sup>* est placé en aval du promoteur et de 1 kb de séquence du gène *EF1 $\alpha$*  humain (cf. ci-dessus); le gène *hygro<sup>R</sup>* est suivi d'un gène de résistance à la néomycine dépourvu de promoteur. Cette stratégie vise à favoriser l'intégration de séquences d'intérêt par RH grâce à l'expression d'une MN dans une lignée modifiée à façon. Elle a été désignée sous l'expression de « système cellulaire de positionnement dans le

génomique » ou cGPS (pour *cellular genome positioning system*). Le principe du cGPS est illustré **Figure 21**. Il faut noter que la région d'homologie en 3' est séparée du site de cassure par environ 1,3 kb, c'est-à-dire la taille du gène *hygro<sup>R</sup>*. Quatre séquences d'intérêt différentes ont été intégrées en utilisant ce système. Les résultats montrent qu'un DSB favorise la RH. Les auteurs de ces expériences ont conclu que ce système cGPS est plus rapide qu'une méthode d'intégration aléatoire et qu'elle permet une expression stable des séquences intégrées.



**Figure 21 : Principe du cGPS**  
(Cabaniols, *et al.*, 2010)

Le locus ciblé contient le site de reconnaissance de la méganucléase I-CreI inséré avec un gène de résistance à l'hygromycine dépourvu de promoteur. Le vecteur de réparation du système cGPS est constitué de deux bras d'homologie au locus ciblé, d'une cassette de résistance à la puromycine, d'une cassette d'expression d'un gène d'intérêt et d'un gène de résistance à la néomycine incomplet. Suite à la création d'un DSB par I-CreI, il y a intégration au locus ciblé de la cassette d'expression du gène d'intérêt et reconstitution de deux gènes de résistance fonctionnels (puromycine et néomycine), par un mécanisme de RH.

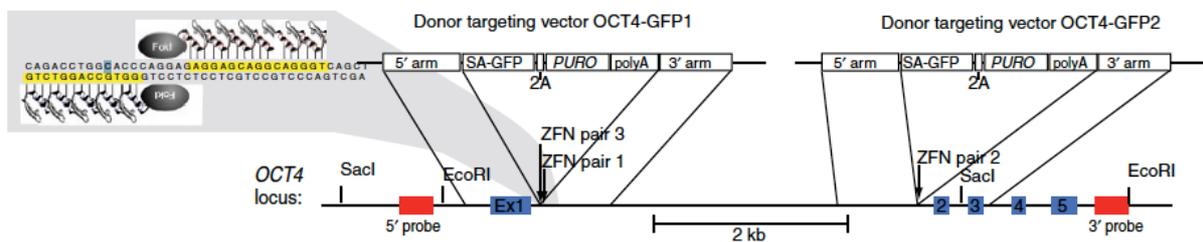
## 5.2 Utilisation de ZFN

À l'heure actuelle, il existe déjà de nombreux exemples d'utilisation de ZFN pour favoriser le ciblage de gène dans les cellules de mammifères, mais aussi directement chez l'animal. Le **Tableau 1** représente un éventail de ciblage de gènes réalisés à l'aide de ZFN dans différents organismes. Je vais m'intéresser plus particulièrement à deux études réalisées sur des cellules de mammifères, l'une chez l'Homme et l'autre chez la Souris.

Organism	Latin name	Method	TM	TGR	References
Animals					
Fruit fly	<i>Drosophila melanogaster</i>	Heat-shock induction	+	+	Bibikova <i>et al.</i> (2002, 2003), Beumer <i>et al.</i> (2006)
		Embryo injection	+	+	Beumer <i>et al.</i> (2008)
Nematode	<i>C. elegans</i>	Gonad injection	+		Morton <i>et al.</i> (2006)
Silkworm	<i>Bombyx mori</i>	Embryo injection	+		Takasu <i>et al.</i> (2010)
Zebrafish	<i>Danio rerio</i>	Zygote injection	+		Meng <i>et al.</i> (2008), Doyon <i>et al.</i> (2008), Foley <i>et al.</i> (2009)
Sea urchin	<i>Hemicentrotus pulcherrimus</i>	Embryo injection	+		Ochiai <i>et al.</i> (2010)
Frog	<i>Xenopus tropicalis</i>	Embryo injection	+		Young <i>et al.</i> (2011)
Rat	<i>Rattus norvegicus</i>	Zygote injection	+	+	Geurts <i>et al.</i> (2009), Mashimo <i>et al.</i> (2010)
Mouse	<i>Mus musculus</i>	Zygote injection	+	+	Meyer <i>et al.</i> (2010), Carbery <i>et al.</i> (2010), Cui <i>et al.</i> (2011)
Plants					
Cress	<i>A. thaliana</i>	Agrobacterium	+		Lloyd <i>et al.</i> (2005), Zhang <i>et al.</i> (2010), Osakabe <i>et al.</i> (2010), De Pater <i>et al.</i> (2009)
Tobacco	<i>Nicotiana sp.</i>	Protoplasts	+	+	Wright <i>et al.</i> (2005), Townsend <i>et al.</i> (2009)
		Agrobacterium	+	+	Cai <i>et al.</i> (2009)
		Viral delivery	+		Marton <i>et al.</i> (2010)
Maize	<i>Zea mays</i>	Cell culture	+	+	Shukla <i>et al.</i> (2009)
Petunia	<i>Petunia sp.</i>	Viral delivery	+		Marton <i>et al.</i> (2010)
Mammalian cells in culture					
Human	<i>Homo sapiens</i>	DNA transformation	+	+	Porteus and Baltimore (2003), Urnov <i>et al.</i> (2005), Alwin <i>et al.</i> (2005), Perez <i>et al.</i> (2008), Hockemeyer <i>et al.</i> (2009), Kim <i>et al.</i> (2009), Zou <i>et al.</i> (2009), Dekelver <i>et al.</i> (2010)
		Viral delivery	+	+	Lombardo <i>et al.</i> (2007)
Mouse	<i>M. musculus</i>	DNA transformation	+	+	Goldberg <i>et al.</i> (2010), Connelly <i>et al.</i> (2010)
Hamster	<i>Cricetulus griseus</i>	DNA transformation	+	+	Santiago <i>et al.</i> (2008), Liu <i>et al.</i> (2010), Cost <i>et al.</i> (2010)
Pig	<i>Sus domestica</i>	DNA transformation	+		Watanabe <i>et al.</i> (2010)

**Tableau 1 : Exemples de ciblage de gène réussis via l'induction d'un DSB par des ZFN (d'après (Carroll, 2011))**

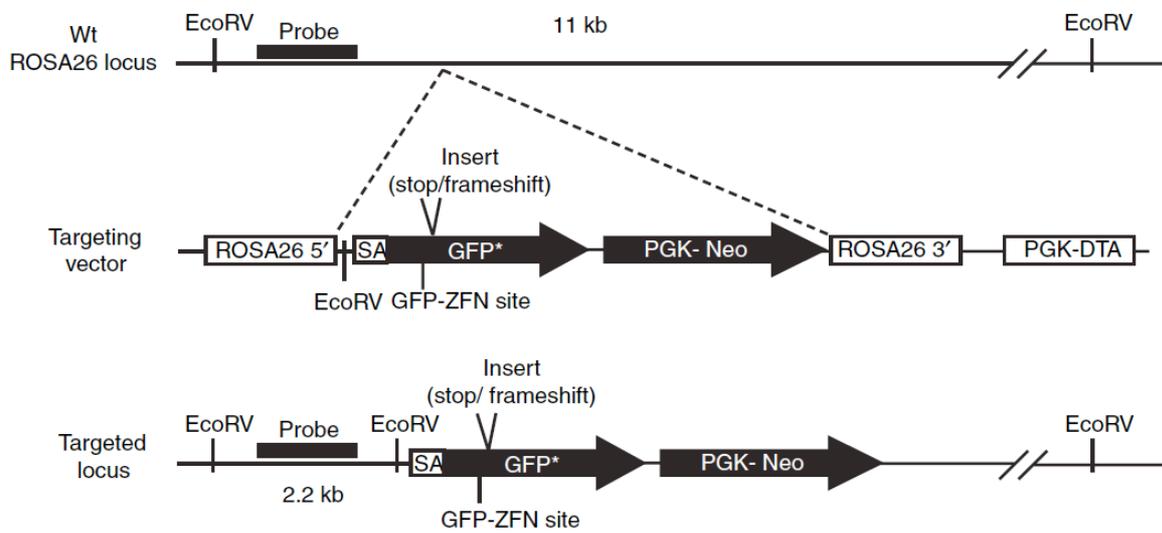
TM (pour targeted mutagenesis) correspond aux réparations par NHEJ. TGR (pour targeted gene replacement) correspond aux ciblage de gène. La liste des références n'est pas exhaustive.



**Figure 22 : Ciblage du gène *Oct4* à l'aide de ZFN**  
(Hockemeyer, *et al.*, 2009)

Les rectangles bleus correspondent aux exons du gène *OCT4*. Les flèches représentent les sites génomiques ciblés par les ZFNs. Les vecteurs de réparation sont représentés sur le haut du schéma. Pour chaque paire de ZFNs, un vecteur de réparation spécifique est construit, contenant approximativement 700 pb de séquences homologues au gène *OCT4*. SA-eGFP (pour splice-acceptor-eGFP) correspond à un site accepteur d'épissage lié à la protéine GFP. La séquence 2A est une séquence de peptide qui s'auto-clive. Ces séquences sont introduites au locus *OCT4* par RH favorisée par une paire de ZFNs, entraînant l'expression de la protéine GFP et du gène de résistance à la puromycine sous contrôle du promoteur du gène *OCT4*. La partie gauche du schéma représente une paire de ZFNs spécifique à une séquence génomique figurée en jaune, ce qui entraîne la dimérisation de la nucléase *FokI*.

Dans une étude publiée en 2009, D. Hockemeyer et collaborateurs ont réalisé des ciblage de gène favorisés par l'induction de DSB à l'aide de ZFN sur trois gènes différents dans des cellules souches embryonnaires humaines (hESCs pour *human embryonic stem cells*) et dans des cellules souches pluripotentes induites (hiPSCs pour *induced pluripotent stem cells*). Les trois gènes endogènes ciblés sont *OCT4*, *AAVS1* et *PITX3* (Hockemeyer, *et al.*, 2009). *OCT4*, le marqueur de pluripotence, est exprimé dans les cellules hESCs. *OCT4* a été ciblé via trois paires de ZFN différentes. Les vecteurs de réparation contenaient, en plus de régions d'homologie 5' et 3' courtes (moins de 1 kb pour chaque bras d'homologie), une cassette de résistance à la puromycine et un gène d'expression de l'eGFP, tous les deux dépourvus de promoteur. Une fois la RH effectuée, ces gènes se retrouvaient sous contrôle du promoteur endogène *OCT4* (**Figure 22**). L'efficacité du ciblage des trois paires de ZFN est mesurée par le rapport du nombre de clones recombinants homologues (sans insertion aléatoire supplémentaire) sur le nombre de clones prélevés. Les efficacités de ciblage étaient supérieures à 94% avec la ZFN n°1, compris entre 36 et 53% avec la ZFN n°2, et très faible avec la ZFN n°3, avec laquelle il y a eu très peu de clones résistants à la puromycine. Remarquons que les ZFN n°1 et 3 ciblent des séquences proches l'une de l'autre et favorisent la réparation avec un même vecteur de réparation. Pourtant leurs efficacités de ciblage diffèrent d'un facteur 2 à 3. *AAVS1* est le deuxième gène ciblé. *AAVS1* est exprimé de façon ubiquitaire. Pour le locus *AAVS1*, les expérimentateurs n'ont utilisé qu'une paire de ZFN, mais plusieurs vecteurs de réparation. Pour tous les vecteurs de réparation, les régions d'homologie 5' et 3' sont courtes (autour de 500 pb). Les différences entre les vecteurs ne seront pas détaillées ici. Que ce soit dans les cellules hESCs ou iPS, les efficacités de ciblage de cette ZFN associé aux différents vecteurs de réparation sont comprises entre 33 et 61%. *PITX3* le troisième gène ciblé n'est pas exprimé dans les cellules hESCs et hiPSCs. Dans ce dernier cas, les auteurs ont utilisé une seule paire de ZFN et un seul vecteur de réparation. L'efficacité de ciblage était égale à 11% pour les cellules hESCs et à 8% pour les cellules hiPSCs. Les auteurs ont conclu que ces efficacités de ciblage, bien que plus faibles que pour les deux autres locus, sont relativement élevées comparées à ce qui est obtenu par RH conventionnelle à des locus non-exprimés dans les cellules hESCs. Il aurait été intéressant, dans cet article, de présenter une RH conventionnelle au locus *PITX3*, pour que les efficacités de ciblage soient directement comparables. En conclusion, le ciblage de gène favorisé par un DSB via une ZFN est dans l'ensemble très efficace et nécessite de courtes homologies de séquences, comparé au ciblage de gène conventionnel.



**Figure 23 : Ciblage de gène au locus ROSA favorisés par des ZFN**  
(Connelly, *et al.*, 2010)

La méthode conventionnelle de RH est utilisée pour insérer au locus ROSA26 un gène GFP muté précédé d'un site accepteur d'épissage. La mutation du gène GFP est due à une séquence de 85 nt insérée dans son cadre de lecture, dont un codon stop. GFP-ZFN site indique le site de coupure de la ZFN utilisée pour favoriser le ciblage de gène.

M.H. Porteus et collaborateurs ont récemment mesuré l'efficacité du ciblage de gène favorisé par un DSB, induit par une ZFN, dans des cellules murines. Ces expériences ont été réalisées dans des cellules de souris de plusieurs types cellulaires : dans des cellules ES et dans des cultures primaires de fibroblastes d'embryon (MEFs), d'adulte et d'astrocytes (Connelly, *et al.*, 2010). Une construction contenant un gène *GFP* dont la séquence de lecture est interrompue par un codon stop a été insérée au locus *ROSA26* par une méthode de *knock in* conventionnelle dans des cellules ES de souris (**Figure 23**). Les ZFN développées pour cette étude ne ciblent pas un locus naturel mais un site situé dans le gène rapporteur *GFP*, en amont du codon stop. Une lignée de souris a été dérivée des cellules ES portant cette construction au locus *ROSA26*. Pour mesurer l'efficacité du ciblage de gène dans les cellules ES, les auteurs ont transfecté simultanément un vecteur de réparation et un plasmide d'expression d'une ZFN. La forme (linéaire ou circulaire) du vecteur de réparation n'est pas précisée dans l'article. Le vecteur de réparation porte une séquence *GFP* incomplète qui permet la correction du gène *GFP* muté, inséré au locus *ROSA26*. Le pourcentage de cellules qui exprimaient la protéine GFP fluorescente par rapport au nombre de cellules transfectées a été mesuré en cytométrie de flux. Pour les auteurs, les cellules qui étaient GFP-positives ont subi un ciblage de gène favorisé par un DSB au locus *ROSA26*. Il était intéressant de déterminer le ratio optimal entre la quantité de vecteur de réparation et le plasmide d'expression d'une ZFN. L'efficacité maximale de ciblage au locus *ROSA26*,  $2,1 \times 10^{-3}$  a été obtenue dans les cellules ES avec un ratio de 90% de vecteur de réparation pour 10% de vecteur d'expression. Le blocage des cellules en phase S/G2 par la vinblastine augmentait l'efficacité de ciblage d'un facteur compris entre 4 et 5, soit  $1,6 \times 10^{-2}$ . Ces expériences montrent qu'une ZFN améliore le ciblage de gène dans les cellules ES de souris par un facteur supérieur à 1 000 (l'efficacité est normalement inférieure à  $10^{-5}$ ). Pour mesurer l'efficacité de ciblage dans d'autres types cellulaires, J.P. Connelly et collaborateurs ont dérivé plusieurs types cellulaires à partir des souris *knock in*. Dans des fibroblastes de type 3T3, le meilleur ratio entre vecteurs de réparation et d'expression était égal à 90%/10% pour une efficacité de ciblage atteignant  $1,8 \times 10^{-2}$  dans des cellules hétérozygotes pour le *knock in*, et s'élevant jusqu'à  $6,7 \times 10^{-2}$  dans des cellules homozygotes pour le *knock in*. Dans des cultures primaires de fibroblastes d'embryon ou d'adulte, les taux de ciblage étaient de l'ordre de  $2 \times 10^{-2}$ . Enfin, dans des cultures d'astrocytes, le taux de ciblage était compris entre  $3 \times 10^{-4}$  et  $1,7 \times 10^{-3}$ . En conclusion, ils ont démontré que cette méthode de ciblage est efficace dans tous ces types cellulaires, avec des fréquences variables. Ils ont aussi montré que le ratio vecteur de

réparation/vecteur d'expression de la ZFN peut être optimisé pour améliorer l'efficacité de ciblage.

Dans ce chapitre, nous avons passé en revue les avancées majeures dans le domaine de la transgénèse chez les Mammifères. Les méganucléases et les nucléases à doigts de zinc offrent de nouvelles possibilités aux chercheurs qui veulent développer des ciblage de gène dans des types cellulaires et dans des espèces animales où ce n'était pas envisageable il y a quelques années. En théorie, les méganucléases et les nucléases à doigts de zinc permettent de réaliser un ciblage de gène dans n'importe quel type cellulaire et à n'importe quel locus. La nouvelle méthode d'expression des nucléases à doigts de zinc, à partir d'ARNm injecté directement dans l'embryon, devrait permettre de créer des mutations ciblées chez l'animal, sans utiliser de cellules souches embryonnaires. Chez d'autres espèces que les Mammifères, les méganucléases et les nucléases à doigts de zinc doivent pouvoir cibler n'importe quel gène chez n'importe quel organisme. Des exemples de ciblage favorisés par des nucléases à doigts de zinc dans de nombreux organismes, sont présentés dans le **Tableau 1**. Ces nucléases ont également d'importantes potentialités en thérapie génique. Un essai clinique en cours teste l'emploi d'une nucléase à doigt de zinc pour invalider le gène *CCR5* (Urnov, *et al.*, 2010). Ce gène code pour un co-recepteur de HIV-1. Les patients qui portent des variants codant une protéine CCR5 non fonctionnelle ne développent pas de SIDA après une infection par le VIH. Le protocole clinique nécessite d'isoler des cellules T précurseurs, de modifier *CCR5* par mutagenèse via une paire de ZFN, d'amplifier des cellules modifiées et de les réimplanter chez le donneur. Le système immunitaire du patient peut ainsi être reconstitué à partir de cette population de cellules résistantes au VIH. C'est la première étape vers les nombreuses applications promises par les nucléases à doigts de zinc. Cependant, les possibilités offertes par les méganucléases et les nucléases à doigts de zinc sont différées, à cause du manque de maîtrise du site de coupures. Les coupures à des sites chromosomiques non attendus (coupures *off-target*) sont mal comprises et représentent un problème important quand l'objectif est de faire de la thérapie génique chez l'homme (Carroll, 2011). Nous reviendrons brièvement sur le phénomène des coupures *off-target* dans la partie Résultats et discussions (1.2.3).

Dans mon projet de thèse, notre objectif était de favoriser la RH au locus *Dct* dans les cellules souches embryonnaires de souris en introduisant un site I-*SceI* dans la séquence de ce

gène. Nous avons inséré deux gènes rapporteurs avec cette méthode et nous avons étudié le profil d'expression de ces gènes intégré au locus *Dct*. Le deuxième chapitre de l'introduction traite donc du lignage mélanocytaire et de l'expression de *Dct* dans les mélanocytes et à leurs précurseurs.



## ***Chapitre 2 : Le gène *Dct*, de sa spécificité au lignage mélanocytaire à son utilisation en transgénèse***

Les cellules pigmentaires ou mélanocytes synthétisent la mélanine, pigment responsable de la couleur de la peau et du pelage des mammifères. Ces cellules sont présentes dans les follicules pileux, l'épiderme, le derme, l'oreille interne et la rétine. Les mélanocytes sont dérivés de la crête neurale. On appelle mélanoblastes les précurseurs des mélanocytes chez l'embryon et chez l'adulte. Dans ce chapitre, nous discuterons de l'origine embryonnaire du mélanocyte à la lumière des récentes découvertes faites sur ce sujet, et du développement du follicule pileux. Nous présenterons ensuite le profil d'expression du gène *Dct*, spécifique des mélanocytes et de ses précurseurs, et les signaux responsables de la régulation de l'expression de ce gène. Enfin, nous étudierons les différents outils de transgénèse construits avec le promoteur du gène *Dct* qui ont permis d'étudier le lignage mélanocytaire.

### **1 Le mélanocyte, origine embryonnaire et développement du follicule pileux**

#### **1.1 Origine embryonnaire et migration**

##### ***1.1.1 Détermination des mélanoblastes dans la crête neurale et migration par la voie dorso-latérale***

La crête neurale est une structure transitoire d'origine neuroectodermique, spécifique des embryons de Vertébrés. Elle se forme tout le long de l'axe antéro-postérieur. Chez l'embryon de poulet, elle se forme entre le premier et le deuxième jour du développement, chez la souris, entre le huitième et le neuvième jour du développement embryonnaire (E8 et E9) et, chez l'homme, au cours de la troisième semaine de gestation. Les cellules de la crête neurale sont induites à la frontière entre le futur épiderme et le futur système nerveux central (le tube neural), au moment où ceux-ci se séparent l'un de l'autre (Moury and Jacobson, 1989, 1990; Selleck and Bronner-Fraser, 1995). Les principaux dérivés de la crête neurale sont : les neurones et les cellules gliales du système nerveux périphérique, les mélanocytes, les cellules musculaires lisses, le tissu conjonctif et enfin le cartilage et les éléments du squelette crânien

(Le Douarin, *et al.*, 2004). La crête neurale est à l'origine de la plupart des cellules pigmentaires du corps des Vertébrés (Rawles, 1947). Les mélanocytes des follicules pileux, de l'épiderme, du derme, de la choroïde, de la glande de Harder et de la stria vascularis de l'oreille interne sont dérivés de la crête neurale (Steel and Barkway, 1989). Seuls les mélanocytes de l'épithélium pigmentaire de la rétine (RPE pour *retinal pigment epithelium*) ont une origine différente ; ils dérivent de la cupule optique.

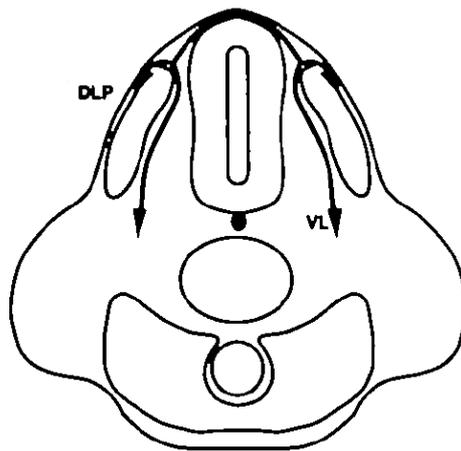
La délamination est le processus qui conduit au détachement et à la migration d'une population de cellules présente au sein d'un épithélium auparavant homogène. Les dérivés des cellules de crête neurale délaminent tous à partir de la surface dorsale du tube neural. La distribution finale de chaque dérivé est, quant à elle, spécifique. Deux modèles sont avancés pour expliquer les différences de distribution des dérivés de crêtes neurales.

Le premier modèle postule que les cellules de la crête neurale quittant le tube neural sont multipotentes. Les cellules en migration subissent une restriction de leurs potentialités en fonction du micro-environnement rencontré. Dans ce modèle, le micro-environnement joue un rôle primordial en orientant les précurseurs multipotents vers un phénotype précis et en permettant la survie et la maturation de sous-populations déjà déterminées (Le Douarin, *et al.*, 1993; Reedy, *et al.*, 1998; Selleck and Bronner-Fraser, 1995; Stemple and Anderson, 1993). Pour une cellule, la détermination correspond à l'acquisition, en principe irréversible, de sa destinée. Selon ce premier modèle, des cellules de crête neurale sont déterminées en tant que mélanoblastes dans la zone de migration constituée par l'espace entre la surface dorsale du tube neural, les somites et l'ectoderme, appelé MSA (*migration staging area*) (Serbedzija, *et al.*, 1990; Serbedzija, *et al.*, 1994).

Au contraire, le deuxième modèle postule que les cellules de crête neurale sont déterminées dès leur émergence du tube neural et empruntent la voie de migration dictée par leur détermination. Ce modèle n'exclut pas que la spécification soit réversible. Chez le poulet, des expériences de greffes et des études sur le développement du lignage mélanocytaire *in vivo* et *in vitro* ont montré que :

- la détermination des cellules de crête neurale en tant que précurseurs de mélanoblastes a lieu avant leur entrée dans la voie de migration dorso-latérale,
- les mélanoblastes migrent préférentiellement selon la voie dorso-latérale,
- seuls les mélanoblastes ont la capacité d'entrer dans la voie de migration dorso-latérale (Erickson and Goins, 1995; Reedy, *et al.*, 1998).

Les mélanoblastes émergent des crêtes neurales, tandis que les mélanocytes fonctionnels se trouvent essentiellement dans l'épiderme et les follicules pileux. Les précurseurs des mélanocytes doivent donc migrer depuis la zone de détermination jusqu'au lieu de différenciation qui correspond à leur emplacement définitif. Chez la souris, le profil de migration des mélanocytes a été mis en évidence par coloration des cellules de crête neurale avec des colorants vitaux (Serbedzija, *et al.*, 1990; Serbedzija, *et al.*, 1994) et par l'étude du profil d'expression du gène *Dct* (*Dopachrome tautomérase*), aussi appelé *Tyrp2* (*Tyrosine related protein 2*) (Steel, *et al.*, 1992). Les mélanoblastes migrent par la voie dorso-latérale, entre le dermomyotome et la lame basale de l'ectoderme (**Figure 24**). Chez la souris, la migration dorso-latérale a lieu entre E8 et E12,5 (Hirobe, 1995; Serbedzija, *et al.*, 1990; Serbedzija, *et al.*, 1994).



**Figure 24 : Représentation schématique de la voie de migration des cellules de crête neurale**  
(Serbedzija, *et al.*, 1990)

Au stade embryonnaire E9, les cellules dérivées de la crête neurale suivent deux voies distinctes : (1) la voie dorso-latérale entre le dermomyotome et l'épiderme (DLP) et (2) la voie ventrale entre le tube neural et le dermomyotome (VL).

### 1.1.2 Les précurseurs des cellules de Schwann, un nouveau réservoir de mélanoblastes

L'ensemble des données présentées ci-dessus a établi que toutes les cellules pigmentaires présomptives suivent la voie dorso-latérale, pour ensuite envahir la peau et se différencier en mélanocytes. Pourtant, une étude récente montre que le système nerveux

périphérique présent dans la peau pendant le développement est à l'origine d'un grand nombre de mélanocytes (Adameyko, *et al.*, 2009). Selon cette étude, les précurseurs des cellules de Schwann ou SCP (*Schwann cell precursors*), en contact avec les axones, donnent naissance non seulement à des cellules gliales, via le processus de gliogenèse, mais aussi à des mélanocytes. Le potentiel de développement et la plasticité des SCP sont les questions au centre de cette étude.

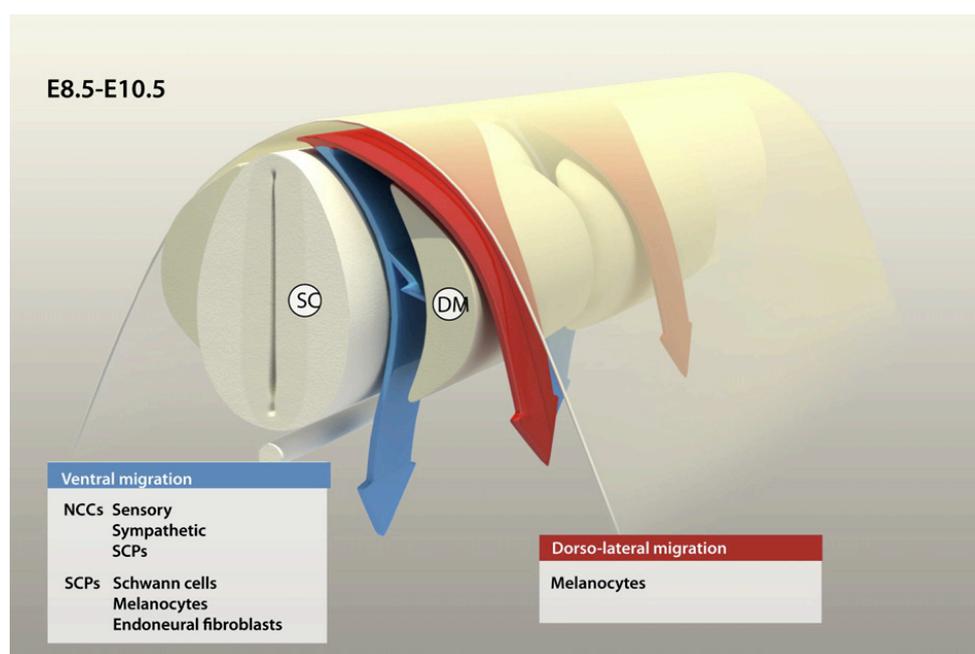
➤ Origine et différenciation des SCP

Les SCP dérivent des cellules de la crête neurale qui migrent pour former le système nerveux périphérique pendant le développement embryonnaire. Chez la souris, durant les stades E9,5 à E14, les cellules dérivées de crête neurale, situées à l'intérieur ou à la surface des nerfs du système nerveux périphérique, sont appelées les SCP. A ce stade, ces cellules ne possèdent plus les caractéristiques biochimiques et morphologiques des cellules de la crête neurale (pour revue voir (Jessen and Mirsky, 2005)). De façon caractéristique, leur survie dépend de leur contact avec le système nerveux (Jessen, *et al.*, 1994). Ces cellules vont donner naissance à toutes les cellules de Schwann du système nerveux adulte via un processus défini. Pendant un stade intermédiaire, des cellules de Schwann dites immatures se distinguent des SCP. En effet, ces cellules de Schwann immatures expriment des marqueurs moléculaires différents des SCP, elles possèdent une lame basale et elles ne dépendent plus d'un contact direct avec les axones pour survivre, puisqu'elles possèdent leur propre circuit autocrine (Jessen and Mirsky, 2005; Meier, *et al.*, 1999).

À l'heure actuelle, il est postulé que les SCP sont générés en l'absence d'autres signaux. Leur voie de différenciation à partir des cellules de crête neurale est considérée comme une voie par défaut. Les signaux impliqués dans le contrôle de la gliogenèse à partir de la crête neurale, comme NRG1 (*neuregulin-1*), BMP2 et 4 (*bone morphogenic protein*) et Notch, ont pour fonction d'inhiber le développement des cellules gliales, ou de supprimer ou d'activer le développement neuronal (Morrison, *et al.*, 2000; Shah, *et al.*, 1994; Shah, *et al.*, 1996). Il est difficile de prouver que ces facteurs initient positivement la gliogenèse. De plus, bien que le facteur de transcription SOX10 (*SRY-box containing gene 10*) soit nécessaire au développement de la glie, il est en fait exprimé par toutes les cellules de la crête neurale. Ce facteur ne semble donc pas faire partie d'une cascade de signaux responsables spécifiquement du développement de la glie (Britsch, *et al.*, 2001). Ainsi, la détermination des SCP n'est pas régulée par des facteurs spécifiques mais par des activités combinatoires de plusieurs signaux étroitement connectés au système nerveux.

➤ Nouveau panorama sur les origines et la migration des mélanoblastes

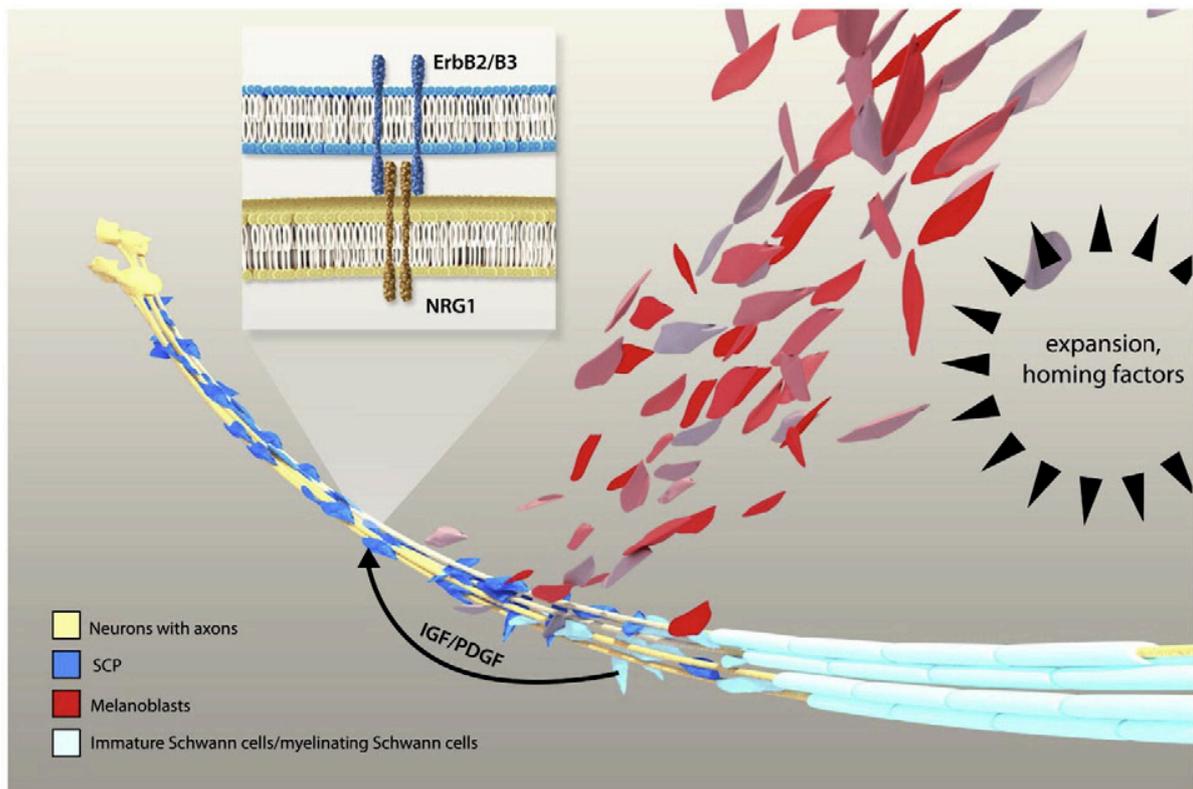
Le dogme, admis jusque récemment, postulait que toutes les cellules pigmentaires présomptives sont uniquement issues de la voie dorso-latérale. Cependant les études qui ont permis de montrer ce fait ne posaient pas la question de comment les mélanoblastes migrent jusqu'au futur derme des régions ventrales du corps et des membres. L'étude réalisée par Adameyko et collaborateurs en 2009 a permis de caractériser plus particulièrement la migration des mélanoblastes vers la peau. Des preuves y sont présentées de l'existence d'une voie de développement des mélanoblastes qui est distincte de la voie dorso-latérale, d'un point de vue spatial et temporel, et qui est associée à l'émergence des nerfs spinaux (Adameyko, *et al.*, 2009). Ces résultats suggèrent que les dérivés de la crête neurale empruntant la voie ventrale sont une source possible de mélanoblastes (**Figure 25**). Des expériences de traçage du lignage mélanocytaire ont permis d'identifier les SCP associés au système nerveux périphérique comme la source principale des mélanoblastes et des mélanocytes matures (pigmentés) dans la peau de l'adulte (Adameyko, *et al.*, 2009).



**Figure 25 : Voies de migration des cellules de crête neurale**  
(Ernfors, 2010)

Les principaux types de cellules dérivées de crête neurale (NCCs) qui migrent par la voie ventrale sont les précurseurs des cellules sensorielles, des cellules sympathiques et des cellules de Schwann (SCPs). Les précurseurs des cellules de Schwann donnent naissance aux cellules de Schwann, aux mélanocytes et aux fibroblastes endoneuraux.

SC (spinal cord) pour la moelle épinière ; DM pour dermomoyotome ; NCCs (neural crest derived cell types) pour cellules dérivées de la crête neurale



**Figure 26 : Illustration schématique des signaux et interactions dans le système nerveux périphérique**  
(Ernfors, 2010)

Il a été démontré qu'une situation dans laquelle il y a diminution du signal NRG1 et augmentation des signaux IGF/PDGF entraîne la cellule à emprunter la voie du lignage mélanocytaire. Ainsi le mécanisme proposé est le suivant : les SCP migrent le long des nerfs et se différencient en cellules de Schwann immatures ou matures, ce qui diminue le contact nerveux et le signal NRG1 et favorise donc la destinée mélanocytaire. Au même moment du développement, les niveaux de IGF/PDGF produits par les cellules de Schwann immatures et matures augmentent et favorisent la détermination des SCP en mélanoblastes. Les mélanoblastes ainsi spécifiés prolifèrent massivement puis migrent vers les follicules pileux et la membrane basale de l'épiderme, sous l'influence des facteurs de localisation (« homing factors ») et d'expansion exprimés dans la peau.

Au vu des résultats obtenus, la question s'est posée de l'existence d'un mécanisme similaire chez l'adulte. Bien que les cellules de Schwann différenciées ne puissent pas donner de mélanoblastes au cours du développement, il a été montré que ces cellules restent compétentes pour produire des cellules pigmentées chez l'adulte. Lorsque des cellules de Schwann différenciées sont confrontées à un nouvel environnement, à la suite par exemple de la perte de contact avec le système nerveux périphérique, elles sont capables de donner des cellules pigmentées (Adameyko, *et al.*, 2009). Le choix de la destinée cellulaire entre la glie et le mélanocyte dépend donc du contact avec les axones.

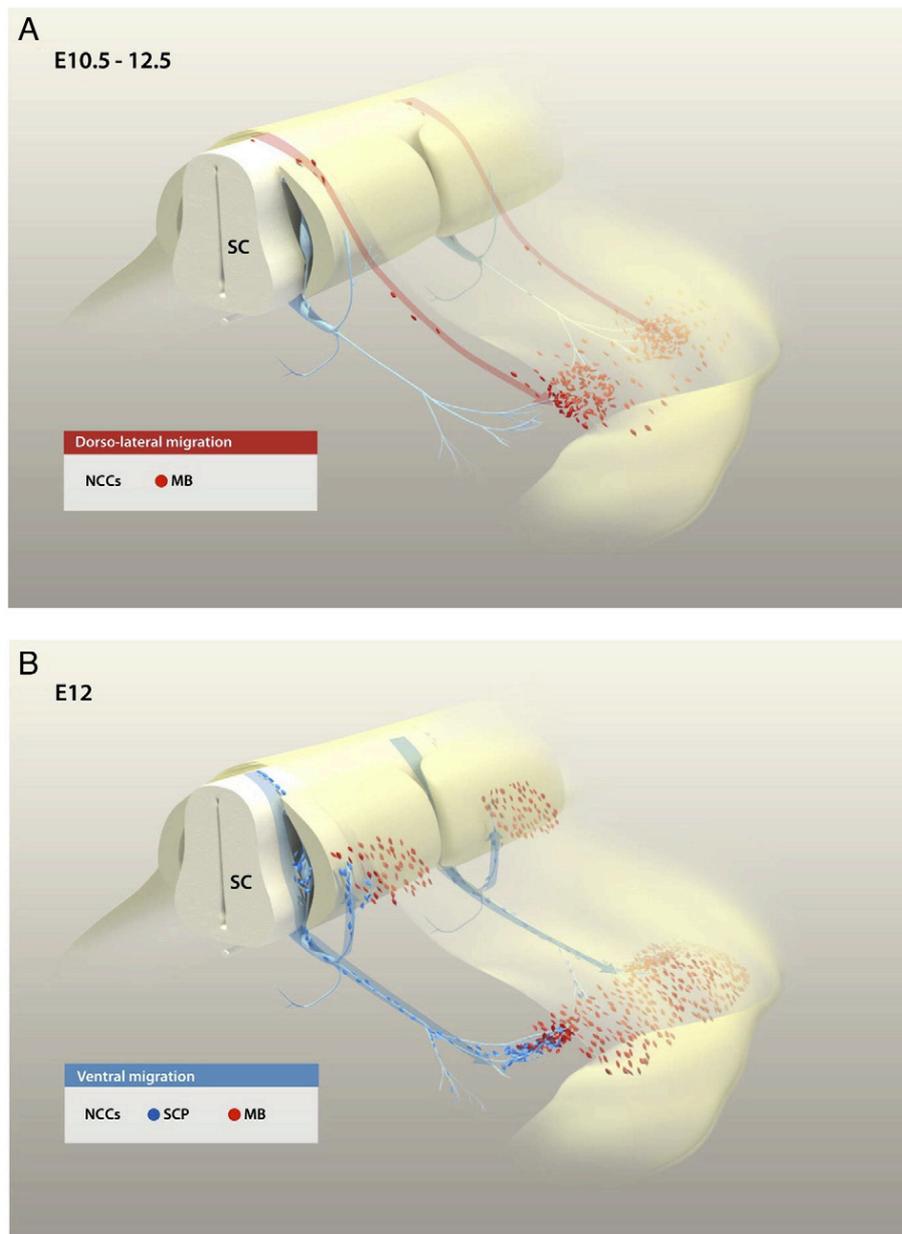
Pour analyser les bases moléculaires de ce phénomène, une lignée de souris délétée pour le gène *ErbB3* a été étudiée (Adameyko, *et al.*, 2009). Ce gène exprime l'un des composants de l'hétérodimère ErbB2/ErbB3, le récepteur tyrosine kinase du ligand NRG1. NRG1 est synthétisé par les neurones sensoriels et moteurs, il est localisé dans la membrane des axones (Bermingham-McDonogh, *et al.*, 1997; Ho, *et al.*, 1995). Ce ligand s'accumule dans les axones au moment où les SCP colonisent les nerfs spinaux. NRG1 est essentiel à la survie de ces précurseurs (Riethmacher, *et al.*, 1997). Il semble réguler la balance entre les cellules gliales et les mélanocytes. L'étude de la délétion du gène *ErbB3* a permis d'analyser le rôle joué par NRG1 dans le développement des mélanocytes. Cette délétion induit une augmentation du nombre de mélanoblastes au stade embryonnaire E12 aux extrémités distales des nerfs spinaux, malgré une diminution globale des SCP le long du système nerveux périphérique (Adameyko, *et al.*, 2009). De plus, une expérience *in vitro* a montré que l'addition de NRG1 sur des cellules gliales en culture réduit le nombre de mélanoblastes nouvellement formés (Adameyko, *et al.*, 2009). Ces deux résultats montrent que NRG1 régule négativement la détermination des SCP en mélanoblastes.

La survie des SCP est sous le contrôle du facteur NRG1 jusqu'à leur transformation en cellules de Schwann immatures, qui produisent et répondent aux facteurs autocrines de survie IGF (*insulin-like growth factor*) et PDGF (*platelet-derived growth factor*) (Meier, *et al.*, 1999). Ces facteurs autocrines entraînent les SCP restants dans la voie de différenciation du lignage mélanocytaire (Adameyko, *et al.*, 2009). Lorsque les SCP se multiplient, le signal NRG1 devient limitant du fait d'une compétition entre les cellules pour le contact nerveux. Parallèlement, ces précurseurs se différencient en cellules de Schwann immatures, ce qui entraîne une augmentation du niveau des facteurs IGF/PDGF. La balance entre NRG1 et IGF/PDGF est alors inversée en faveur de la détermination des SCP en mélanocytes (**Figure 26**).



Chez le poisson zèbre, l'existence d'une autre population de précurseurs de mélanoblastes que celle issue de la voie dorso-latérale avait déjà été postulée. Yang et Johnson avaient fait l'hypothèse que les cellules de la crête neurale du poisson zèbre engendrent non seulement des mélanoblastes et des mélanocytes matures, mais aussi une population de cellules quiescentes de réserve. Cette population peut être recrutée pendant le développement larvaire ou la régénération d'un organe, par un mécanisme encore inconnu, pour donner de nouveaux mélanocytes (Yang and Johnson, 2006). L'identité de ces cellules, appelées « cellules souches de mélanocytes » est inconnue. Il a tout de même été découvert que les mélanocytes nouvellement régénérés proviennent d'une source non identifiée de cellules non pigmentées, dont le processus de différenciation dépend du récepteur ErbB (Hultman, *et al.*, 2009). Chez le poisson zèbre adulte, il existe donc des mélanocytes qui ont une origine indépendante des dérivés connus de la crête neurale. Ils sont déjà présents dans la larve et leur différenciation dépend de la voie de signalisation NRG. Les SCP, identifiés comme une source de mélanocytes par Adameyko et collaborateurs, ressemblent aux cellules souches de mélanocytes, dont l'existence est postulée chez le poisson zèbre (Adameyko and Lallemand, 2010).

Jusque récemment, il était admis que les mélanoblastes de la voie dorso-latérale migrent sous l'épiderme pour coloniser l'ensemble de la peau du tronc et des membres, puis traversent la lame basale de l'épiderme. Beatrice Mintz avait fait l'hypothèse que les mélanocytes du pelage étaient des dérivés clonaux d'un petit nombre de mélanoblastes et que tous les clones proliféraient de façon identique pour générer une bande de couleur transversale du pelage. Ces travaux suggéraient également un faible mélange des cellules entre les différents clones (Mintz, 1967). Une étude plus récente, basée sur l'analyse clonale d'une seule cellule dérivée de crête neurale marquée génétiquement (Wilkie, *et al.*, 2002), montre qu'il existe un grand nombre de progéniteurs de mélanoblastes et que les mélanocytes dérivés d'un progéniteur forment une grappe de cellules dans la peau. Ces grappes de mélanocytes s'entremêlent les unes avec les autres, entraînant souvent un chevauchement des mélanocytes descendants de différents progéniteurs (Wilkie, *et al.*, 2002). Wilkie et collaborateurs ont déduit de ces résultats qu'il existe trois phases au cours du développement des mélanoblastes : une première phase de prolifération, suivie d'une phase de migration sans



**Figure 27 : Illustration conceptuelle des origines cellulaires et des voies de migration des mélanocytes**  
(Ernfors, 2010)

- A-** La voie de migration dorso-latérale telle qu'elle est décrite originellement. Les cellules dérivées de crête neurale adoptent un destin mélanocytaire peu après la délamination, puis migrent sans proliférer sous l'épiderme vers une zone précise de la peau. Une fois la destination atteinte, les mélanocytes prolifèrent massivement.
  - B-** Les SCP comme précurseurs de mélanocytes. Les cellules dérivées de crête neurale qui n'empruntent pas la voie neuronale s'associent avec les nerfs et adoptent la voie des SCP. Ceux-ci migrent le long des ramifications dorsale et ventrale vers des localisations cutanées. Aux extrémités nerveuses, quelques SCP empruntent une destinée mélanocytaire : ils se détachent des nerfs et se dirigent vers la peau où ils prolifèrent massivement.
- SC (spinal cord) pour moëlle épinière ; NCCs (neural crest derived cell types) pour cellules dérivées de la crête neurale.

prolifération et enfin une seconde phase de prolifération des cellules dispersées dans le corps (**Figure 27A**).

Les SCP ont été identifiés comme un deuxième groupe de précurseurs de mélanoblastes présents au niveau du système nerveux périphérique (Adameyko, *et al.*, 2009). Tout en proliférant, les SCP migrent le long des nerfs nouvellement formés. Certains se détachent de l'extrémité des nerfs pour donner naissance à des mélanoblastes (**Figure 27B**). Il a été montré que la majorité des mélanocytes de l'épiderme et des follicules pileux, du tronc et des membres, dérivent des SCP (Adameyko, *et al.*, 2009). De plus, des analyses histologiques ont révélé que les souris perdent un grand nombre de mélanoblastes dérivés de la voie dorso-latérale peu après que la migration ait débuté (Adameyko, *et al.*, 2009). Ces résultats montrent une faible contribution des mélanoblastes de la voie dorso-latérale dans la production des mélanocytes de l'épiderme et des follicules pileux de la souris. Wilkie et collaborateurs avaient observé que la différenciation des mélanocytes avait lieu en grappes. Ils expliquaient ce phénomène par l'absence ou la quasi-absence de prolifération des mélanoblastes pendant leur migration. L'existence des SCP en tant que précurseurs des mélanocytes est finalement la raison probable de ce phénomène de différenciation en grappes. En effet, les SCP migrent et prolifèrent le long des nerfs qui se ramifient jusque dans la peau des différentes régions du corps (**Figure 27B**). L'ensemble des mélanocytes, différenciés à partir d'un seul SCP situé dans une région de la peau, peut donc former une grappe de cellules.

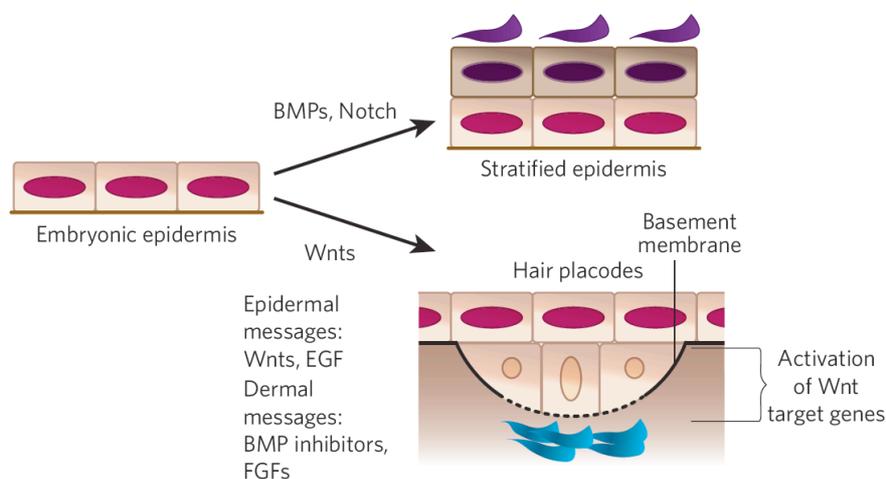
Chez la souris, des mélanoblastes migrent dans toutes les régions du corps, entre E14,5 et E16, puis ils se déplacent vers les bourgeons des follicules pileux pour coloniser la couche la plus interne de la matrice du bulbe pileux (Hirobe, 1992, 1995; Jordan and Jackson, 2000). Des expériences de culture de peau de souris transgéniques *Dct-lacZ* qui expriment le rapporteur *LacZ* dans les mélanoblastes et les mélanocytes (voir plus bas) ont permis de montrer que la migration des mélanoblastes est liée à la formation du follicule pileux (Jordan and Jackson, 2000). La colonisation du follicule pileux par les mélanoblastes est un mécanisme actif qui fait intervenir un chimio-tactisme positif vers la base du follicule. Lorsque les mélanoblastes ont colonisé les follicules pileux, leur développement est étroitement lié à celui du follicule pileux.



## 1.2 Développement du follicule pileux et cycle de vie

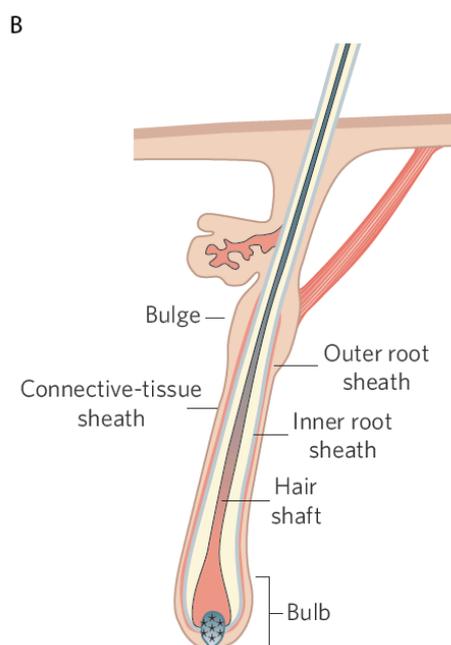
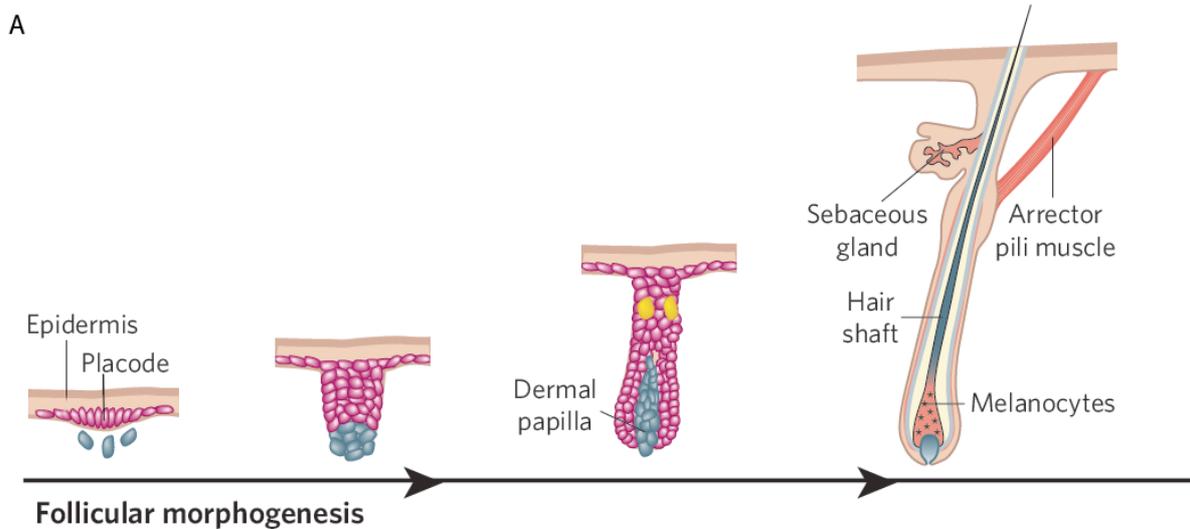
### 1.2.1 Morphogenèse du follicule pileux

Le développement du follicule pileux dépend d'une série de signaux échangés entre les cellules mésenchymateuses du derme et les cellules épithéliales de l'ectoderme (Fuchs, *et al.*, 2001; Hardy, 1992; Schmidt-Ullrich and Paus, 2005). Le derme contrôle également l'espace inter-folliculaire (Oro and Scott, 1998; Zhang, *et al.*, 2008). Chez la souris, un signal dermique induit à E14,5 la formation de la placode dans l'épiderme. L'épiderme s'épaissit et s'invagine dans le derme sous-jacent pour former le bourgeon épidermique du follicule pileux. Plusieurs familles de signaux moléculaires participent à l'induction de la placode. Parmi eux, l'ectodysplasine-A (EDA), appartenant à la famille du TNF (*tumor necrosis factor*), favorise la formation de la placode via son récepteur EDAR et le facteur NFκB (*nuclear factor kappa B*). Les membres de la famille BMP (*bone morphogenic protein*) inhibent la formation de la placode en entrant en compétition avec la signalisation de l'EDA. Parallèlement à ces signaux dermiques, les signaux ectodermiques Wnt/β-catenine entraînent la prolifération des cellules épidermiques vers le derme pour former la placode (**Figure 28**).



**Figure 28 : Etapes de signalisations précoces pour la spécification de la peau de l'embryon**  
(Fuchs, 2007)

Au cours du développement, l'épiderme embryonnaire composé d'une seule couche de cellules exprime les signaux Wnt. Quelques cellules ne répondent pas aux signaux Wnt et deviennent des cellules épidermiques via les signaux BMP, EGF et Notch. Les cellules qui répondent aux signaux Wnt reçoivent de manière sous-jacente par le mésenchyme les signaux d'inhibition FGF et BMP. Ceux-ci commandent aux cellules de former un appendice. Ensemble, les signaux d'inhibition BMP et les signaux d'activation Wnt favorisent la création de la placode. Les signaux sous-jacents provenant du derme ordonnent à la placode de former le follicule pileux.



**Figure 29 : La morphogénèse du follicule pileux**  
(d'après (Fuchs, 2007))

- A-** La morphogénèse du follicule pileux commence par la formation de la placode. Celle-ci entraîne la condensation des cellules dermiques pour générer la papille dermique qui sera associée au follicule pileux de façon permanente. Une fois formée, la papille dermique favorise la prolifération rapide des cellules épithéliales de la placode et induit la croissance du bulbe vers le derme et l'hypoderme pour former le follicule pileux.
- B-** Structure du follicule pileux complètement différencié. Il est constitué en partie de compartiments d'origine dermique : la papille dermique et la gaine conjonctive (« connective-tissue sheath »). Les autres compartiments sont d'origine épidermique : la tige pileuse (« hair shaft »), la gaine épithéliale interne (« inner root sheath ») et la gaine épithéliale externe (« outer root sheath ») dans laquelle le bulge est situé. Le bulbe (« bulb ») est la région située à la base du follicule pileux.

*Sonic hedgehog (Shh)* est un des premiers gènes exprimés dans la placode. *Shh* joue un rôle crucial dans l'organisation des cellules dermiques en permettant leur condensation pour donner une papille dermique qui sera associée au follicule pileux de façon permanente (pour revue, voir (Fuchs, 2007)). La papille dermique une fois formée favorise la prolifération rapide des cellules épithéliales de la placode et induit la croissance du bulbe vers le derme et l'hypoderme pour générer le follicule pileux différencié (**Figure 29A**). Ces cellules épidermiques ayant une importante capacité mitotique forment la matrice du bulbe pileux. Les cellules filles de la tige pileuse sont poussées vers l'extrémité distale du follicule pileux en formation. Les cellules de la matrice se différencient pour donner naissance aux différents types cellulaires du follicule pileux. Le type cellulaire produit dépend de la localisation de la cellule dans le bulbe (Legue and Nicolas, 2005).

Sur la partie supérieure du follicule pileux, des renflements apparaissent : le renflement le plus superficiel correspond à l'ébauche des glandes sébacées et le plus profond correspond à l'insertion du muscle arrecteur du poil. Ce muscle s'insère de façon oblique sur la face postérieure, de sorte qu'il redresse le poil lorsqu'il se contracte.

### 1.2.2 Structure du follicule pileux

Une fois différencié, le follicule pileux possède une structure complexe, composée de compartiments parfaitement individualisés, les uns d'origine dermique (la papille dermique et la gaine conjonctive), les autres d'origine épidermique (la matrice, les gaines épithéliales interne et externe, la tige pileuse et la glande sébacée) (**Figure 29B**).

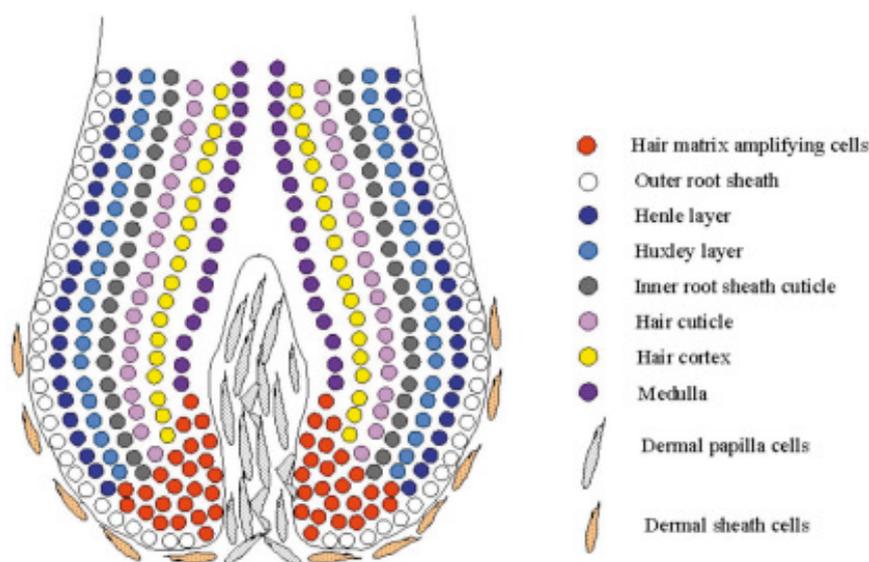
#### ➤ Les compartiments d'origine dermique

La **papille dermique** est une région constituée de tissu conjonctif qui s'invagine à l'intérieur de la matrice qui l'enveloppe. L'interaction physique entre la papille dermique et la matrice est l'élément moteur de la croissance du poil. Cette masse ovoïde peu dense sécrète une matrice extracellulaire synthétisée par les fibroblastes. La **gaine conjonctive** est située dans la continuité de la papille dermique et enveloppe le follicule pileux. Cette gaine est constituée d'un tissu conjonctif composé de fibroblastes et de matrice extracellulaire. Une **membrane basale** sépare le compartiment dermique du compartiment épidermique.

➤ Les compartiments d'origine épidermique

La **matrice** est une zone germinative située à la base du bulbe qui entoure la papille dermique. Elle est le siège d'une intense activité mitotique. Les cellules de la matrice forment la tige pileaire et la gaine épithéliale interne.

La **gaine épithéliale interne**, ou IRS (pour *inner root sheath*), résulte de la prolifération et de la différenciation des cellules de la matrice. Elle entoure la tige pileaire pour la guider de la matrice du bulbe vers l'épiderme. Elle s'étend jusqu'au point de jonction avec la glande sébacée. Elle comporte de l'intérieur vers l'extérieur trois couches concentriques : la cuticule, la couche de Huxley et la couche de Henle (**Figure 30**) (Porter, 2003). La kératinisation de la gaine épithéliale interne s'accompagne de la mort des cellules au niveau de la jonction avec la glande sébacée, ce qui libère la tige pileaire des parois folliculaires (Tobin, *et al.*, 2002).



**Figure 30 : Schéma d'une coupe de bulbe de follicule pileux**  
(Rogers, 2004)

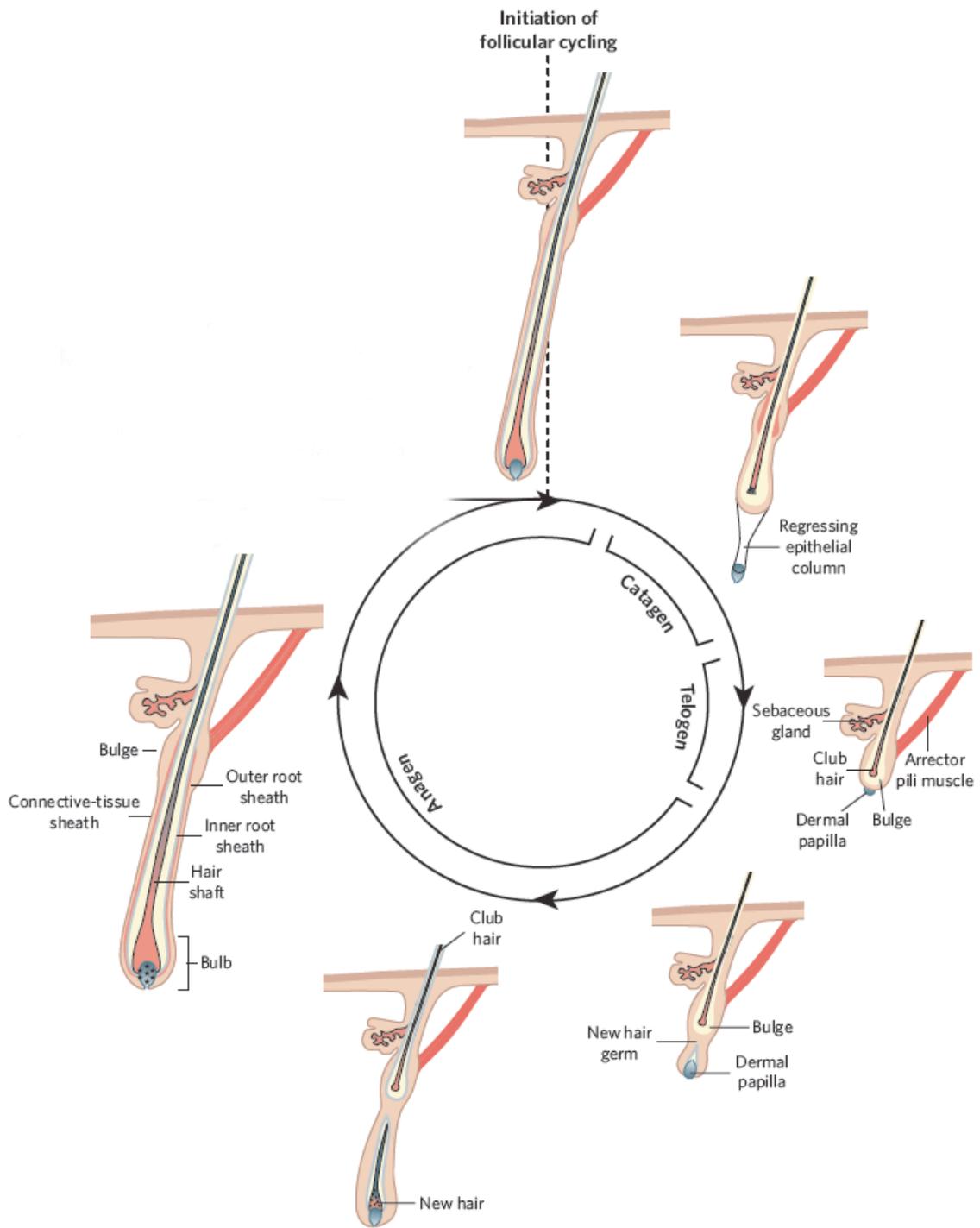
Dans la région du bulbe, les cellules de la matrice qui entourent la papille dermique prolifèrent activement (rouge). Les cellules filles forment d'une part la tige pileaire constituée de la médulla (violet), du cortex (jaune) et la cuticule (mauve), d'autre part la gaine épithéliale interne constituée d'une cuticule (gris foncé), de la couche de Huxley (bleu clair) et de la couche de Henle (bleu foncé). La gaine épithéliale interne est entourée par la gaine épithéliale externe (blanc). Par ailleurs, les cellules d'origine dermique, de forme allongée, représentent la papille dermique (gris clair) et la gaine conjonctive (rose).

La **tige pileaire** comporte de l'intérieur vers l'extérieur trois couches concentriques : la médulla, le cortex et la cuticule (**Figure 30**) (Porter, 2003). Les cellules du cortex synthétisent des kératines et des protéines d'agrégation qui forment un ciment intercellulaire. On y trouve également les grains de mélanine responsables de la couleur du poil. La cuticule superficielle protège essentiellement la fibre pileaire et maintient son intégrité.

La **gaine épithéliale externe**, ou ORS (pour *outer root sheath*), apparaît au cours de la morphogenèse du follicule pileux et semble issue de l'invagination de l'épithélium de surface, avec lequel elle est en continuité. Le « bulge » est un épaissement de la gaine épithéliale externe, situé au point d'insertion du muscle arrecteur du poil. Le **bulge** contient les cellules responsables de la reconstruction d'un nouveau follicule pileux (Cotsarelis, *et al.*, 1989). Il est la niche anatomique des cellules souches des kératinocytes (Oshima, *et al.*, 2001) et des mélanocytes (Nishimura, *et al.*, 2002). Les mélanocytes matures, responsables de la pigmentation du poil, sont localisés dans le bulbe des follicules pileux anagènes. Ils reposent sur la lame basale qui entoure la papille dermique. Ces cellules synthétisent la mélanine au sein de vésicules, appelées mélanosomes. Les dendrites des mélanocytes permettent le transfert des mélanosomes vers les kératinocytes pré-corticaux qui formeront le futur cortex de la tige pileaire.

### 1.2.3 Cycle de vie du follicule pileux

Le follicule pileux est un mini-organe producteur de poils. Les poils poussent de façon cyclique. Chaque cycle comprend trois phases. Pendant la phase de croissance, ou anagène, le follicule pileux croît et le poil est formé à la fin de cette phase. Pendant la phase de régression, ou catagène, le follicule s'involue et le poil meurt. La phase télogène correspond à une période de repos du follicule (Chase, 1954; Paus and Cotsarelis, 1999; Stenn and Paus, 2001; Sundberg, *et al.*, 2005). Seule la partie profonde du follicule pileux subit des phases de croissance et de régression. Chez une jeune souris, un cycle complet dure environ un mois (Muller-Rover, *et al.*, 2001). Chez les souris âgées, la phase télogène est plus longue, si bien qu'un cycle peut durer plusieurs mois (Chase, 1954).



**Figure 31 : Cycle de vie du follicule pileux**  
(Fuchs, 2007)

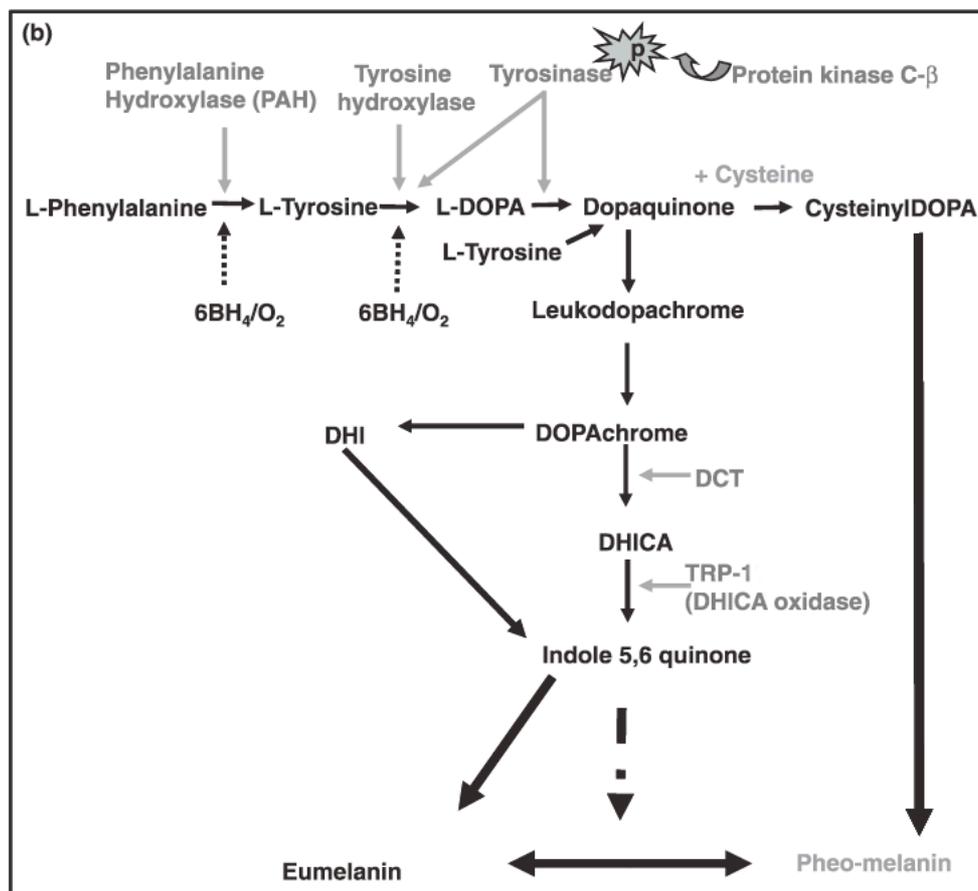
La **phase anagène** est caractérisée par une activité mitotique intense des kératinocytes dans le bulbe du poil (matrice et ORS). Elle est subdivisée en six stades. Les stades I à III correspondent aux phases de développement du bulbe grâce à une forte activité mitotique dans la région du germe, située sous le poil précédent qui a la forme d'un club de golf (voir ci-dessous). Le jeune bulbe croît autour de la papille dermique. Au stade III, le bulbe est complètement formé et la papille se resserre à sa base. A ce moment, les mélanocytes se placent au-dessus de la papille et développent des dendrites et produisent des grains de mélanine. Lors des phases IV et V, la tige pileuse se forme mais est contenue dans les gaines internes et externes. La tige pileuse de l'épiderme émerge au stade VI (Muller-Rover, *et al.*, 2001).

Au cours de la **phase catagène**, les cellules de la matrice cessent de proliférer brusquement. Il s'agit d'un processus contrôlé de différenciation cellulaire et d'apoptose. La résorption de la partie transitoire du follicule pileux se traduit par une rétractation du bulbe vers le derme et par une diminution de la taille de la papille dermique. Les cellules de la gaine épithéliale externe situées au niveau du bulge se kératinisent intensément.

Pendant la **phase télogène**, le follicule pileux est au repos et atteint sa taille minimale dans le derme. Il s'agit de la partie permanente du follicule. La kératine compacte au niveau du bulge retient le poil mort qui reste attaché au follicule par une structure qui rappelle celle d'un club de golf (en anglais *club hair*). Pendant la transition télogène/nouvelle anagène, des cellules souches quiescentes localisées près de la papille dermique sont activées pour produire une nouvelle tige pileuse (Blanpain, *et al.*, 2004; Tumber, *et al.*, 2004). Le nouveau poil en croissance pousse l'ancien qui tombe en fin d'anagène IV (Muller-Rover, *et al.*, 2001). Les cellules souches du bulge produisent des cellules à fort index mitotique qui sont les futures cellules de la matrice.

Le cycle de vie du follicule pileux est illustré sur la **Figure 31**.

Les mélanosomes sont des lysosomes modifiés, spécifiques des mélanocytes (Orlow, 1995). La mélanogénèse est la voie de biosynthèse des pigments de mélanine (**Figure 32**). Elle a lieu exclusivement dans les mélanosomes. Deux types distincts de mélanines sont produits au cours de cette biosynthèse : les eumélanines (pigments bruns et noirs) et les phéomélanines (pigments rouges et jaunes).



**Figure 32 : Voie de biosynthèse des mélanines**  
(d'après (Tobin, 2008))

DHICA : acide 5,6-dihydroxyindole-2-carboxylique ; DOPA : dihydroxyphenylalanine

## 2 Expression et régulation du gène *Dct*

### 2.1 La DOPAchrome tautomérase, une enzyme de la famille des Tyrosinases

#### 2.1.1 *La protéine Tyrosinase*

Le gène *tyrosinase* (*Tyr*) est localisé sur le chromosome 7 de la souris, au locus *albinos*. La tyrosinase est une métalloprotéine à cuivre, membranaire. La région transmembranaire permet l'intégration de la protéine dans le mélanosome. La tyrosinase est l'enzyme qui initie la mélanogenèse. Cette enzyme est limitante pour l'ensemble de cette voie métabolique. L'allèle le plus connu au locus *Tyrosinase* (*Tyr*) est la mutation *albinos* (*Tyr<sup>c</sup>*). Les souris homozygotes pour cette mutation sont blanches aux yeux rouges, car elles ne produisent pas de mélanine.

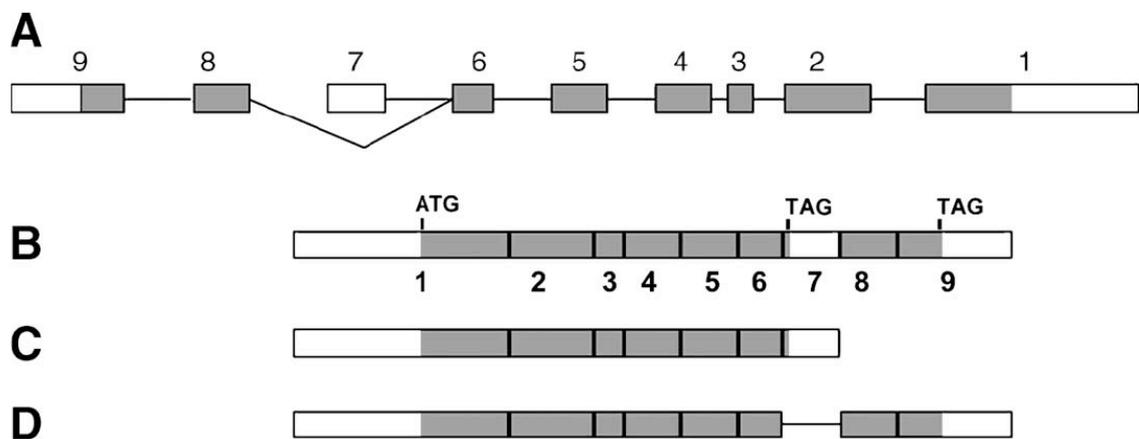
Différentes équipes ont utilisé un anticorps de lapin pour détecter la protéine correspondante chez la souris. Cet anticorps était en fait capable de reconnaître deux autres protéines présentes dans des extraits cellulaires de mélanocytes. Ces trois protéines partagent 40% d'identité en acides aminés (Jackson, 1988; Jackson, *et al.*, 1992; Kwon, *et al.*, 1987; Shibahara, *et al.*, 1986). Les protéines ainsi découvertes ont été nommées TRP1 et TRP2 pour *Tyrosinase-Related Protein 1 et 2*.

#### 2.1.2 *La protéine TRP1*

Le gène *Tyrp1* est localisé sur le chromosome 4 de la souris, au locus *brown*. La protéine TRP1 est une métalloprotéine membranaire. Chez la souris, TRP1 possède une faible activité dopachrome tautomérase (Winder, *et al.*, 1993) et une forte activité DHICA oxydase. Elle oxyde le DHICA en carboxyl-indole quinone (Kobayashi, *et al.*, 1994). Chez l'homme, TRP1 ne possède pas d'activité DHICA oxydase (Boissy, *et al.*, 1998). TRP1 pourrait jouer un rôle dans la prolifération et la morphologie des mélanocytes. Elle pourrait aussi modifier l'activité de la tyrosinase (Li, *et al.*, 2004).

#### 2.1.3 *La protéine TRP2 ou DCT*

Le gène *Tyrp2* est localisé sur le chromosome 14 de la souris, au locus *slaty* (Jackson, *et al.*, 1992). La protéine TRP2 est une métalloprotéine membranaire. TRP2 possède une forte activité DOPAchrome tautomérase. Elle catalyse l'isomérisation de la DOPAchrome en DHICA (Tsukamoto, *et al.*, 1992). Spontanément, la DOPAchrome se transforme en DHI.



**Figure 33 : Représentation schématique du gène *Dct* et de ses transcrits**  
(Jiao, *et al.*, 2006)

- A- Schéma de la structure des exons du gène *Dct* dans le chromosome 14 de la souris.
- B- Schéma du transcrit *Dct* en entier.
- C- L'isoforme du transcrit *Dct* qui contient l'exon 7 mais pas les exons 8 et 9 est présente dans le RPE.
- D- L'isoforme du transcrit *Dct* qui ne contient pas l'exon 7 est présente dans les mélanocytes.

Pour cette raison, la protéine TRP2 et le gène *Tyrp2* ont été rebaptisés DCT (DOPAchrome tautomérase) et *Dct*. Le DHI est rapidement métabolisé en indole-quinone puis en mélanine. Il est donc fortement toxique pour le mélanocyte. Le DHICA est beaucoup plus stable et il est incorporé lentement dans le polymère de mélanine. La protéine DCT, en favorisant la formation du DHICA, minimise les effets cytotoxiques de la mélanogenèse (Tsukamoto, *et al.*, 1992). Il existe trois mutations au locus *Dct* (Budd and Jackson, 1995). Les souris mutantes *slaty* produisent un pigment d'eumélanine de couleur gris foncé/marron plutôt que noir, dû à une activité DOPAchrome tautomérase significativement plus faible que chez les souris sauvages (Jackson, *et al.*, 1992).

## 2.2 Le profil d'expression du gène *Dct* en fait un bon marqueur du lignage mélanocytaire

Chez la souris, le gène *Dct* est localisé sur le chromosome 14 et est composé de 9 exons. La séquence codante de ce gène est située sur le brin complémentaire, soit dans la direction téloméro-centrique. Il existe deux isoformes du transcrit *Dct* qui résultent d'un épissage alternatif de l'exon 7 (Figure 33). Le transcrit présent dans les mélanocytes ne contient pas l'exon 7, tandis que le transcrit des cellules du RPE contient l'exon 7 mais pas les exons 8 et 9 (Jiao, *et al.*, 2006).

### *2.2.1 Expression attendue dans les mélanoblastes et les mélanocytes*

Par des expériences d'hybridation *in situ* sur des embryons de souris de stades tardifs et des nouveaux-nés, il a été montré qu'une sonde spécifique à l'ARN messager (ARNm) du gène *Dct* s'hybride fortement avec toutes les cellules contenant de la mélanine (Steel, *et al.*, 1992). Ces résultats ont été confirmés par des expériences d'immuno-histochimie avec un anticorps dirigé contre la protéine DCT (Pavan and Tilghman, 1994). L'analyse histologique classique ne permet pas de distinguer les mélanoblastes des cellules du mésenchyme du fait de l'absence de mélanine. La sonde dirigée contre l'ARNm de *Dct*, appelée sonde *Dct*, a permis d'identifier les mélanoblastes dès le stade E10. Les cellules exprimant l'ARNm *Dct* ont été identifiées comme les mélanoblastes pour leurs caractéristiques : leur nombre, leur localisation et le fait qu'à partir du stade E16,5 la sonde *Dct* soit spécifique aux mélanocytes pigmentés (Steel, *et al.*, 1992). La sonde *Dct* et l'anticorps anti-DCT ne marquent pas toutes les cellules dérivées de la crête neurale, car des dérivés de la crête neurale, comme les cellules

mésenchymateuses de la mâchoire inférieure, n'expriment pas DCT (Steel, *et al.*, 1992). Steel et collaborateurs ont conclu que la protéine DCT est spécifique des mélanoblastes et des mélanocytes.

Au stade E9,5, la sonde *Dct* s'hybride avec la vésicule optique et l'épithélium présomptif pigmentaire de la rétine. Au fur et à mesure du développement de l'œil, entre E9,5 et E11,5, l'hybridation avec la sonde *Dct* est de plus en plus intense. À E11,5, la sonde *Dct* s'hybride sur toute la largeur du RPE. L'expression de *Dct* est détectée dans le mésenchyme postérieur à l'oreille interne à partir de E10, ce qui correspond à la détection la plus précoce de mélanoblastes en migration. L'expression de *Dct* au niveau du canal cochléaire et de la stria vascularis augmente progressivement jusqu'au stade E16,5 (Steel, *et al.*, 1992). À partir de E12, les mélanoblastes sont détectés par hybridation avec la sonde *Dct* dans le derme et le mésenchyme sous-cutané. Cependant leur distribution est inégale et de nombreuses régions de la peau sont dépourvues de mélanoblastes. Aux stades plus tardifs, les mélanocytes sont essentiellement localisés dans des follicules pileux, mais beaucoup sont encore observés dans le derme (Steel, *et al.*, 1992).

Les résultats d'immuno-histochimie montrent que la protéine DCT est spécifique des mélanocytes interfolliculaires chez l'embryon à E18,5 et des mélanocytes des follicules pileux d'un souriceau de 7 jours (Pavan and Tilghman, 1994). Les sites d'expression de la protéine DCT chez l'embryon tels qu'ils sont révélés par immuno-histochimie, reflètent la distribution de l'ARNm *Dct* observée par hybridation *in situ* (Pavan and Tilghman, 1994). Les résultats obtenus par immuno-histochimie révèlent également une distribution qui n'est pas uniforme des cellules positives pour DCT : les cellules DCT+ sont observées le long de l'axe antéro-postérieur, spécialement au niveau de la tête, des membres postérieurs et de la base de la queue. Au contraire, avant le stade E15,5, très peu de mélanoblastes sont observés dans la partie du tronc située entre les membres. De même, la distribution dorso-ventrale n'est pas uniforme (Pavan and Tilghman, 1994). Ces observations sur la distribution non-uniforme des cellules DCT+ rejoignent celles de Steel et collaborateurs utilisant l'hybridation *in situ*.

Dans la peau des souris adultes, DCT est la seule enzyme de la famille des tyrosinases à être exprimée dans les cellules souches des mélanocytes situées dans le bulge du follicule pileux (Botchkareva, *et al.*, 2001; Nishimura, *et al.*, 2002). Contrairement aux gènes *Tyr* et *Tyrp1*, *Dct* est toujours exprimé dans les mélanomes et même dans les tumeurs amélanotiques (Orlow, 1995; Pak, *et al.*, 2001).

### 2.2.2 Expression inattendue dans le télencéphale

La protéine DCT est également exprimée dans la partie frontale du télencéphale à partir du stade E10,5. Ce marquage, détectable jusqu'à E14 est surprenant car les cellules DCT+ du télencéphale ne sont pas destinées à devenir des mélanocytes (Steel, *et al.*, 1992). L'expression du gène *Dct* dans le cerveau a été étudiée chez la souris transgénique *Dct-LacZ* (Mackenzie, *et al.*, 1997). Au cours du développement, l'expression du rapporteur *LacZ* est localisée dans la partie dorsale du télencéphale entre E10,5 et E12,5. Entre E15,5 et E17,5, l'expression du rapporteur s'est déplacée dans la zone ventriculaire, où sont situées les cellules souches neuronales (Jiao, *et al.*, 2006). Chez l'adulte, le niveau d'expression du gène *Dct* diminue fortement. Chez les souris *Dct-LacZ*, la  $\beta$ -galactosidase est observable dans le cortex, le septum, la zone subventriculaire, le gyrus denté et le bulbe olfactif. Des expériences d'immuno-histochimie chez la souris âgée de 15 jours (à P15) corroborent ces résultats (Jiao, *et al.*, 2006). L'expression de *Dct* dans le cerveau change donc dans l'espace et dans le temps.

Des expériences de RT-PCR ont été réalisées sur des cDNA issus de télencéphales de souris embryonnaires, de mélanocytes et de télencéphales de nouveau-nés de souris (Jiao, *et al.*, 2006). L'isoforme du transcrit *Dct* présente dans le télencéphale est la même que celle présente dans les mélanocytes (Figure 33D), c'est-à-dire qu'elle est dépourvue de l'exon 7 (Jiao, *et al.*, 2006).

Des expériences d'inhibition de l'expression du gène *Dct* dans des cellules en culture dérivées du télencéphale d'un embryon à E13,5 ont été réalisées avec des siRNA (*small-interfering RNA*). L'inhibition de *Dct* entraîne une diminution du nombre de cellules d'environ 25%. Au contraire, la surexpression de *Dct* dans des cellules de la zone subventriculaire du cerveau induit une augmentation du nombre de cellules de 260% (Jiao, *et al.*, 2006). Le facteur de transcription SOX10 est exprimé dans les cellules de la crête neurale, alors qu'il n'est plus exprimé dans les progéniteurs des cellules non-gliales (Kim, *et al.*, 2003). Par ailleurs, SOX10 active l'expression du gène *Dct* dans les mélanoblastes et les mélanocytes en interagissant directement avec son promoteur (Jiao, *et al.*, 2004; Potterf, *et al.*, 2001). SOX10 intervient également dans le développement du système nerveux périphérique (Britsch, *et al.*, 2001). Dans des progéniteurs neuronaux, la surexpression de *Sox10* augmente le niveau d'expression de *Dct*, tandis que l'inhibition de *Sox10* le diminue d'une manière dose-dépendante (Jiao, *et al.*, 2006). L'ensemble des données obtenues par Jiao et collaborateurs montre que l'expression spatio-temporelle du gène *Dct* est liée à la



**Figure 34 : Hybridation *in situ* sur un embryon de souris entier à E11,5**  
(Baxter and Pavan, 2002)

La sonde *Dct* s'hybride avec l'épithélium pigmentaire de la rétine, le télencéphale, les mélanoblastes en migration et les ganglions dorsaux.

	TRP-2	TRP-1	Tyrosinase	Pigment
PRE	9.5	11.5	13.5	11.5
Melanoblasts	10	14.5	14.5	16.5
Telencephalon	10.5	Never	Never	Never

The times are given in days post coitum and are nominal ages taken from the equivalent stages of development in Theiler (1989).

**Tableau 2 : Stades d'apparition des marqueurs des mélanocytes dans l'épithélium pigmentaire de la rétine, les mélanoblastes en migration et le télencéphale**  
(Steel, *et al.*, 1992)

PRE ou RPE (en anglais « pigment retinal epithelium ») pour épithélium pigmentaire de la rétine.

neurogenèse du cortex cérébral. *Dct* joue donc un rôle dans la prolifération des progéniteurs neuronaux, en plus de son implication dans la biogenèse de la mélanine (Jiao, *et al.*, 2006).

Une expérience d'hybridation *in situ* sur l'embryon *in toto* à E11,5 illustre l'ensemble des résultats décrits dans cette section 2.2 (**Figure 34**) (Baxter and Pavan, 2002). DCT est donc un bon marqueur des mélanoblastes en migration à partir de E10, c'est-à-dire 4 jours avant que les sondes dirigées contre les ARNm codant la tyrosinase et TRP1 ne détectent les mélanocytes et 6 jours avant que la mélanine ne soit produite. Le **Tableau 2** présente les différents stades d'apparition des marqueurs des mélanocytes dans le RPE, le télencéphale et les mélanoblastes.

### 2.3 Régulation du gène *Dct*

Le facteur de transcription MITF (*microphthalmia-associated transcription factor*) est le régulateur majeur de la spécification des mélanoblastes (pour revue voir (Levy, *et al.*, 2006; Steingrimsson, *et al.*, 2004)). MITF est un facteur de transcription de type hélice basique - boucle-hélice leucine zipper (bHLH-ZIP) (Hodgkinson, *et al.*, 1993). MITF est exprimé dans les mélanoblastes juste après leur émergence du tube neural et bien avant leur entrée dans la voie de migration dorso-latérale, ce qui correspond à la mise en place du lignage mélanocytaire. Chez la souris, des mutations au locus *Mitf* empêchent les cellules dérivées de la crête neurale de devenir des cellules DCT positives (Opdecamp, *et al.*, 1997). Les promoteurs des gènes de la famille des tyrosinases possèdent tous un motif M-box (AGTCATGTGCT) dans leur région proximale. Ce motif contient la séquence consensus CANNTG qui est le site de fixation de nombreux facteurs de transcription à structure hélice-boucle-hélice (Massari and Murre, 2000). Le facteur de transcription MITF se lie à ces différents motifs et transactive les promoteurs associés à chacun des sites (Bentley, *et al.*, 1994). MITF régule également l'expression des gènes de la famille des tyrosinases via un motif E-box (Aksan and Goding, 1998).

Des analyses génétiques ont mis en évidence une relation épistatique entre les gènes *Sox10* et *Mitf* : SOX10 régule l'expression du gène *Mitf* en se fixant au promoteur et en l'activant (Lee, *et al.*, 2000b; Verastegui, *et al.*, 2000). SOX10 régule également l'expression du gène *Dct* en se liant directement à son promoteur (Potterf, *et al.*, 2001). SOX10 interagit avec au moins deux motifs consensuels présents dans la région proximale du promoteur du gène *Dct*, définis



par les séquences AACAAAG (-141 ; -135) et GTCAAAG (-292 ; -298) (Ludwig, *et al.*, 2004).

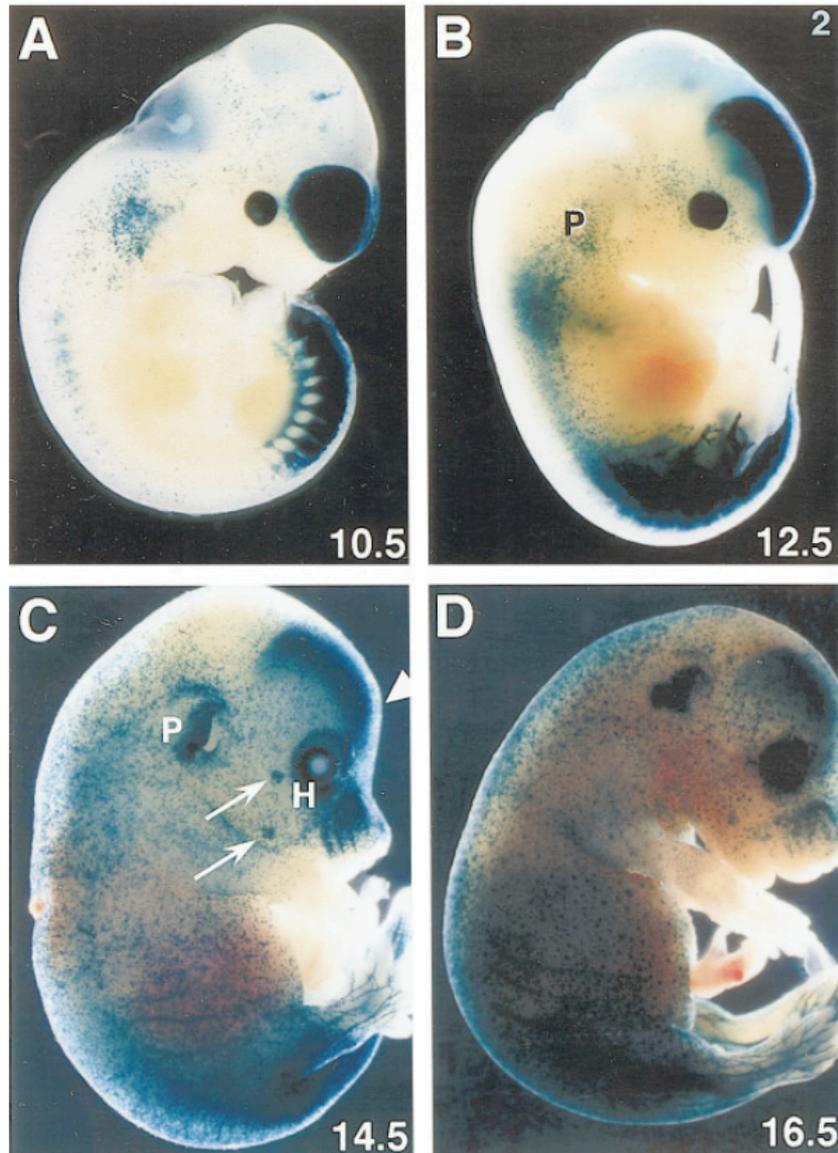
Les promoteurs des gènes *Tyr* et *Tyrp1* sont caractérisés par des éléments de régulation situés loin en 5', à 15 kb en amont du premier exon. Ces séquences sont conservées au cours de l'évolution. Ces éléments de régulation sont nécessaires à l'expression des gènes *Tyr* et *Tyrp1* dans les mélanocytes et le RPE (Giraldo, *et al.*, 2003; Murisier, *et al.*, 2007). Aucune séquence de régulation en position distale n'a été découverte dans le promoteur du gène *Dct* (Murisier, *et al.*, 2008).

Chez l'homme, DCT n'est pas exprimé dans les mélanocytes actifs du bulbe pileux. Cependant DCT est exprimé dans l'épiderme et dans des cultures primaires de mélanocytes de cheveux et de peau (Commo, *et al.*, 2004). Cette carence spécifique du follicule pileux humain résulte probablement d'une régulation transcriptionnelle liée à l'absence de la protéine SOX10 (Commo, *et al.*, 2004). L'absence de la protéine DCT dans les mélanocytes du follicule pileux est surprenante compte tenu du rôle présumé de l'enzyme dans la voie de biosynthèse des mélanines.

### **3 Utilisation du gène *Dct* en transgénèse**

#### **3.1 La souris transgénique *Dct-LacZ***

Une lignée de souris transgéniques portant le gène rapporteur *LacZ* sous le contrôle du promoteur *Dct* a été produite en 1997 (Mackenzie, *et al.*, 1997). Une séquence de 3,4 kb du promoteur *Dct* a été utilisée pour mimer l'expression du gène endogène. L'objectif était d'identifier les premiers mélanoblastes avant l'apparition de pigment. Chez l'embryon, le promoteur *Dct* est actif à partir du stade E9. L'expression du rapporteur *LacZ* est d'abord observé à E9 au niveau de l'œil, puis à E9,5 dans le télencéphale. A E10,5, l'expression de *LacZ* s'étend à l'ensemble du RPE. Le marquage dans le cerveau est observable de E9,5 à E17,5. L'expression du transgène au niveau du bulbe olfactif est détecté entre E14,5 et E16,5 (Mackenzie, *et al.*, 1997). Ces observations sont en accord avec les résultats obtenus par hybridation *in situ* et par immuno-histochimie (Pavan and Tilghman, 1994; Steel, *et al.*, 1992).



**Figure 35 : Coloration X-gal de souris transgéniques *Dct-LacZ***  
(Mackenzie, *et al.*, 1997)

- A-** A E10,5, les mélanoblastes sont clairsemés le long de la voie dorso-latérale du tronc et de la queue. Au niveau de la région crânienne, les mélanoblastes sont présents dans le mésencéphale et dans l'œil.
- B-** A E12,5, les mélanoblastes sont présents dans l'œil et autour. Les mélanoblastes occupent également le mésenchyme latéral à la moelle épinière jusque dans la queue. La partie ventrale de l'embryon ne révèle aucun marquage.
- C-** A E14,5, pratiquement toute la surface de l'embryon est recouverte du marquage X-gal avec une forte concentration du marquage au niveau de l'oreille (P pour pinna). Les vibrisses et les autres follicules pileux au niveau de la tête (indiqués par des flèches) contiennent un grand nombre de mélanoblastes. Une forte concentration de cellules marquées encadre l'œil et la glande de Harder (H). A ce stade, les mélanoblastes commencent à migrer dans les membres postérieurs et antérieurs, mais pas encore dans les pattes.
- D-** A E16,5, les mélanoblastes commencent à envahir les pattes, mais pas encore la surface ventrale des membres.

La coloration X-gal identifie les mélanoblastes après qu'ils aient quitté la crête neurale à partir de E10,5. Ils sont alors situés dans le mésenchyme en position latérale par rapport à la crête neurale (Figure 35A). A E12,5, le nombre de cellules marquées augmente dans la partie caudale de l'embryon. A ce stade, la partie dorsale contient beaucoup plus de cellules *LacZ* positives que la partie ventrale. Des mélanoblastes commencent à migrer dans les membres antérieurs, puis un ou deux jours plus tard, dans les membres postérieurs. Au niveau de la tête, des mélanoblastes s'accumulent autour de l'œil à partir de E12,5. Ils colonisent la choroïde, la glande de Harder, la strie vasculaire de l'oreille interne et la cochlée (Figure 35B). Vers E14,5, des mélanoblastes sont présents sur l'ensemble de la surface de la tête (Figure 35C). A E16,5, seules les surfaces ventrales des membres ne sont pas marquées (Figure 35D). Après E17,5, la coloration X-gal ne peut plus pénétrer la peau de l'embryon. L'expression du transgène est ensuite détectée dans des mélanocytes des follicules pileux chez les souris adultes transgéniques (Mackenzie, *et al.*, 1997).

A partir du stade E10,5, les ganglions dorsaux, du tronc et de la queue, et les nerfs rachidiens de la région caudale sont nettement marqués par la coloration X-gal (Mackenzie, *et al.*, 1997). Pourtant aucune expression de la protéine DCT n'a été détectée par immunohistochimie dans les nerfs de la région caudale, et l'ARNm *Dct* n'est pas observé par hybridation *in situ* entre E11,5 et E15,5 (Mackenzie, *et al.*, 1997). MacKenzie et collaborateurs ont donc considéré que l'expression du transgène *LacZ* dans ces régions était ectopique. A E12,5, le transgène est exprimé dans la tige optique, ce qui n'est pas corroboré par les expériences d'hybridation *in situ*. MacKenzie et collaborateurs en ont conclu que l'expression du transgène *LacZ* chez les souris *Dct-LacZ* s'étend au-delà de l'expression normale du gène *Dct* chez l'embryon. La Figure 34 présente le marquage par hybridation *in situ* avec un ADNc *Dct* d'un embryon entier à E11,5. D'après ce cliché, la sonde s'hybride avec les ganglions dorsaux contrairement à ce qui est décrit par MacKenzie et collaborateurs. Ce résultat suggère que l'expression de *LacZ* au niveau des ganglions dorsaux des souris *Dct-LacZ* n'est pas ectopique. De plus, il a été récemment montré que les SCP sont des précurseurs de mélanocytes. L'expression du gène *Dct* dans des ganglions dorsaux à partir de E10,5 pourrait correspondre à la détermination des SCP en mélanoblastes.

La lignée *Dct-LacZ* a été largement utilisée pour identifier les mélanocytes et leurs précurseurs chez l'embryon et chez l'adulte (Aubin-Houzelstein, *et al.*, 1998; Aubin-Houzelstein, *et al.*, 2008; Delmas, *et al.*, 2007; Jordan and Jackson, 2000; Nishimura, *et al.*, 2002; Schouwey, *et al.*, 2007; Takemoto, *et al.*, 2006). Les résultats présentés dans cette

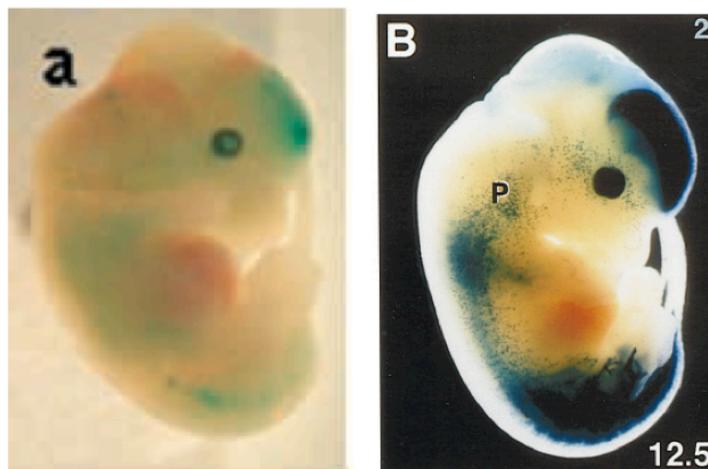
section 3.1 montrent que la lignée *Dct-LacZ* exprime à un niveau élevé le rapporteur *LacZ*. Le niveau élevé d'expression du gène rapporteur *LacZ* chez les embryons et les souris transgénique *Dct-LacZ* s'explique au moins en partie par un grand nombre de copies du transgène (Ian Jackson, communication personnelle).

### 3.2 La souris *Dct<sup>Cre/+</sup>* : knock in et knock out au locus *Dct*

La souris transgénique *Dct-Cre* et la souris *knock in Dct<sup>tm1(Cre)Bee/+</sup>* (appelée *Dct<sup>Cre/+</sup>* dans ce manuscrit) ont été décrites en 2002 (Guyonneau, *et al.*, 2002). Ce sont les premières souris qui expriment la Cre recombinase dans les mélanocytes. Le profil d'expression de la Cre chez ces souris a été étudié par croisement avec des souris des lignées *Rosa26<sup>tm1Sor</sup>* (R26R) et Z/EG. La lignée R26R possède un rapporteur *LacZ* interrompu par une cassette Stop encadrée de sites *loxP*. Dans les cellules où la Cre recombinase est exprimée, la cassette Stop est perdue et le rapporteur *LacZ* est traduit. La lignée R26R exprime donc le rapporteur *LacZ* sous condition d'expression de la Cre. Les souris Z/EG possèdent un gène *LacZ* encadré de sites *loxP* ; lorsque la Cre recombinase est active, le gène *LacZ* est perdu ce qui permet l'expression d'une protéine GFP. Les embryons hybrides *Dct-Cre* × R26R présentent les mêmes marquages par coloration X-gal que ceux observés chez les embryons *Dct-LacZ*. A E12,5, le marquage est présent dans le RPE, les mélanoblastes en migration et le télencéphale (Guyonneau, *et al.*, 2002). Comme pour les embryons *Dct-LacZ*, une expression du transgène *LacZ* est observée au niveau des ganglions et des nerfs rachidiens. Cette expression est alors considérée comme ectopique. Les embryons hybrides *Dct-Cre* × Z/EG ont également été étudiés. Des cellules fluorescentes sont détectées dans le RPE et dans le télencéphale à partir de E10,5. A E12,5, des cellules fluorescentes sont visibles dans la région latérale du tube neural, principalement dans le tronc. Ces cellules correspondent aux mélanoblastes en migration (Guyonneau, *et al.*, 2002). De façon surprenante, aucun marquage GFP n'est détecté dans les ganglions et les nerfs rachidiens (Guyonneau, *et al.*, 2002).

Le gène *Cre* a été intégré à la place du premier exon du gène *Dct* par recombinaison homologue dans des cellules ES. Des souris chimères ont été obtenues après injection des cellules ES recombinées homologues dans des blastocystes. Enfin, suite au croisement des souris chimères avec des femelles C57BL/6J, des souris *knock in Dct<sup>Cre/+</sup>* ont été engendrées. Le profil d'expression de la Cre chez les souris *knock in Dct<sup>Cre/+</sup>* a été étudié (Guyonneau, *et al.*, 2002). Les souris *knock in* ont été croisées avec les souris R26R. Chez les embryons et les

adultes hybrides  $Dct^{Cre/+} \times R26R$  une expression du rapporteur  $LacZ$  a été détectée dans les mélanoblastes et les mélanocytes. Cependant cette expression a été observée plus tardivement (à E12,5) que chez les souris  $Dct-LacZ$  et  $Dct-Cre \times R26R$ , où elle est observée dès E10,5 (Guyonneau, *et al.*, 2002). A E12,5, la coloration X-gal n'est pas identique chez les souris  $Dct^{Cre/+} \times R26R$  et chez les transgéniques  $Dct-LacZ$  (Figure 36). Bien que la coloration bleue permette de visualiser les mélanoblastes/cytes chez les embryons  $Dct^{Cre/+} \times R26R$ , quelques follicules pileux seulement sont colorés en bleu chez les adultes  $Dct^{Cre/+} \times R26R$  (Guyonneau, *et al.*, 2002). Ces résultats inattendus ne sont pas discutés dans l'article de Guyonneau et collaborateurs.



**Figure 36 : Comparaison du profil d'expression de LacZ**  
(Guyonneau, *et al.*, 2002) et (Mackenzie, *et al.*, 1997)

- A- Expression du gène rapporteur  $LacZ$  au stade E12,5 chez les souris  $Dct^{Cre/+} \times R26R$ .
- B- Expression du gène rapporteur  $LacZ$  chez les souris transgéniques  $Dct-LacZ$  au même stade.

Les souris *knock in*  $Dct^{Cre/+}$  ont été croisées entre elles pour donner naissance à des souris homozygotes  $Dct^{Cre/Cre}$  (Guyonneau, *et al.*, 2004). Ces souris n'expriment pas l'ARNm  $Dct$  ni la protéine DCT (Guyonneau, *et al.*, 2004). Les souris homozygotes  $Dct^{Cre/Cre}$  ont un pelage d'une couleur plus claire et diluée que les souris  $Dct^{Cre/+}$  et de génotype sauvage. Ce phénotype est observé sur fonds agouti et nonagouti. Les organes dans lesquels  $Dct$  est normalement exprimé ont été analysés : le cerveau, les yeux et la peau. Aucune différence significative n'a été détectée entre les organes de souris  $Dct^{Cre/Cre}$  et sauvages. D'autres organes qui n'expriment pas  $Dct$  sont normaux. Les conséquences phénotypiques dues à une



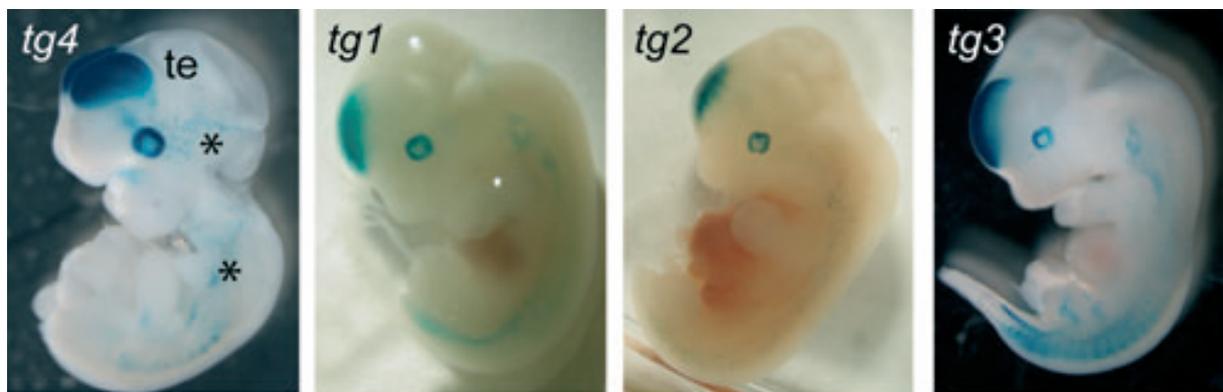
absence d'expression de DCT sont restreintes aux mélanocytes, avec un effet décrit sur la couleur du pelage uniquement (Guyonneau, *et al.*, 2004). Des cultures primaires de mélanocytes ont pu être dérivées de souris homozygotes  $Dct^{Cre/Cre}$ . L'absence d'expression de DCT ne semble affecter ni la viabilité, ni la prolifération de mélanocytes *in vitro* (Guyonneau, *et al.*, 2004).

D'après Jiao et collaborateurs, la protéine DCT joue un rôle dans la prolifération des progéniteurs neuronaux (Jiao, *et al.*, 2006). La souris  $Dct^{Cre/Cre}$  ne présente cependant aucun défaut dans le cerveau ou dans le système nerveux périphérique. Ces résultats sont en contradiction avec l'hypothèse d'un rôle de *Dct* dans les progéniteurs neuronaux. De nombreux facteurs de croissance régulent la prolifération des progéniteurs neuronaux, comme le bFGF (*basic fibroblast growth factor*) et l'EGF (*epidermal growth factor*) (Cairns and Finklestein, 2003; Deleyrolle, *et al.*, 2006; Wang, *et al.*, 2004). D'autres facteurs de croissance peuvent donc compenser l'absence de *Dct* pendant le développement embryonnaire (Jiao, *et al.*, 2006).

Les souris  $Dct^{Cre/Cre}$  ont été comparées aux souris portant la mutation *slaty* ( $Dct^{sl}$ ). Les phénotypes de couleur du pelage sont très similaires. Les doubles hétérozygotes  $Dct^{Cre/sl}$  issus du croisement entre les lignées  $Dct^{Cre/Cre}$  et  $Dct^{sl/sl}$  ne peuvent être distinguée des homozygotes  $Dct^{Cre/Cre}$  et  $Dct^{sl/sl}$  (Guyonneau, *et al.*, 2004). Ce test de complémentation démontre que les deux mutations affectent le même gène et ont des effets similaires sur la synthèse d'eumélanine. Les follicules pileux  $Dct^{Cre/Cre}$  et  $Dct^{sl/sl}$  ont une quantité de mélanine significativement plus faible que les souris de génotype sauvage. Cependant, la quantité relative de mélanines chez les souris  $Dct^{sl/sl}$  est significativement plus faible que chez les souris  $Dct^{Cre/Cre}$  sans que cela soit visible au niveau du pelage (Guyonneau, *et al.*, 2004). Les souris  $Dct^{sl/sl}$  produisent une protéine DCT mutée. Aucune culture primaire de mélanocytes n'a pu être dérivée de souris  $Dct^{sl/sl}$ .

### 3.3 La lignée de souris avec un chromosome artificiel bactérien (BAC) *Dct-LacZ*

Comme nous l'avons vu, les gènes *Tyr* et *Tyrp1* possèdent des éléments de régulation très loin en amont du premier exon. Ces éléments de régulation sont impliqués dans l'expression des gènes *Tyr* et *Tyrp1* dans les mélanocytes et dans le RPE



**Figure 37 : Profils d'expression de *LacZ* chez les embryons de souris BAC *Dct-LacZ***  
(Murisier, *et al.*, 2008)

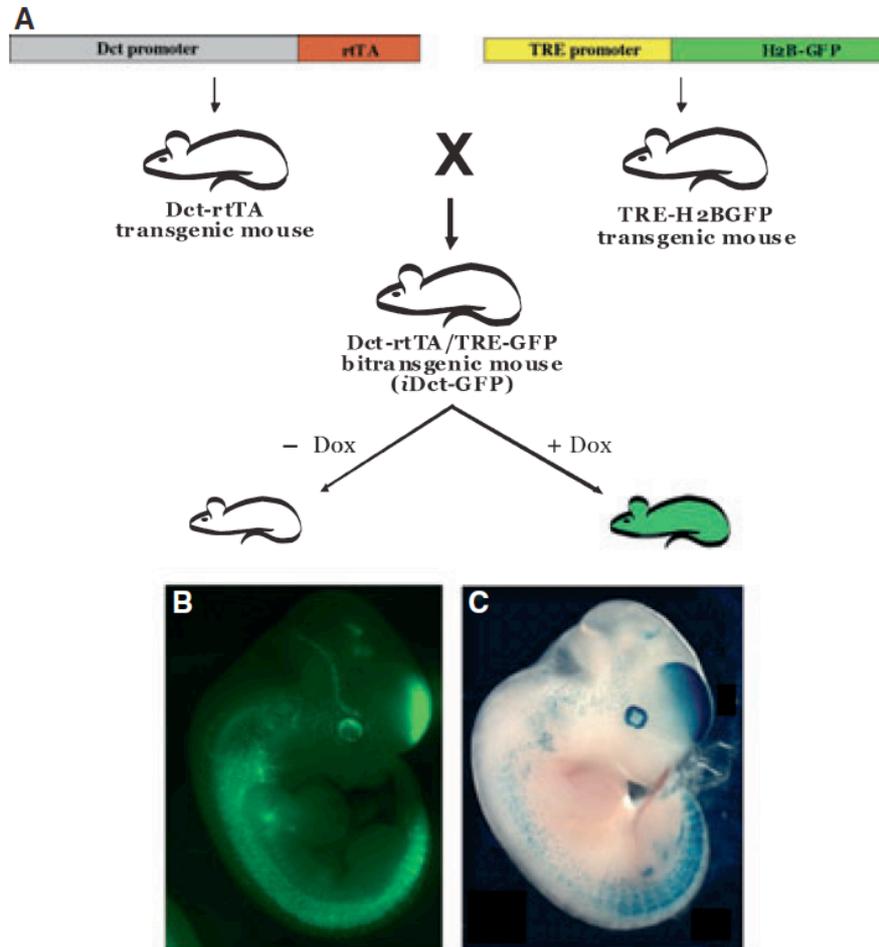
Coloration X-gal d'embryons de stade E10,5 (tg4), E12,5 (tg3) et E13,5 (tg1 et 2). Le gène rapporteur *LacZ* est exprimé dans les mélanoblastes, le RPE et le télencéphale. Dans certaines lignées, *LacZ* est également exprimé dans les ganglions dorsaux.

(Giraldo, *et al.*, 2003; Murisier, *et al.*, 2007). En effet, chez des souris transgéniques *Tyr-LacZ*, construites avec 6 kb du promoteur du gène *Tyr*, le transgène *Tyr-LacZ* n'a pas le profil d'expression du gène *Tyr* endogène dans les mélanoblastes et le RPE. En revanche, chez des souris transgéniques BAC *Tyr-LacZ*, portant un BAC de 186 kb qui comprend le gène *Tyr*, le transgène a le même profil d'expression que celui du gène endogène (Murisier, *et al.*, 2007).

Dans le cas de la souris transgénique *Dct-LacZ*, nous avons vu qu'une séquence de 3,4 kb du promoteur *Dct* est suffisante pour mimer l'expression du gène endogène dans le télencéphale, les mélanoblastes/mélanocytes et le RPE. D'après MacKenzie *et al.*, cette souris transgénique présente néanmoins des expressions ectopiques du transgène dans les nerfs et les ganglions dorsaux (Mackenzie, *et al.*, 1997). Pour apporter un nouvel outil dans l'étude du lignage mélanocytaire, un BAC *Dct* qui comprend 147 kb de séquence du gène *Dct* a été utilisé. Le gène rapporteur *LacZ* a été inséré par recombinaison homologue dans le BAC à la place de l'exon 1 du gène *Dct* pour produire un BAC *Dct-LacZ*. Puis, des souris transgéniques ont été fabriquées avec le BAC *Dct-LacZ* (Murisier, *et al.*, 2008). Le profil d'expression du rapporteur *LacZ* chez les souris transgéniques BAC *Dct-LacZ* correspond à celui du gène *Dct* endogène. En effet, l'expression du rapporteur est observée dans le télencéphale, les mélanoblastes/mélanocytes et le RPE. Quelques colorations ectopiques sont également observées dans l'ensemble de l'embryon pendant le développement. Cependant, elles sont plus faibles que les expressions ectopiques rapportées chez les souris transgéniques *Dct-LacZ*, que ce soit dans les ganglions dorsaux, dans les nerfs rachidiens ou dans la tige optique (Figure 37) (Murisier, *et al.*, 2008). Le profil d'expression du rapporteur chez les souris BAC n'est finalement pas très différent du profil d'expression du gène *LacZ* chez des souris transgéniques *Dct-LacZ* avec le promoteur de 3,4 kb. L'utilisation du BAC permet une expression à priori plus spécifique du transgène, mais il faut tenir compte de la difficulté de construire un BAC et d'obtenir des souris transgéniques BAC (Murisier, *et al.*, 2008). A travers cette analyse, il a été montré que le gène *Dct* n'a pas d'éléments de régulation très en amont ou très en aval du premier et du dernier exons, contrairement aux gènes *Tyr* et *Tyrp1*.

### 3.4 Les souris transgéniques inductibles *iDct-LacZ* et *iDct-GFP*

Récemment, de nouvelles lignées transgéniques utilisant le promoteur *Dct* ont été décrites. Le promoteur de 3,4 kb du gène *Dct* utilisé est celui présent dans le transgène



**Figure 38 : Génération de la souris *iDct-GFP***  
(Zaidi, *et al.*, 2011b)

- A-** Génération des souris *iDct-GFP*. Les souris transgéniques *Dct-rtTA* ont été croisées avec les souris TRE-H2BGFP pour donner naissance aux souris bi-transgéniques *iDct-GFP*. Ces dernières expriment de la GFP de façon strictement régulée par la doxycycline.
- B-** Un embryon *iDct-GFP* au stade E11,5.
- C-** Un embryon *Dct-LacZ* au même stade.

*Dct-LacZ*. Ce promoteur contrôle l'expression du gène transactivateur réverse de la tétracycline ou *rtTA* (*reverse tetracycline transactivator*) (Woods and Bishop, 2011). En présence de tétracycline ou de doxycycline la protéine *rtTA* active un transgène sous le contrôle de TRE, un élément de réponse à la tétracycline (TRE, *Tet-response element*). Les souris *Dct-rtTA* ont été croisées avec des lignées de souris TRE-*LacZ* et TRE-H2B<sup>GFP</sup>. Les double-transgéniques *Dct-rtTA* TRE-*LacZ* et *Dct-rtTA* TRE-H2B<sup>GFP</sup> sont nommées *iDct-LacZ* et *iDct-GFP*, respectivement. Chez les souris *iDct-LacZ*, l'expression de *LacZ* est induite par la doxycycline et révélée par coloration X-gal chez l'adulte. On observe une coloration bleue dans les mélanocytes des follicules pileux et de la peau, du dos et de la queue (Woods and Bishop, 2011). Le gène rapporteur est également exprimé dans le cerveau et dans l'œil (Woods and Bishop, 2011). Ces résultats sont cohérents avec l'expression attendue du gène *Dct*. La souris transgénique *Dct-rtTA* est la première qui permette une expression inductible et réversible d'un transgène dans le lignage mélanocytaire (Woods and Bishop, 2011).

Une deuxième équipe a également généré une lignée *Dct-rtTA* et l'a croisée avec la lignée TRE-H2B<sup>GFP</sup> (Figure 38A) (Zaidi, *et al.*, 2011a). Les souris *Dct-rtTA* TRE-H2B<sup>GFP</sup>, nommées *iDct-GFP*, possèdent des cellules pigmentées fluorescentes, ce qui permet non seulement de les étudier au niveau cellulaire, mais aussi au niveau moléculaire après un tri par cytométrie de flux. Chez les souris *iDct-GFP* adultes, la protéine GFP est localisée dans le bulbe, le bulge et la gaine épithéliale externe du follicule pileux, où sont situés les mélanocytes et leurs précurseurs (Zaidi, *et al.*, 2011a). Par immunohistochimie avec un anticorps anti-DCT sur coupe de peau, il a été montré que les marquages GFP et DCT se superposent. La protéine GFP est donc bien exprimée exclusivement dans les cellules DCT+ (Zaidi, *et al.*, 2011a). L'administration de doxycycline à une femelle *Dct-rtTA* à partir du 10,5<sup>ème</sup> jour de gestation (E10,5) active l'expression de la GFP chez l'embryon. Une injection par voie intrapéritonéale de doxycycline à un nouveau-né ou à un adulte active l'expression de la GFP dans les 12 à 18 heures qui suivent. A E11,5, l'embryon *iDct-GFP* exprime la GFP dans le télencéphale, le RPE, les mélanoblastes en migration et les ganglions dorsaux. Le profil d'expression de la GFP est comparable à celui du gène *LacZ* chez des embryons *Dct-LacZ* au même stade (Figure 38B et C) (Zaidi, *et al.*, 2011b). Les souris *iDct-GFP* offrent un nouveau modèle pour suivre les cellules du lignage mélanocytaire chez l'embryon et chez l'adulte et pour isoler par tri en cytométrie de flux les mélanocytes et leurs précurseurs.



Dans ce chapitre, nous avons vu que les dérivés de la crête neurale qui donneront des mélanocytes ne migrent pas tous selon la voie dorso-latérale. Les précurseurs des cellules de Schwann, qui suivent la voie ventrale, sont une seconde source de mélanocytes. Ces cellules sont des précurseurs communs aux mélanocytes, aux cellules de Schwann et aux fibroblastes neuronaux. Il se trouve que de nombreuses pathologies de l'homme sont caractérisées par des troubles de la pigmentation associés à des anomalies du système nerveux périphérique. Ces pathologies sont classées parmi les maladies neurocutanées. L'existence d'un précurseur commun aux mélanocytes et aux cellules de Schwann fournit une nouvelle voie de recherche pour comprendre ces maladies (Adameyko and Lallemand, 2010). Nous avons vu également que le gène *Dct* est un bon marqueur du lignage mélanocytaire, bien qu'il soit également exprimé dans le télencéphale. La question de son expression dans les ganglions dorsaux reste à élucider. Enfin, nous avons passé en revue les différentes utilisations du gène *Dct* en transgénèse. *Dct* s'avère être un outil très intéressant d'une part pour étudier la biologie des mélanocytes et d'autre part pour exprimer des gènes d'intérêt dans les cellules productrices de pigments. Les mécanismes qui régissent la progression du mélanome malin cutané sont encore mal connus. Pourtant de nombreux gènes ont été identifiés sur la base de leur surexpression dans les mélanomes. L'importance fonctionnelle de leur surexpression est généralement inconnue. L'utilisation d'un système de transgénèse sous contrôle du promoteur *Dct* permettrait de tester les effets d'une expression forcée dans les cellules pigmentaires de tels gènes et ainsi de mieux comprendre leurs fonctions dans les mélanomes malins cutanés.



# OBJECTIFS DE LA THESE

## *Chapitre 1 : Elaboration d'un nouvel outil d'ingénierie génétique*

L'objectif de ma thèse était de fabriquer un outil original d'ingénierie génétique qui permette d'exprimer toute séquence d'intérêt dans les mélanocytes et leurs précurseurs chez l'embryon et chez la souris adulte. Pour cibler au mieux le lignage mélanocytaire, nous avons choisi de tirer parti du gène *Dct/Trp2* codant la dopachrome tautomérase. Le but était de faciliter l'étude de gènes impliqués dans la biologie des mélanocytes ou dans le développement et la progression des mélanomes malins.

Des gènes et des régions chromosomiques sont soupçonnés d'avoir une importance fonctionnelle lors de la progression du mélanome malin cutané. Ils ont été identifiés par des méthodes diverses, comme les puces d'expression (Hoek, *et al.*, 2004), les bibliothèques SAGE (Weeraratna, *et al.*, 2004), des études de transcriptome (Berger, *et al.*, 2010; Luo, *et al.*, 2011) ou de génomique intégrative (Kabbarah, *et al.*, 2010). Ces analyses reposent sur une comparaison des profils d'expression des gènes de mélanocytes normaux et de lignées de mélanomes, ou de mélanomes primaires et métastatiques. Les résultats permettent d'identifier de nombreux gènes candidats préférentiellement exprimés dans les mélanomes malins cutanés ou dans le mélanome en cours de développement. Pour mieux comprendre le rôle de ces gènes, il est nécessaire de développer des études fonctionnelles. L'utilisation de notre approche permettrait d'analyser les effets d'une expression forcée de ces gènes. Nous souhaitons participer à ce type d'étude en intégrant le gène *Rack1* au locus *Dct*, en collaboration avec l'UMR955 INRA de Génétique Fonctionnelle et Médicale de l'Ecole Nationale Vétérinaire d'Alfort. Le gène *RACK1* est surexprimé dans les métastases des mélanomes malins cutanés de l'homme (Egidy, *et al.*, 2008). L'utilisation itérative de notre système permettrait d'intégrer de nombreux gènes candidats au locus *Dct* en parallèle. Cela apporterait un gain de temps considérable dans l'étude de la progression du mélanome malin cutané.

Il était donc important de développer un système sans risque d'expression ectopique du transgène. Comme nous l'avons vu, les souris transgéniques portent, insérées au hasard, de multiples copies en tandem de la molécule d'ADN injectée. Il en résulte parfois une

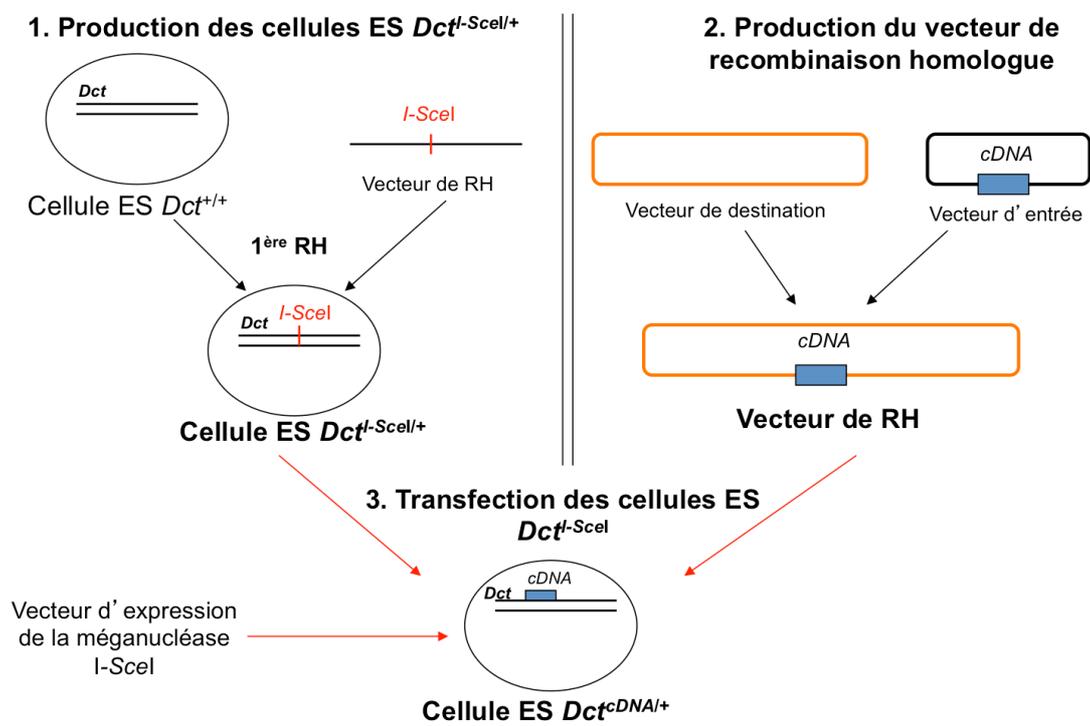


Figure 39 : Principe du projet d'élaboration d'un nouvel outil d'ingénierie génétique

expression ectopique du transgène qui rend nécessaire l'étude de plusieurs lignées indépendantes de souris transgéniques, afin de distinguer les effets spécifiques du transgène des effets de position. Au contraire, la technique du *knock in* élimine les problèmes d'expression ectopique du transgène. Dans ce cas, le facteur limitant est la fréquence faible de la recombinaison homologue (RH) qui dépend en outre du locus ciblé. Pour s'affranchir de ce problème, nous avons choisi de favoriser la RH en forçant la production d'une cassure double-brin au locus ciblé.

Le développement de l'outil s'est déroulé en trois étapes (Figure 39) :

1. *La production d'une lignée de cellules ES de génotype  $Dct^{I-SceI/+}$*

Ces cellules qui portent le site de reconnaissance de la méganucléase I-SceI inséré au locus *Dct* ont été produites en réalisant une expérience de RH conventionnelle.

2. *La fabrication d'un vecteur permettant une RH au locus *Dct**

Ce vecteur doit en outre pouvoir être utilisé de façon itérative pour permettre l'intégration au locus *Dct* de n'importe quelle séquence d'intérêt. Dans ce but, nous avons tiré profit de la technologie Gateway<sup>®</sup> (Invitrogen).

3. *Transfection des cellules ES  $Dct^{I-SceI/+}$  avec un vecteur de RH, contenant une séquence d'intérêt, et un vecteur d'expression de la méganucléase I-SceI*

Il est établi que l'induction d'un DSB dans le génome d'une cellule favorise le mécanisme de réparation par RH au niveau de la cassure (Choulika, *et al.*, 1995; Cohen-Tannoudji, *et al.*, 1998; Donoho, *et al.*, 1998). L'induction d'un DSB au locus *Villin* dans des cellules ES en culture a favorisé l'intégration d'une séquence étrangère par RH et l'efficacité de l'obtention de clones recombinants homologues a été améliorée d'un facteur 100 (Cohen-Tannoudji, *et al.*, 1998). Notre hypothèse était que l'induction d'un DSB au locus *Dct* par la méganucléase I-SceI favoriserait la RH dans des cellules ES en culture et permettrait d'obtenir avec une fréquence élevée des clones portant la séquence d'intérêt insérée au locus *Dct*. Nous avons testé notre hypothèse en intégrant le gène rapporteur *Lago1*, dérivé du gène *LacZ* d'*Escherichia coli*. A partir du gène bactérien *LacZ*, tous les îlots CpG ont été supprimés par des mutagenèses successives pour donner naissance au gène *LagoZ* (Chevalier-Mariette, *et al.*, 2003). Dans une seconde étape, les codons ont été « humanisés » et une séquence NLS a été ajoutée pour aboutir au gène *Lago1* (non publié). Le rapporteur *LacZ* est utilisé de longue date en transgénèse. Les modifications réalisées pour donner le gène *Lago1* devraient permettre une expression plus forte de la  $\beta$ -galactosidase. Au cours d'un deuxième

test, nous avons intégré un second gène rapporteur au locus *Dct*, le gène *H<sub>2</sub>BmCherry* qui exprime une protéine fluorescente rouge localisée au niveau du noyau.

Les résultats de cette étude sont présentés sous forme d'un article (en préparation) dans le **Chapitre 1 de la partie Résultats et discussions**.

## ***Chapitre 2 : Etude du patron d'expression des gènes rapporteurs *Lago1* et *H<sub>2</sub>BmCherry* intégrés au locus *Dct****

La souris transgénique *Dct-LacZ* exprime le transgène dans les mélanocytes et ses précurseurs. Chez l'embryon, le transgène est exprimé dans le télencéphale, les ganglions dorsaux et les mélanoblastes en migration (Mackenzie, *et al.*, 1997). L'expression du gène *LacZ* dans les ganglions dorsaux a longtemps été considérée comme un marquage ectopique. Des observations récentes ont montré l'existence d'un précurseur commun aux mélanocytes et aux cellules de Schwann, appelé précurseur de cellules de Schwann. Chez les souris *Dct-LacZ*, le marquage X-gal dans les ganglions dorsaux est peut-être dû à l'existence de ces précurseurs. Nous souhaitons donc documenter le profil d'expression du gène *Dct* avec un rapporteur inséré par RH à la place de l'exon 1 du gène *Dct*.

Comme nous l'avons vu dans le Chapitre 2, section 3.2, de la partie Introduction, une seule souris *knock in* utilisant le promoteur *Dct* a été publiée jusqu'à présent, la souris *Dct<sup>Cre/+</sup>* (Guyonneau, *et al.*, 2002). La souris *Dct<sup>Cre/+</sup>* a été croisée avec la souris R26R et le profil d'expression du rapporteur *LacZ* observé chez les souris hybrides n'est pas semblable à celui observé chez les souris transgéniques *Dct-LacZ*. La différence entre les profils d'expression n'est pas discutée dans l'article en question. Le gène inséré par *knock in* est le gène codant pour la Cre recombinase. Le profil d'expression de ce gène est étudié via la délétion d'une cassette floxée au locus *Rosa26* qui entraîne l'expression du rapporteur *LacZ*. Le profil d'expression de ce rapporteur chez les souris hybrides *Dct<sup>Cre/+</sup> × R26R* dépend de l'efficacité de recombinaison de la cassette floxée. L'emploi de cet outil de recombinaison peut aboutir à des efficacités variables selon le site ciblé et la disposition des séquences *loxP*. De plus, les diverses utilisations de la Cre recombinase ont montré que cette protéine pouvait avoir des effets toxiques sur le tissu ciblé (Higashi, *et al.*, 2009; Huh, *et al.*, 2010). Il convient donc d'étudier le profil d'expression du gène *Dct* via un autre rapporteur inséré à ce locus par recombinaison homologue.

Nous avons utilisé notre outil pour générer des cellules ES *knock in Dct<sup>Lago1-Néo/+</sup>*, puis *Dct<sup>H2BmCherry-Néo/+</sup>* (cf. article M. Fenina, en préparation). L'objectif était d'abord de vérifier que les gènes rapporteurs étaient exprimés dans des cellules pigmentées *in vitro*. Pour cela, nous avons utilisé une technique de différenciation des cellules ES en mélanocytes, puis nous avons observé l'expression des gènes rapporteurs dans les cellules en culture. Nous avons également généré des souris chimères (*Dct<sup>Lago1-Néo/+</sup> ↔ Dct<sup>+/+</sup>*), avec lesquelles nous avons produit une lignée de souris *Dct<sup>Lago1-Néo/+</sup>*. Nous avons croisé les souris hétérozygotes entre elles et obtenu des souris homozygotes *Dct<sup>Lago1-Néo/Lago1-Néo</sup>*. Nous avons observé le phénotype des souris homozygotes et comparé nos résultats à ceux de Guyonneau et collaborateurs sur les souris *Dct<sup>Cre/Cre</sup>* (Guyonneau, *et al.*, 2004). Enfin, des souris chimères mâles (*Dct<sup>Lago1-Néo/+</sup> ↔ Dct<sup>+/+</sup>*) ont été croisées avec des souris femelles *PGK-Cre*. Nous avons ainsi créé une lignée de souris *Dct<sup>Lago1/+</sup>*, pour lesquelles la cassette Néo est déléetée. Nous avons analysé le profil d'expression de *Lago1* chez les souris *Dct<sup>Lago1/+</sup>* chez l'embryon et chez l'adulte et l'avons comparé au profil du rapporteur *LacZ* des souris transgéniques *Dct-LacZ*. Les résultats de ces études sont présentés dans le **Chapitre 2 de la partie Résultats et discussions**.



# RESULTATS ET DISCUSSIONS

## *Chapitre 1 : Elaboration d'un nouvel outil d'ingénierie génétique*

### **1 Article : Une cassure double-brin induite par I-SceI ne favorise pas la recombinaison homologe au locus *Dct* dans les cellules embryonnaires souches de souris**

Cet article est en préparation et sera soumis à *Nucleic Acids Research*.

L'induction d'une cassure double-brin à un locus favorise le ciblage de gène à ce locus chez la levure, les plantes et les cellules de Mammifères. Chez la souris, le gène *Dct* est exprimé de façon assez spécifique dans les mélanocytes et leurs précurseurs. Nous souhaitons développer un outil de ciblage de gène au locus *Dct* dans les cellules ES de souris qui permette d'insérer n'importe quelle séquence nucléotidique d'intérêt à ce locus et d'étudier les conséquence de son expression forcée sur le devenir des cellules pigmentées chez l'embryon et chez l'adulte. Cet outil permettrait en particulier d'étudier l'effet d'une expression forcée de certains des gènes surexprimés pendant la progression du mélanome malin et dont la surexpression pourrait avoir une importance fonctionnelle sur la sévérité de la tumeur.

Nous avons d'abord inséré le site de reconnaissance d'I-SceI dans l'intron 1 du gène *Dct* par une recombinaison homologe conventionnelle. Nous avons obtenu un clone ES recombinant homologue de génotype  $Dct^{I-SceI-Neo/+}$ . La Cre recombinase a été exprimée transitoirement dans ces cellules pour éliminer la cassette de sélection à la Néomycine insérée au locus *Dct*. Deux clones ES MF1 et MF2  $Dct^{I-SceI/+}$  ainsi générés ont été validés pour le passage à la lignée germinale. Nous avons évalué l'efficacité de coupure de la méganucléase I-SceI au locus *Dct* *in vivo*. La technique dite de *ligation-mediated-PCR* (LM-PCR) permet de détecter d'éventuelles coupures dans une région chromosomique donnée. Nous avons testé par LM-PCR l'existence de coupures au site I-SceI dans des cellules ES  $Dct^{I-SceI/+}$ . Deux plasmides d'expression de I-SceI, pCMV-I-SceI et pCAG-I-SceI, ont été introduits dans des

cellules ES  $Dct^{I-SceI/+}$  et les coupures ont été recherchées. Les résultats ont montré que des cassures au locus  $Dct$  sont effectivement induites. Nous avons dès lors voulu tirer profit des cellules  $Dct^{I-SceI/+}$  pour introduire virtuellement n'importe quelle séquence d'intérêt à la place du gène  $Dct$ .

Un premier vecteur de réparation de la cassure appelé HR1 a été construit. Dans le vecteur HR1, la séquence à insérer au locus  $Dct$  est le gène  $LagoI$ , un gène dérivé du gène  $LacZ$ . Les cellules  $Dct^{I-SceI/+}$  ont été électroporées simultanément avec un plasmide d'expression de  $I-SceI$  et le vecteur de réparation HR1. Les expériences effectuées avec le plasmide d'expression pCMV- $I-SceI$  n'ont pas généré de clone recombinant homologue. Les expériences réalisées avec le plasmide pCAG- $I-SceI$  ont produit un clone recombinant homologue. La fréquence de recombinaison homologue en présence de  $I-SceI$ , exprimée par le plasmide pCAG- $I-SceI$ , est estimée à  $1,4 \times 10^{-7}$ . Cette efficacité est la même que celle d'une expérience de recombinaison homologue conventionnelle. En analysant ces résultats, nous avons identifié une particularité de la construction HR1 qui pouvait expliquer une faible fréquence de RH. En effet, en plus des régions d'homologie nécessaire à la RH, le plasmide HR1 contenait deux courtes régions, de 183 pb au total, qui sont homologues aux séquences situées autour du site  $I-SceI$  de l'allèle  $Dct^{I-SceI}$ . Or il a été montré que la RH favorisée par une cassure double-brin a lieu majoritairement dans les 200 pb encadrant le site de coupure. Notre hypothèse était que la RH s'est produite avec ces courtes régions d'homologie plutôt qu'avec les séquences homologues qui permettaient l'intégration du rapporteur  $LagoI$ . Nous avons donc construit un deuxième vecteur de réparation de la cassure, HR2. Le vecteur HR2 n'a plus le défaut du vecteur HR1. Le plasmide HR2 contient le gène rapporteur  $H_2B-mCherry$ . Nous avons introduit dans les expériences suivantes une recombinaison homologue conventionnelle qui implique de transfecter les cellules  $Dct^{I-SceI/+}$  avec le plasmide HR2 sous une forme linéaire. Nous avons en parallèle transfecté simultanément dans les cellules  $Dct^{I-SceI/+}$  un plasmide d'expression de  $I-SceI$  (pCMV- $I-SceI$  ou pCAG- $I-SceI$ ) et le vecteur HR2. L'expérience de recombinaison homologue conventionnelle a produit un clone recombinant homologue. L'expérience avec pCMV- $I-SceI$  (et HR2) n'a pas produit de clone recombinant homologue. L'expérience avec pCAG- $I-SceI$  (et HR2) a produit un clone recombinant homologue. Les fréquences de recombinaison homologue sont estimées à  $1,6 \times 10^{-7}$  lorsque l'enzyme  $I-SceI$  est exprimée (pCAG- $I-SceI$ ) et à  $2,7 \times 10^{-7}$  dans le cas d'une recombinaison homologue conventionnelle. La fréquence de recombinaison homologue

lorsque la méganucléase I-*SceI* est exprimée n'est donc pas supérieure à la fréquence de recombinaison homologue conventionnelle au locus *Dct*.

A travers cette étude, nous avons montré qu'un site de reconnaissance de I-*SceI* peut être coupé *in vivo* par la méganucléase lorsqu'il est présent au locus *Dct*. Nous avons également montré que cette coupure ne suffit pas à augmenter la fréquence de recombinaison homologue par rapport à une recombinaison homologue conventionnelle. La recombinaison homologue n'est donc pas systématiquement plus efficace lorsqu'une coupure double-brin est induite par I-*SceI* au locus cible que lorsque l'expérience est effectuée avec un vecteur de réparation linéaire sans cassure double-brin préalable. La structure de la chromatine au locus cible pourrait expliquer en partie cette différence.



I-*SceI*-mediated double-strand break does not increase the frequency of homologous recombination at the *Dct* locus in mouse embryonic stem cells

Myriam FENINA<sup>1, 2, 4</sup>, Dominique SIMON-CHAZOTTES<sup>1, 2</sup>, Sandrine VANDORMAEL-POURNIN<sup>1, 2</sup>, Francina LANGA-VIVES<sup>2, 3</sup>, Michel COHEN-TANNOUDJI<sup>1, 2</sup>, Bruno A. BERNARD<sup>4</sup>, and Jean-Jacques PANTHIER<sup>1, 2</sup>

1. Mouse functional Genetics, Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris cedex France.

2. CNRS URA 2578, Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris cedex France.

3. Mouse Genetics Engineering Center, Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris cedex France

4. Life Sciences Department, L'Oréal Recherche and Innovation, 90 rue du Général Roguet 92583 Clichy, France.

Short title: Homologous recombination at the *Dct* locus

Key words: Endonuclease cleavage; gene targeting; ES cells; DSB; DNA repair; knockin; gene correction

## **ABSTRACT**

Target-site double-strand breaks (DSBs) at natural endogenous loci were shown to increase the rate of gene replacement by homologous recombination in mouse embryonic stem cells. The effect of DSB at the dopachrome tautomerase (*Dct*) locus on the efficiency of gene targeting was studied using an embryonic stem cell line carrying the recognition site for the yeast *I-SceI* meganuclease embedded in the *Dct* gene. The targeted integration was measured after transfection of a repair vector with or without *I-SceI* expression constructs. The *I-SceI* meganuclease was indeed able to introduce a DSB at the *Dct* locus in live ES cells. However, the level of gene targeting was not improved by the DSB induction, indicating a limited capacity of *I-SceI* to increase the efficiency of homologous recombination at the *Dct* locus. These data suggest that targeted recombination by meganuclease-induced DSB may be locus dependent in mammalian cells.

## Introduction

The natural efficiency of the delivery of donor sequences at a specific locus of the mouse genome in embryonic stem (ES) cells by homologous recombination varies between  $1 \times 10^{-5}$  and  $1 \times 10^{-8}$ . Such a frequency is too low to consider the iterative introduction of a number of genes of interest at a given locus in standard practice. This problem can be overcome by enhancing recombination reactions at the target site through the induction of a double-strand break (DSB) (1). Such DSBs can be induced with the yeast mitochondrial I-SceI meganuclease which has an 18-bp recognition site, absent normally in the mammalian genome but that may be added to the genome of recipient cells. In previous studies performed with I-SceI, specific DSBs were shown to stimulate homologous recombination in mouse 3T3 fibroblasts, PCC7-S multipotent cells, and in several ES cell lines (2-6). Similar results were recently observed in Chinese hamster ovary cells after lentiviral delivery of the I-SceI protein (7). Expression of I-SceI was also shown to be relevant to improve the efficiency of gene targeting in other organisms, including plants (8). Altogether, it is generally admitted that a system based on the introduction of an I-SceI recognition site close to the targeted locus in the genome of recipient ES cells combined with transient expression of the I-SceI meganuclease to create a DSB, should enhance the precise delivery of donor sequences at this site.

The *Dct* gene encodes the dopachrome tautomerase, a melanogenic enzyme. In the embryo, *Dct* is expressed in pigment cell precursors, i.e. melanoblasts, derived from the neural crest, in cells of the retinal pigment epithelium (RPE) and in the developing forebrain (9,10). In the adult epidermis, *Dct* is expressed in pigment cells at all differentiation states: in stem cells that reside in the bulge region of the hair follicle, in progenitors of the outer root sheath and in melanocytes of the hair matrix (11). Studies performed in the mouse with a *LacZ* reporter gene expressed under the control of 3.4 kb of the *Dct* promoter depicted *LacZ* expression in melanoblasts and melanocytes, in the RPE, forebrain, dorsal root ganglia and caudal nerves (12). Hence, *Dct* promoter was shown to mimic largely the endogenous expression pattern of the gene although the expression of *LacZ* in the dorsal root ganglia is thought to be an ectopic staining (12). The *Dct* promoter has thus been used in transgenic mouse lines to drive the expression of genes in melanocytes and their precursors (13-16). However, a drawback of using combination of regulatory region from *Dct* and the coding regions of exogenous genes, either reporter genes or genes whose function is to be evaluated, is that several independent lines are required to distinguish the specific expression of the transgene from ectopic expression (17). Besides, *Dct* knockout mice are viable and fertile, and exhibit no defects,

with the exception of a diluted coat color (18), making *Dct* an interesting gene target for knockin experiments. We thus became interested by constructing a genetic tool that would allow to insert with a high efficiency virtually any gene of interest in place of the *Dct* gene. Such a tool would be useful, to study the effects of expression of genes that are significantly upregulated in cutaneous melanomas, as an example.

Our approach relied on the combination of three components: an ES cell line carrying a *Dct* allele with the *I-SceI* recognition site, an *I-SceI* expressing plasmid and a template for the DSB-repair process carrying sequence homologies to the target. The targeted integration at the *Dct* locus was tested after transfection of both the *I-SceI* expressing plasmid and the repair construct in the modified ES cell line. We report here that an *I-SceI* recognition site embedded within the *Dct* gene sequence can be cleaved by transiently expressed *I-SceI* meganuclease in ES cells. We further show, using two different repair vectors, that *I-SceI*-mediated DSB did not increase the frequency of homologous recombination at the *Dct* locus compared to conventional homologous recombination.

## **Materials and methods**

### **Plasmids and cells**

Plasmid pPBSK $\lambda$ B#4 was given by I.J. Jackson (MRC, Edinburgh, UK). Plasmids pL253, pL452, pSW23 were provided by N. Copeland and N. Jenkins (Frederick, MD, USA). CK35 ES cells (19) were grown on mitomycin C-treated Neo<sup>R</sup> primary fibroblasts in Dulbecco's modified Eagle's Medium + GlutaMAX (Invitrogen) supplemented with 15% fetal bovine serum, 0.1 mM  $\beta$ -mercaptoethanol (Sigma), 10<sup>3</sup> U/ml murine LIF (PAA) as previously described (20).

### **Design of repair vectors**

To insert an *I-SceI* recognition site at the *Dct* locus, a replacement vector was constructed. A 6.5 kb *SnaI-HincII* fragment from pPBSK $\lambda$ B#4 that contains 18 kb of *Dct* gene (21) was inserted into the *SpeI* site of pL253 to produce pL253-*Dct* plasmid. To introduce an *I-SceI* recognition site and a Neo<sup>R</sup> cassette into the 6.5 kb *Dct* fragment, a minitarget vector was constructed. A 5' *Dct* fragment containing a *NheI* site was amplified and flanked with *KpnI* and *EcoRI* sites using the following primers: 5' *KpnI* forward 5'-ATAGGTACCTCCCAATTAAGAAGGCATGG-3' and 5' *EcoRI* reverse 5'-GCGGAATTCCGCCTTTCTGAGTGAAGAG-3'. The amplicon was inserted into pCR2.1 plasmid (TOPO TA Cloning<sup>®</sup>, Invitrogen). The *I-SceI* recognition site was added at

the *NheI* site. Besides, a 3' *Dct* fragment was amplified and flanked with *BamHI* and *SacII* sites using the following primers: 3'*BamHI* forward 5'-GTAGGATCCACCTTTGGCTTGTTTGTGG-3' and 3'*SacII* reverse 5'-ATACCGCGGAGGACATGAGAACCCCAAGA-3'. The amplicon was inserted into a pCR2.1 plasmid. pSW23 plasmid was digested by *KpnI* and *SacII*, and filled in with the three fragments: the 5' *KpnI-EcoRI* fragment containing the *I-SceI* site, an *EcoRI-BamHI* Neo<sup>R</sup> cassette from pL452 plasmid and the 3' *BamHI-SacII* fragment. The replacement vector was produced by a recombineering reaction between the modified pSW23 and pL253-*Dct* plasmid. The replacement vector carries a herpes simplex virus-thymidine kinase (HSV-TK) negative selection cassette downstream of the 6.5 kb *Dct* fragment.

The construction of homologous recombination (HR) repair vectors HR1 and HR2 relied on the Gateway<sup>®</sup> technology (Invitrogen). Entry and destination vectors were produced. pENTR1A entry vector (Invitrogen) contains *ccdB* flanked with multi-cloning sites (MCS). The SV40 polyadenylation sequence (pA) was inserted at the *EcoRV* site to give pENTR1ApA. The *Lago1* gene with its start codon (provided by J.F. Nicolas, Institut Pasteur, Paris, France) was inserted in pENTR1ApA in place of *ccdB* (Fig. 1). The first destination vector (DV1) was constructed as follows. Starting from pL253-*Dct* plasmid, a *SexAI-AvrII* 400 bp fragment containing the ATG start codon was removed from *Dct* sequence. A linker made of the following primers was used to fill the gap in *Dct* sequences: *HpaPme* forward 5'-CTAGGTAAACGTTTAAA-3' and *HpaPme* reverse 5'-CCTGGTTTAAACGTTAAC-3'. The linker allowed the insertion of a unique *HpaI* recognition site into pL253-*Dct*, giving pL253-*Dct-HpaI* plasmid. To insert the Neo<sup>R</sup> cassette into pL253-*Dct-HpaI*, a recombineering reaction was performed and gave pL253-*Dct-HpaI-Neo<sup>R</sup>* plasmid. Finally, a reading frame cassette A (RfA) (Gateway<sup>®</sup> technology) that contains the Cm<sup>R</sup>-*ccdB* cassette flanked by *attR1* and *attR2* sites, was inserted at the *HpaI* site into pL253-*Dct-HpaI-Neo<sup>R</sup>* to obtain DV1. In DV1, the negative selection cassette HSV-TK from pL253 is downstream of the 3' sequence homology of *Dct* (Fig. 1). To produce the first repair vector (HR1), the Cm<sup>R</sup>-*ccdB* cassette in DV1 was replaced by *Lago1* gene using LR reaction (Gateway<sup>®</sup> technology) (Fig. 1).

To remove the short regions of homology between the HR1 repair vector and the first intron of *Dct* gene, a second destination vector (DV2) was produced. The H<sub>2</sub>B-*mCherry* reporter gene with its start codon (provided by J. Artus, Sloan-Kettering Institute, New-York, USA) was inserted in place of *ccdB* in pENTR1ApA entry vector (Fig. 2). To obtain DV2, a 2.8 kb

*AvrII-Bsu36I* fragment was synthesized. It contains a Neo<sup>R</sup> cassette framed with *FRT* sites and 1 kb of *Dct* genomic sequence. A 1 kb *AvrII-Bsu36I* fragment was removed from the pL253-*Dct-HpaI* plasmid and replaced by the 2.8 kb synthesized *AvrII-Bsu36I* fragment. Then RfA was ligated at the *HpaI* site. Figure 2 shows the map of DV2. To produce the second repair vector (HR2), the Cm<sup>R</sup>-*ccdB* cassette of DV2 was replaced by the H<sub>2</sub>B-*mCherry* sequences using LR reaction (Fig. 2). The integrity of the repair vectors was verified by sequence analysis.

### **Homologous recombination assay**

To insert a unique *I-SceI* site at the *Dct* locus, CK35 ES cells were electroporated with the *NotI*-linearized replacement vector. G418 (300 µg/mL) was added 48 h after plating for 12 days and gancyclovir (2 µM) was added 96 h after plating for 4 days. The Neo<sup>R</sup> cassette was removed using pIC-Cre plasmid (22) in which the transcription of Cre is driven by a synthetic HSV-tk promoter and enhancer. Fifteen micrograms of pIC-Cre plasmid were electroporated into approximately  $1 \times 10^7$  ES clone 4 cells and the cells were cultured without G418. For the gene targeting with HR1 repair vector, MF1 ES cells were electroporated with 20 µg of supercoiled HR1 alone or with 33 µg pCMV-*I-SceI* (23) or 37.5 µg pCAG-*I-SceI* (24). The ratio of expression plasmid to repair vector was 5 to 1. For the experiment using HR2, MF1 ES cells were electroporated with 30 µg of supercoiled HR2 plasmid alone or with either 11 µg pCMV-*I-SceI* or 13 µg pCAG-*I-SceI* expression plasmids. The ratio of expression plasmid to repair vector was 1 to 1. Approximately  $1.6 \times 10^7$  MF1 ES cells were independently electroporated with 30 µg of *NotI*-linearized HR2 plasmid, as a control.

### **DNA analysis in selected clones**

Genomic DNAs of ES clones obtained after selection with G418 and gancyclovir were digested with *BamHI*. Correct gene targeting was analyzed by Southern blot using a 1 kb 5' external probe produced by PCR amplification with the following primers: 5DCTF forward 5'-TTGGGGTCAGGGAGATACAG-3' and 5DCTR reverse 5'-TGAGCAGCAGTGAAGTTTGG-3'.

### **Ligation-mediated PCR (LM-PCR) analysis**

Approximately  $1.6 \times 10^7$  MF1 ES cells were electroporated with 50 µg pCMV-*I-SceI*, pCAG-*I-SceI* or mock plasmid. Four hours later, genomic DNA was extracted. Two micrograms of genomic DNA from MF1 cells transfected with the mock plasmid were digested with *PstI* or *I-SceI* and precipitated. LM-PCRs were performed with these *PstI*- and

I-SceI-digested DNAs, and with undigested DNA from MF1 cells transfected with mock plasmid, pCMV-I-SceI or pCAG-I-SceI. LM-PCRs were performed as follows. LM-C1 primer 5'-AATTCTTCAACCGGACAT-3' was used for the first extension. The extension product was ligated to an asymmetrical synthetic double-stranded linker constituted of two primers: a 25-mer linkerF primer forward 5'-GCGGTGACCCGGGAGATCTGAATTC-3' and an 11-mer linkerR primer reverse 5'-GAATTCAGATC-3'. A first PCR amplification was done on this ligation product using primer LM-C2: 5'-CGGACATGCAAATGCACAGGTGAGG-3' and the linkerF. The PCR-amplified product was then exponentially amplified by nested PCR using LM-C3 primer and linkerF (LM-C3 5'-CCCTTGGGCAGACCCAGATGTCACT-3'). The PCR products were separated on an agarose gel, then alkaline blotted to a nylon membrane, and hybridized to the following 36-mer radioactive probe LM 5'-CTTCTGAGGAGAGGGCGACACTGGTGACAAACTGTTA-3'.

### **Generation of *Dct*<sup>I-SceI/+</sup> mice**

Two independent *Dct*<sup>I-SceI/+</sup> 129/Sv ES clones, named MF1 and MF2, were introduced into C57BL/6J blastocysts, which were transferred to pseudo-pregnant C57BL/6J females. Male chimeras (129/Sv *Dct*<sup>I-SceI/+</sup> ↔ C57BL/6J *Dct*<sup>+/+</sup>) were mated with C57BL/6J females. The progeny was genotyped at the *Dct* locus to evaluate the percentage of *Dct*<sup>I-SceI/+</sup> mice.

### **Results**

Our experiments aimed at testing the efficiency of a ready-to-use tool to produce ES cells, and eventually mice, carrying any sequence of interest inserted in place of the *Dct* gene. Our strategy relied on the reported stimulation of gene targeting frequency at a natural locus associated with a DSB induced by the yeast meganuclease I-SceI in ES cells (5). We performed a two-step experiment. In a first step, the I-SceI restriction site was inserted to the *Dct* gene in ES cells using conventional gene targeting procedures. The *Dct* gene carrying a unique I-SceI restriction site was thus considered as a landing pad. In a second step, an I-SceI-expression plasmid was introduced together with a repair vector sharing a 5.9-kb of *Dct* isogenic DNA by electroporation in the engineered ES cells, and the efficiency of gene targeting at the *Dct* locus was assayed. Cabaniols and colleagues (2010) have coined the term cellular genome positioning system (cGPS) to designate such a gene targeting strategy.

### **Production of a landing pad at the *Dct* locus in ES cells**

To create a landing pad at the *Dct* locus, the I-*SceI* restriction site was inserted within the first intron of the *Dct* gene in ES cells. A replacement vector containing a unique I-*SceI* restriction site and a positive selection (Neo<sup>R</sup>) cassette, framed by *loxP* sites, all flanked with 1.9 and 4.5 kb genomic sequences from the *Dct* gene was constructed. A HSV-TK selection cassette was present outside the region of sequence homology between the vector and the *Dct* gene (Fig. 3A). The replacement vector was linearized and electroporated into approximately  $1.6 \times 10^7$  CK35 ES cells. The cells were cultured in the presence of G418 and gancyclovir. Out of 107 subcultured colonies, one clone (ES clone 4) was correctly targeted with the replacement vector as shown by PCR (data not shown) and Southern blot analysis (Fig. 3B). To test whether the meganuclease I-*SceI* is able to specifically cleave the new *Dct*<sup>I-*SceI*-Neo</sup> allele, genomic DNA of the ES clone 4 was treated with I-*SceI* and *Bam*HI restriction enzymes. Southern blot analysis using an external 5' probe revealed the 4.5 kb I-*SceI*-*Bam*HI distinctive fragment indicating that the I-*SceI* site inserted at the *Dct* locus was indeed cut *in vitro* by the meganuclease (Fig. 3C).

To delete the Neo<sup>R</sup> cassette, a Cre recombinase-expressing plasmid (pIC-Cre) was electroporated into *Dct*<sup>I-*SceI*-Neo/+</sup> ES (clone 4) cells. A preliminary experiment indicated that approximately 36% of the cells transfected with pIC-Cre plasmid died in presence of G418, presumably because they had lost the Neo<sup>R</sup> cassette. pIC-Cre plasmid was electroporated into approximately  $1 \times 10^7$  *Dct*<sup>I-*SceI*-Neo/+</sup> ES cells and the cells were cultured without G418. Twenty-four clones were subcultured and their sensitivity was assessed by adding G418 on a duplicate plate: 8 clones were Neo<sup>S</sup>. All eight clones had lost the Neo<sup>R</sup> cassette as shown by PCR analysis (data not shown). Two ES clones (MF1 and MF2) were further selected on morphological criteria. Southern blot analysis further confirmed the deletion of the Neo<sup>R</sup> cassette in MF1 and MF2 clones (Fig. 3D).

To test whether *Dct*<sup>I-*SceI*+/+</sup> MF1 and MF2 clones are able to colonize the germ line, we introduced MF1 and MF2 cells into C57BL/6J blastocysts, which were transferred to pseudo-pregnant females. Twelve and ten chimeras were produced from MF1 and MF2 cells, respectively. Altogether 18 chimeras were more than 95% chimeric, based on their coat color pattern. Several male chimeras were mated to C57BL/6J females. Half of their progeny was *Dct*<sup>I-*SceI*+/+</sup>, indicating that the genome of MF1 and MF2 ES cells was transmitted via the germ line.

#### **Assay of double-strand break at the *Dct* locus**

To test whether the meganuclease could trigger DSB *in vivo* at the *Dct* locus in *Dct*<sup>I-SceI/+</sup> ES cells, we electroporated I-SceI-expressing plasmids into MF1 ES cells and assayed the DNA lesion at the I-SceI site using a sensitive technique, known as ligation-mediated PCR (LM-PCR), that allows the specific detection of breaks in a defined region of genomic DNA (25). Two different I-SceI-expressing plasmids were tested: (i) pCMV-I-SceI in which I-SceI expression is driven by the cytomegalovirus promoter, and (ii) pCAG-I-SceI, where I-SceI is expressed under the control of the CAG composite promoter (5,23,24). Approximately 50 µg of pCMV-I-SceI, pCAG-I-SceI and a mock plasmid were independently electroporated into  $1.6 \times 10^7$  *Dct*<sup>I-SceI/+</sup> MF1 ES cells and four hours later the genomic DNAs were extracted.

To test the specificity and sensibility of the LM-PCR on transfected ES cells, we used two sites recognized by restriction endonucleases. The first is the I-SceI site, whose cleavage was under evaluation; the second is a *Pst*I site at position +52 relative to the I-SceI site. Approximately 2 µg of extracted genomic DNA from mock plasmid-transfected *Dct*<sup>I-SceI/+</sup> MF1 ES cells were digested with *Pst*I or I-SceI restriction endonuclease respectively. Then the digested DNA was denatured to allow annealing with a first *Dct* gene-specific primer (LM-C1) located at position -185 relative to the I-SceI site. This was followed by LM-C1 primer extension that terminated at the site of a break to produce a blunt-ended DNA, which was then ligated to an asymmetrical synthetic double-stranded linker. The newly synthesized DNA molecule was denatured to allow annealing with a second *Dct* gene-specific primer (LM-C2) located at position -174 relative to the I-SceI site and amplification in a PCR reaction with linker primer. The PCR-amplified products were exponentially amplified by nested PCR using a third *Dct*-gene-specific primer (LM-C3) located at position -148 relative to the I-SceI site and linker primer, as shown in Fig. 4A. Finally, the PCR products were separated on an agarose gel, alkaline blotted to a nylon membrane, and hybridized with a radioactive probe which does not overlap the primer sequences. PCR products of the predicted sizes, 148 bp for I-SceI digestion and 200 bp for *Pst*I digestion, were seen (Fig.4B). These data indicate that the LM-PCR technique allowed the specific detection of DNA cleavage in MF1 ES cells.

The ability of the meganuclease to trigger a DSB *in vivo* at the *Dct* locus was then evaluated. Approximately 2 µg of extracted genomic DNA from mock plasmid-transfected MF1 ES cells, pCMV-I-SceI- transfected MF1 ES cells and pCAG-I-SceI-transfected MF1 ES cells were directly analyzed by LM-PCR. Figure 4C shows that no DNA lesions occurred at significant level at the I-SceI site in the absence of I-SceI expression. When the genomic DNA

from pCMV-I-*SceI*-transfected MF1 ES cells was used as a template in the LM-PCR reaction, a 148 bp amplification product was detected, showing that expression of the meganuclease in MF1 ES cells triggered DSB at the I-*SceI* restriction site. LM-PCR in which the genomic DNA from pCAG-I-*SceI*-transfected cells was used as a template produced similar results (Fig. 4C), suggesting that both I-*SceI* expression vectors were equally efficient in triggering DSB at the target locus.

### **Landing of *Lago1* at the *Dct* locus**

We wished to develop a cellular genome positioning system (cGPS) to repeatedly target a single locus, *Dct*, in the mouse. To test the cGPS, we first used the *Lago1* gene. *Lago1* was derived from the *LacZ* gene by two successive steps. First, the 291 CpG sequences of *LacZ*, with the exception of two, were mutated to reduce the CpG content of the bacterial gene, while conserving the amino acid sequence (26). Second, codon usage was humanized to promote translation in mammals and a nuclear location signal (*nls*) sequence was added (J.F. Nicolas, unpublished data). The construction of the HR1 repair vector relied on the Gateway<sup>®</sup> technology (Fig. 1). The HR1 repair vector contained *Lago1*, a Neo<sup>R</sup> cassette framed with *loxP* sites, two regions of homology with the *Dct*<sup>I-*SceI*</sup> allele, 1.4 and 4.5 kb in length, and a HSV-TK negative selection cassette (Fig. 5A).

We assessed the rate of insertion of *Lago1* gene at the *Dct* locus following DSB-induced homologous recombination. Approximately  $1.3 \times 10^7$  MF1 ES cells were electroporated with 20  $\mu$ g of supercoiled HR1 either with or without an I-*SceI* expressing plasmid. Both pCMV-I-*SceI* and pCAG-I-*SceI* were tested. The cells were exposed to G418 and gancyclovir. A total of 215, 235 and 252 colonies resistant to both antibiotics were obtained when MF1 ES cells were transfected with HR1 alone, and in combination with pCMV-I-*SceI* or pCAG-I-*SceI* respectively. For each experiment, 136 colonies were individually picked up and PCR tested. Transfection with either HR1 alone or with HR1 and pCMV-I-*SceI* gave no targeted colonies. Transfection with HR1 and pCAG-I-*SceI* gave a positive PCR signal (data not shown), which was confirmed by Southern blot analysis (Fig. 5B and data not shown). Thus gene targeting using the HR1 repair vector and pCAG-I-*SceI* led to a frequency of homologous recombination that could be estimated at  $1.4 \times 10^{-7}$ . This frequency is not higher than that obtained with the conventional gene targeting procedures (generally range between  $10^{-5}$  and  $10^{-8}$ ).

Elliott and colleagues (1998) reported previously that, in I-*SceI*-induced gene targeting with a transfected circular plasmid, the majority of recombination events occurred within 100 bp

from the cleavage site. Actually, the HR1 repair vector contains two short regions of homology with the targeted *Dct*<sup>I-SceI</sup> allele next to the I-SceI site. These regions are shown in Figure 5C. In the *Dct*<sup>I-SceI</sup> allele, the first short region of homology is located between the exon 1 of *Dct* gene and the I-SceI site. It encompasses 58 bp of *Dct* intron 1 sequence. Still in the *Dct*<sup>I-SceI</sup> allele, a second region of homology is located between I-SceI site and the end of *loxP* site. It encompasses 125 bp of *Dct* intron 1 sequence and *loxP* sequence. In HR1 repair vector, both 58 bp and 125 bp regions are located between the *attB2* site and the Neo<sup>R</sup> cassette. We hypothesized that these homology regions, 183 bp in length, could be used as an efficient repair template and would produce by homologous recombination a recombinant allele (*Dct*<sup>loxP</sup>) harboring neither a *Lago1* gene nor a Neo<sup>R</sup> cassette. Hence, homologous recombinant clones carrying the *Dct*<sup>loxP</sup> allele would die in the presence of G418.

### **Landing of H<sub>2</sub>B-*mCherry* at the *Dct* locus**

We thus decided to remove both short regions of homology (including the *loxP* site) and to construct a novel repair vector. Therefore, a second destination vector (DV2) was produced (see Materials and Methods). To test anew the cGPS, we used the H<sub>2</sub>B-*mCherry* reporter gene. The HR2 repair vector was constructed (Fig. 2). It carries H<sub>2</sub>B-*mCherry*, a Neo<sup>R</sup> cassette framed with *FRT* sites, two regions of homology with the *Dct*<sup>I-SceI</sup> allele, 1.4 and 4.5 kb in length, and a HSV-TK negative selection cassette (Fig. 6A). By contrast with HR1, HR2 displays neither a short region of homology with the *Dct* gene next to the I-SceI site nor a *loxP* site.

We assessed the rate of targeted insertion of H<sub>2</sub>B-*mCherry* at the *Dct* locus by I-SceI-induced homologous recombination. Approximately  $1.6 \times 10^7$  MF1 ES cells were electroporated with 30 µg of supercoiled HR2 plasmid alone or with either pCMV-I-SceI or pCAG-I-SceI expression plasmids (Fig. 6A). As an additional control, approximately  $1.6 \times 10^7$  MF1 ES cells were independently electroporated with 30 µg of linearized HR2 plasmid. The cells were cultured in the presence of G418 and gancyclovir. Colony counting revealed 397 and 370 colonies in presence of pCMV-I-SceI and pCAG-I-SceI plasmids respectively, and 653 colonies in the absence of the meganuclease. Electroporation with linear HR2 plasmid, representative of a conventional gene targeting experiment, revealed 635 resistant colonies to both antibiotics. For each experiment, 144 colonies were individually picked up, amplified and PCR tested. Electroporation with linear HR2 plasmid resulted in one targeted clone (clone 7E). No targeted clones were seen in the supercoiled HR2-electroporated MF1 ES cells. The same results were obtained in the MF1 ES cells electroporated with both supercoiled HR2 and

pCMV-I-*SceI* plasmids. Electroporation with HR2 and pCAG-I-*SceI* led to one targeted clone (clone 11E). Both conventional (linear HR2) and I-*SceI*-mediated (supercoiled HR2 and pCAG-I-*SceI*) *Dct* locus targetings were confirmed by Southern blot analysis (Fig. 6B). Thus gene targeting using the HR2 repair vector and pCAG-I-*SceI* led to a frequency of homologous recombination that could be estimated at  $1.6 \times 10^{-7}$  whereas conventional gene targeting led to a frequency of  $2.7 \times 10^{-7}$ . Therefore, I-*SceI*-induced DSB at the *Dct* locus did not provide a higher frequency of homologous recombination compared to conventional gene targeting in ES cells.

## Discussion

In this report, we have presented evidence that I-*SceI*-induced DSB in ES cells does not improve the efficacy of the gene targeting methodology at the *Dct* locus compared to the conventional approach. Electroporation was used to introduce the I-*SceI* expressing plasmids into ES cells and a low efficiency of transfection could explain these results. However, we were able to detect GFP expression in at least 65% of a population of CK35 ES cells electroporated with a plasmid containing the CMV promoter driving the expression of the GFP, indicating efficient electroporation (data not shown). Since the first repair vector (HR1) contained short regions of homology in the vicinity of the I-*SceI* site, we first hypothesized that these regions were preferentially used to repair the DSB, thus generating homologous recombinant clones that did not integrate the Neo<sup>R</sup> cassette and died eventually in the presence of G418. Therefore, a second repair vector (HR2) with no homology to the sequence surrounding the I-*SceI* site was generated, but we still failed to demonstrate improvement of the gene targeting. Actually it has been previously reported in mouse Ltk<sup>-</sup> fibroblasts that chromosomal DSBs introduced by I-*SceI* at a specific site was also unable to stimulate gene targeting (27).

Distinct differences in frequencies of targeted integration driven by a DSB among cell types have been reported. Homologous recombination after cleavage by a zinc-finger nuclease (ZFN) at the *CCR5* locus in presence of cognate donor linear and circular episomes was more efficient in a panel of immortalized cell lines from human leukemia than in human stem cells, such as cord blood CD34<sup>+</sup> hematopoietic cells and human ES cells (28). It has also been reported that the rate of ZFN-mediated gene targeting at the *Rosa26* locus was higher in primary fibroblasts from adult mice than in murine ES cells (29). These results suggest that DSB-induced gene targeting may be less efficient in ES cells than in somatic cells. This contention seems inconsistent with reports showing that homologous recombination is the

predominant pathway to repair DSBs, whereas somatic cells utilize non-homologous end joining (NHEJ) (30,31). It has also been reported that ES cells that had been allowed to differentiate preferred the error-prone NHEJ pathway to the high-fidelity homologous recombination to repair DNA DSBs (30). Because ES cells and somatic cells are intrinsically different in the extent to which they retain genomic integrity (32), it was important to assess that our experiments were made with genuine ES culture rather than differentiated ES culture. We confirmed that the *Dct*<sup>I-SceI/+</sup> CK35 cells are truly pluripotent ES cells, able to colonize the germ line. Furthermore, CK35 ES cells have been previously used to demonstrate highly efficient gene targeting after DSB (5).

We observed that the repair vector was inserted at random chromosomal sites rather than at the I-SceI site in the majority of neomycin- and gancyclovir-resistant clones. One possible explanation for the random integration was that I-SceI did not cleave its recognition sequence in *Dct*<sup>I-SceI/+</sup> ES cells. However, we clearly showed the existence of I-SceI-induced DSBs at the *Dct* locus. The DSB induced by I-SceI at the *Dct* locus is not expected to be unique in a given *Dct*<sup>I-SceI/+</sup> ES cell for two reasons. First, there is a high basal level of DSBs in exponentially growing murine ES cells. Approximately 20 phosphorylated histone 2AX ( $\gamma$ H2AX) foci, that allow the visualization of DSBs (33), may be observed in ES cells (31). Second, I-SceI was recently shown to cleave *in vivo* I-SceI-like sites where the nucleotide sequence was identical to the consensus I-SceI cleavage sequence at 14 to 15 out of 18 base pairs (bp), even when the wild-type cleavage site was present in the same cell (34). Such random integrations could not be detected in previous studies performed with I-SceI in mouse ES cells where selection strategies were specifically designed to single out the gene targeting events and eliminate the nonrecombinant recombination events (4-6). Seminal experiments of gene targeting in ES cells based on an I-SceI-induced gene replacement system were first performed with mutated resistance genes integrated in chromosomal sequences (4), and later extended to natural endogenous genes, *Hprt*, *Villin* and *Dbx1* genes (5,6,35). However, altogether, the number of actual genes that have been efficiently targeted following I-SceI-mediated homologous recombination in ES cells is limited. It is widely accepted that the efficiency of conventional homologous recombination depends on the target locus. Our results suggest that, similarly, gene correction by DSB-induced homologous recombination may be poorly efficient depending on the targeted locus, although such negative results are generally not reported in the literature. Consequently we anticipate that deeper analysis of the meganuclease, repair vector, target locus and cells that fail homologous recombination by

DSB may also shed light on the nature of the factors that contribute to gene targeting in mammalian cells.

### **Acknowledgements**

We thank Jérôme Artus, Neal Copeland, Ian Jackson, Nancy Jenkins, and Jean-François Nicolas for the gift of plasmids. We are grateful to Stéphane Commo, Olivier Danos, and Jean-Stéphane Joly for helpful discussions. This work was supported by grants from the *Association Nationale de la Recherche et de la Technologie* (ANRT), research contract C080052, and *L'Oréal Recherche*. MF was awarded fellowships from the *ANRT*, *L'Oréal Recherche* and *Fondation pour la Recherche Médicale*.

## References

1. Orr-Weaver, T.L., Szostak, J.W. and Rothstein, R.J. (1981) Yeast transformation: a model system for the study of recombination. *Proc Natl Acad Sci U S A*, **78**, 6354-6358.
2. Rouet, P., Smih, F. and Jasin, M. (1994) Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol*, **14**, 8096-8106.
3. Choulika, A., Perrin, A., Dujon, B. and Nicolas, J.F. (1995) Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccharomyces cerevisiae*. *Mol Cell Biol*, **15**, 1968-1973.
4. Smih, F., Rouet, P., Romanienko, P.J. and Jasin, M. (1995) Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. *Nucleic Acids Res*, **23**, 5012-5019.
5. Cohen-Tannoudji, M., Robine, S., Choulika, A., Pinto, D., El Marjou, F., Babinet, C., Louvard, D. and Jaissier, F. (1998) I-SceI-induced gene replacement at a natural locus in embryonic stem cells. *Mol Cell Biol*, **18**, 1444-1448.
6. Donoho, G., Jasin, M. and Berg, P. (1998) Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. *Mol Cell Biol*, **18**, 4070-4078.
7. Izmiryan, A., Basmaciogullari, S., Henry, A., Paques, F. and Danos, O. (2011) Efficient gene targeting mediated by a lentiviral vector-associated meganuclease. *Nucleic Acids Res*, **39**, 7610-7619.
8. Puchta, H., Dujon, B. and Hohn, B. (1996) Two different but related mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. *Proc Natl Acad Sci U S A*, **93**, 5055-5060.
9. Steel, K.P., Davidson, D.R. and Jackson, I.J. (1992) TRP-2/DT, a new early melanoblast marker, shows that steel growth factor (c-kit ligand) is a survival factor. *Development*, **115**, 1111-1119.
10. Pavan, W.J. and Tilghman, S.M. (1994) Piebald lethal (sl) acts early to disrupt the development of neural crest-derived melanocytes. *Proc Natl Acad Sci U S A*, **91**, 7159-7163.
11. Nishimura, E.K., Jordan, S.A., Oshima, H., Yoshida, H., Osawa, M., Moriyama, M., Jackson, I.J., Barrandon, Y., Miyachi, Y. and Nishikawa, S. (2002) Dominant role of the niche in melanocyte stem-cell fate determination. *Nature*, **416**, 854-860.
12. Mackenzie, M.A., Jordan, S.A., Budd, P.S. and Jackson, I.J. (1997) Activation of the receptor tyrosine kinase Kit is required for the proliferation of melanoblasts in the mouse embryo. *Dev Biol*, **192**, 99-107.
13. Guyonneau, L., Rossier, A., Richard, C., Hummler, E. and Beermann, F. (2002) Expression of Cre recombinase in pigment cells. *Pigment Cell Res*, **15**, 305-309.
14. Dunn, K.J., Brady, M., Ochsenbauer-Jambor, C., Snyder, S., Incao, A. and Pavan, W.J. (2005) WNT1 and WNT3a promote expansion of melanocytes through distinct modes of action. *Pigment Cell Res*, **18**, 167-180.
15. Lanning, J.L., Wallace, J.S., Zhang, D., Diwakar, G., Jiao, Z. and Hornyak, T.J. (2005) Altered melanocyte differentiation and retinal pigmented epithelium transdifferentiation induced by Mash1 expression in pigment cell precursors. *J Invest Dermatol*, **125**, 805-817.
16. Woods, S.L. and Bishop, J.M. (2011) A new transgenic mouse line for tetracycline inducible transgene expression in mature melanocytes and the melanocyte stem cells using the Dopachrome tautomerase promoter. *Transgenic Res*, **20**, 421-428.

17. De Sepulveda, P., Salaun, P., Maas, N., Andre, C. and Panthier, J.J. (1995) SARs do not impair position-dependent expression of a kit/lacZ transgene. *Biochem Biophys Res Commun*, **211**, 735-741.
18. Guyonneau, L., Murisier, F., Rossier, A., Moulin, A. and Beermann, F. (2004) Melanocytes and pigmentation are affected in dopachrome tautomerase knockout mice. *Mol Cell Biol*, **24**, 3396-3403.
19. Kress, C., Vandormael-Pournin, S., Baldacci, P., Cohen-Tannoudji, M. and Babinet, C. (1998) Nonpermissiveness for mouse embryonic stem (ES) cell derivation circumvented by a single backcross to 129/Sv strain: establishment of ES cell lines bearing the Omd conditional lethal mutation. *Mamm Genome*, **9**, 998-1001.
20. Robertson, E., Bradley, A., Kuehn, M. and Evans, M. (1986) Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. *Nature*, **323**, 445-448.
21. Budd, P.S. and Jackson, I.J. (1995) Structure of the mouse tyrosinase-related protein-2/dopachrome tautomerase (Tyrp2/Dct) gene and sequence of two novel slaty alleles. *Genomics*, **29**, 35-43.
22. Gu, H., Zou, Y.R. and Rajewsky, K. (1993) Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. *Cell*, **73**, 1155-1164.
23. Choulika, A., Perrin, A., Dujon, B. and Nicolas, J.F. (1994) The yeast I-Sce I meganuclease induces site-directed chromosomal recombination in mammalian cells. *C R Acad Sci III*, **317**, 1013-1019.
24. Okabe, M., Ikawa, M., Kominami, K., Nakanishi, T. and Nishimune, Y. (1997) 'Green mice' as a source of ubiquitous green cells. *FEBS Lett*, **407**, 313-319.
25. Besaratinia, A. and Pfeifer, G.P. (2009) DNA-lesion mapping in mammalian cells. *Methods*, **48**, 35-39.
26. Chevalier-Mariette, C., Henry, I., Montfort, L., Capgras, S., Forlani, S., Muschler, J. and Nicolas, J.F. (2003) CpG content affects gene silencing in mice: evidence from novel transgenes. *Genome Biol*, **4**, R53.
27. Lukacsovich, T., Yang, D. and Waldman, A.S. (1994) Repair of a specific double-strand break generated within a mammalian chromosome by yeast endonuclease I-SceI. *Nucleic Acids Res*, **22**, 5649-5657.
28. Lombardo, A., Genovese, P., Beausejour, C.M., Colleoni, S., Lee, Y.L., Kim, K.A., Ando, D., Urnov, F.D., Galli, C., Gregory, P.D. *et al.* (2007) Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat Biotechnol*, **25**, 1298-1306.
29. Connelly, J.P., Barker, J.C., Pruett-Miller, S. and Porteus, M.H. (2010) Gene correction by homologous recombination with zinc finger nucleases in primary cells from a mouse model of a generic recessive genetic disease. *Mol Ther*, **18**, 1103-1110.
30. Tichy, E.D., Pillai, R., Deng, L., Liang, L., Tischfield, J., Schwemberger, S.J., Babcock, G.F. and Stambrook, P.J. (2010) Mouse embryonic stem cells, but not somatic cells, predominantly use homologous recombination to repair double-strand DNA breaks. *Stem Cells Dev*, **19**, 1699-1711.
31. Serrano, L., Liang, L., Chang, Y., Deng, L., Maulion, C., Nguyen, S. and Tischfield, J.A. (2011) Homologous recombination conserves DNA sequence integrity throughout the cell cycle in embryonic stem cells. *Stem Cells Dev*, **20**, 363-374.
32. Stambrook, P.J. and Tichy, E.D. (2010) Preservation of genomic integrity in mouse embryonic stem cells. *Adv Exp Med Biol*, **695**, 59-75.
33. Bonner, W.M., Redon, C.E., Dickey, J.S., Nakamura, A.J., Sedelnikova, O.A., Solier, S. and Pommier, Y. (2008) GammaH2AX and cancer. *Nat Rev Cancer*, **8**, 957-967.

34. Petek, L.M., Russell, D.W. and Miller, D.G. (2010) Frequent endonuclease cleavage at off-target locations in vivo. *Mol Ther*, **18**, 983-986.
35. Bielle, F., Griveau, A., Narboux-Neme, N., Vigneau, S., Sigrist, M., Arber, S., Wassef, M. and Pierani, A. (2005) Multiple origins of Cajal-Retzius cells at the borders of the developing pallium. *Nat Neurosci*, **8**, 1002-1012.

## Figure legends

**FIG. 1. Construction of the HR1 repair vector.** The pENTR1ApA entry vector is represented in the upper left. Black circles flanking *ccdB* represent multi-cloning sites (MCSs). The SV40 polyadenylation sequence (pA) was introduced at the 3' end of the *ccdB* gene, before *attL2* sequence. The entry vector contains a Kan<sup>R</sup> cassette for selection in bacterial culture. The entry vector carrying *Lagol* gene is represented in the upper right. The destination vector (DV1) contains 1.4 and 4.5 kb of *Dct* homologous arms depicted as grey rectangles. DV1 also contains the Neo<sup>R</sup> and HSV-TK cassettes used in cell culture, and an Amp<sup>R</sup> cassette for selection in bacterial culture. The repair vector (HR1) is produced by LR reaction, allowing the replacement of Cm<sup>R</sup>-*ccdB* cassette by the *Lagol* gene.

**FIG. 2. Construction of the HR2 repair vector.** The pENTR1ApA entry vector is represented in the upper left. The entry vector containing the H<sub>2</sub>B-*mCherry* gene is represented in the upper right. The DV2 destination vector contains the *Dct* homologous arms, 1.4 and 4.5 kb in length, depicted as grey rectangles. The black circle denotes 109 bp of *Dct* intron absent in DV1 destination vector that were inserted in DV2 destination vector. DV2 also contains a Neo<sup>R</sup> cassette flanked with *FRT* sites depicted as white diamond symbols. The repair vector (HR2) is produced by LR reaction, allowing the replacement of Cm<sup>R</sup>-*ccdB* cassette by the H<sub>2</sub>B-*mCherry* gene.

**FIG. 3. Production of a landing pad at the *Dct* locus.** (A) Introduction of an I-*SceI* site at the *Dct* locus. From top to bottom are represented the *Dct* wild-type allele (*Dct*<sup>+</sup>), the repair vector, the *Dct*<sup>I-*SceI*-Neo</sup> targeted allele, and the *Dct*<sup>I-*SceI*</sup> allele produced after the deletion of the Neo<sup>R</sup> cassette. The grey boxes represent the exons 1 and 2 of *Dct* gene. The black circle represents 109 bp of *Dct* intron sequence that are lost during an homologous recombination event. The horizontal dark grey bar represents the external 5' probe used for the Southern blots. The Neo<sup>R</sup> and HSV-TK cassettes are depicted as white rectangles. *loxP* sites are represented by white triangles. The *Dct* homologous arms, 1.9 and 4.5 kb in length, are denoted as grey rectangles. I-*SceI* and *Bam*HI restriction sites are indicated. (B) Southern blot analysis of *Dct*<sup>+/+</sup> ES cells and targeted ES cells (clone 4). Genomic DNAs of ES cells were digested with *Bam*HI. The 11.7 and 6.4 kb fragments are distinctive of the *Dct*<sup>+</sup> and *Dct*<sup>I-*SceI*-Neo</sup> alleles, respectively. (C) Test of the ability of I-*SceI* meganuclease to specifically cleave *Dct*<sup>I-*SceI*-Neo/+</sup> ES cells. Southern blot analysis of *Dct*<sup>+/+</sup> ES cells and clone 4. Genomic

DNAs were digested with *I-SceI* and *BamHI*. The 4.5 kb fragment is distinctive of the *Dct<sup>I-SceI-Neo</sup>* allele. (D) Southern blot analysis of *Dct<sup>+/+</sup>* ES cells, clone 4, MF1 and MF2 clones. Genomic DNAs were digested with *BamHI*. The 4.5 kb fragment is distinctive of the *Dct<sup>I-SceI</sup>* allele.

**FIG.4. Assay of double-strand break induced by *I-SceI* at the *Dct* locus.** (A) Diagram of the ligation-mediated PCR (LM-PCR) technique to analyze a lesion at the *I-SceI* site. The successive steps of LM-PCR are represented from top to bottom. Arrows indicate the *Dct* gene-specific primers LM-C1, LM-C2 and LM-C3. Two black bars depict the asymmetrical synthetic double-stranded linker constituted of linkerF and linkerR primers. After cleavage and denaturation of genomic DNA, LM-C1 primer was annealed and extended. Then, the double-stranded linker was ligated to the blunt-ended fragment. Fragments were PCR amplified using LM-C2 and linkerF primer. A second nested PCR amplification was performed using LM-C3 and linkerF primer. A lesion at the *I-SceI* site would lead to a 148 bp LM-PCR product. (B) LM-PCR assay revealed with a radioactive probe. MF1 ES cells were transfected with mock plasmid expression plasmid. Genomic DNA from mock plasmid-transfected ES cells was digested with *PstI* or *I-SceI*. The 200 bp and 148 bp fragments are distinctive of the LM-PCR product after digestion with *PstI* and *I-SceI*, respectively. (C) LM-PCR results revealed using a radioactive probe. Undigested genomic DNAs from mock plasmid-, pCMV-*I-SceI*- and pCAG-*I-SceI*-transfected MF1 ES cells were analyzed by LM-PCR. No LM-PCR products were observed when undigested DNA from mock plasmid MF1 ES cells was used as a template. A distinctive 148 bp fragment was obtained after LM-PCR-amplification of undigested DNA from MF1 ES cells transfected with pCMV-*I-SceI* and pCAG-*I-SceI*.

**FIG.5. Landing of *Lago1* gene at the *Dct* locus.** (A) Insertion of *Lago1* gene at the *Dct* locus. The *Dct<sup>I-SceI</sup>* allele, the HR1 repair vector and the *Dct<sup>Lago1-Neo</sup>* targeted allele are represented from top to bottom. A lightning denotes *I-SceI* expression from pCMV-*I-SceI* or pCAG-*I-SceI* plasmid. The 1.4 and 4.5 kb of *Dct* isogenic DNA are depicted by grey rectangles. (B) Southern blot analysis of *Dct<sup>+/+</sup>*, *Dct<sup>I-SceI/+</sup>* and *Dct<sup>Lago1-Neo/+</sup>* ES cells. Genomic DNAs were digested with *BamHI*. The probe used for the hybridization is the external 5' probe depicted by a dark grey bar. The 11.7, 4.5, and 10.1 kb fragments are distinctive of the *Dct<sup>+</sup>*, *Dct<sup>I-SceI</sup>*, and *Dct<sup>Lago1-Neo</sup>* alleles, respectively. (C) Diagram of DSB-induced homologous recombination with no insertion of the *Lago1* gene. The *Dct<sup>I-SceI</sup>* allele, the HR1 plasmid and

the  $Dct^{loxP}$  recombinant allele are represented from top to bottom. Short regions of homology, 58 bp and 125 bp in length, between the HR1 repair vector and the genomic DNA at the  $Dct$  locus in  $Dct^{I-SceI/+}$  cells, are depicted by dark grey rectangles.

**FIG.6. Landing of H<sub>2</sub>B-*mCherry* gene at the *Dct* locus.** (A) Insertion of H<sub>2</sub>B-*mCherry* gene at the  $Dct$  locus. The  $Dct^{I-SceI}$  allele, the HR2 repair vector and the  $Dct^{H2B-mCherry-Neo}$  allele are represented from top to bottom. *AvrII* sites are indicated. There are no homologous sequences between  $Dct^{I-SceI}$  and HR2 close to the *I-SceI* site. A lightning denotes *I-SceI* expression from pCMV-*I-SceI* or pCAG-*I-SceI* plasmid. The 1.4 and 4.5 kb of  $Dct$  isogenic DNA are depicted by grey rectangles. (B) Southern blot analysis of  $Dct^{+/+}$ ,  $Dct^{I-SceI/+}$  ES cells, and  $Dct^{H2B-mCherry-Neo/+}$  ES clones 7E and 11E. ES clones 7E and 11E were obtained after transfection with the linear HR2 repair vector, and with both the supercoiled HR2 repair vector and pCAG-*I-SceI* expression plasmid, respectively. Genomic DNAs were digested with *AvrII*. The probe used for the hybridization is the external 5' probe depicted by a dark grey bar. The 3.9 kb fragment is carried by the  $Dct^{+}$  and  $Dct^{I-SceI}$  alleles, and the 6.5 kb fragment is distinctive of the  $Dct^{H2B-mCherry-Neo}$  targeted allele.

Figures

FIG.1.

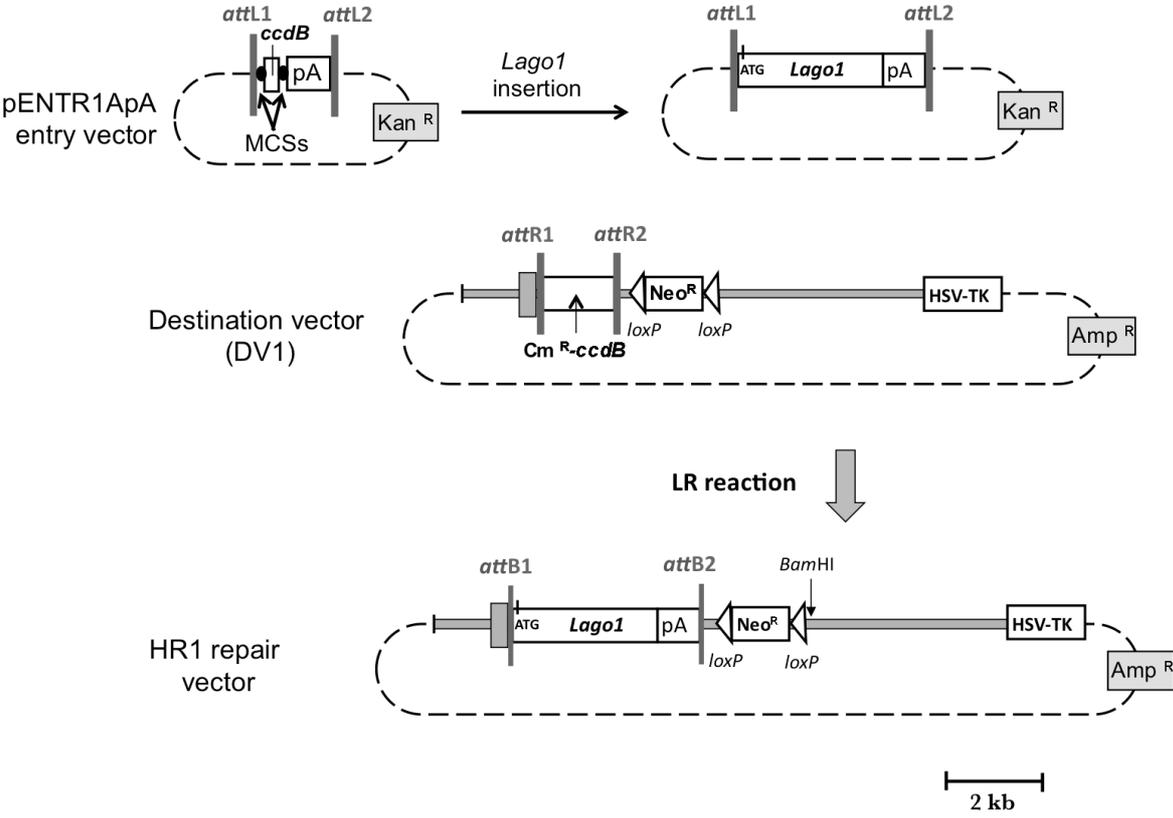
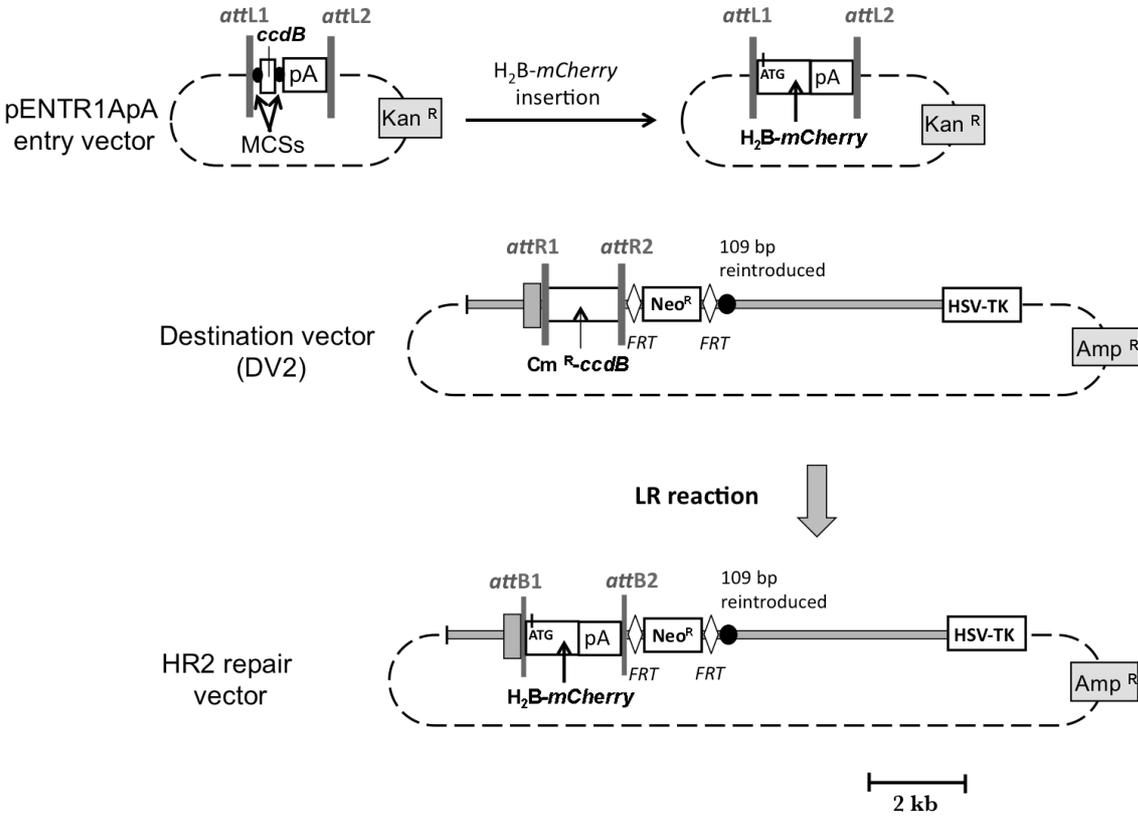
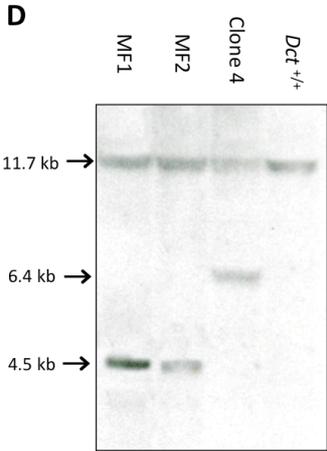
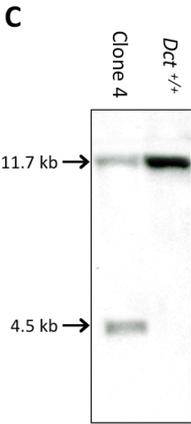
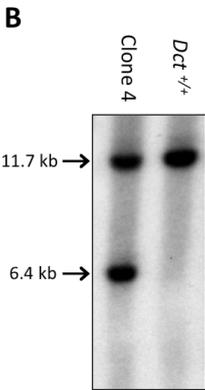
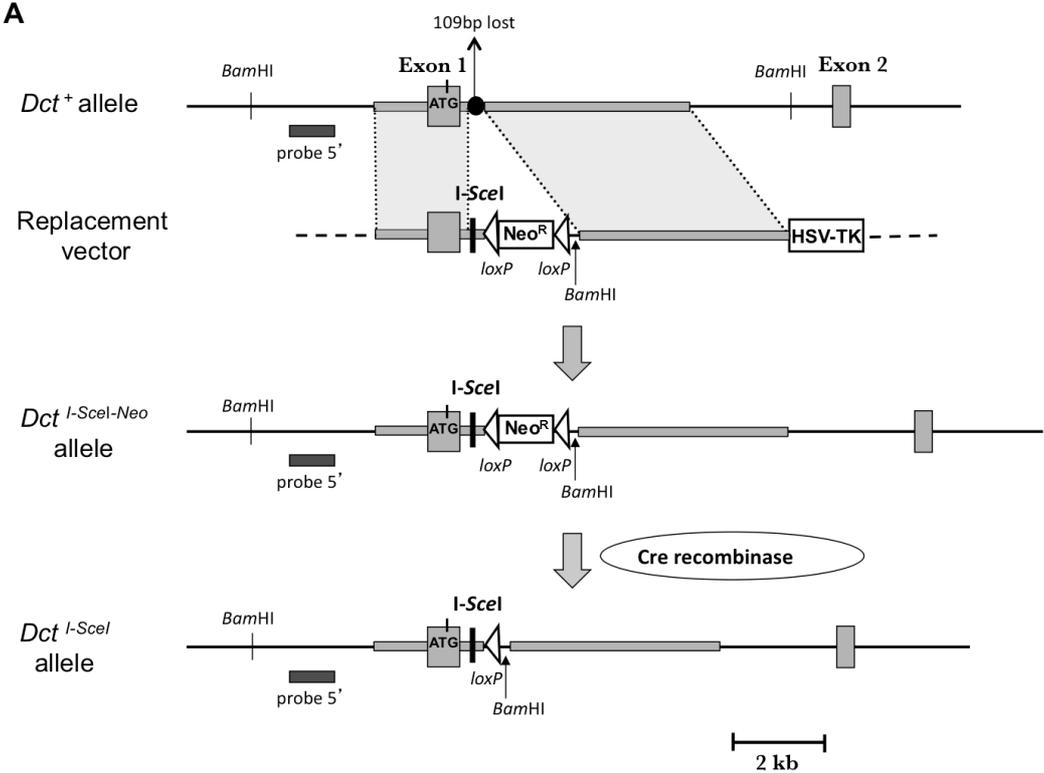


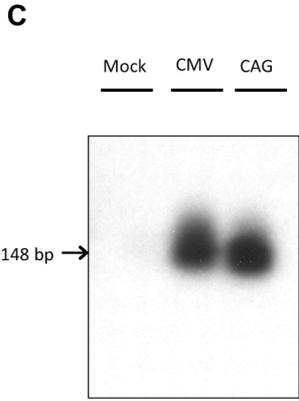
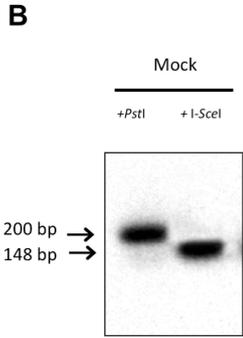
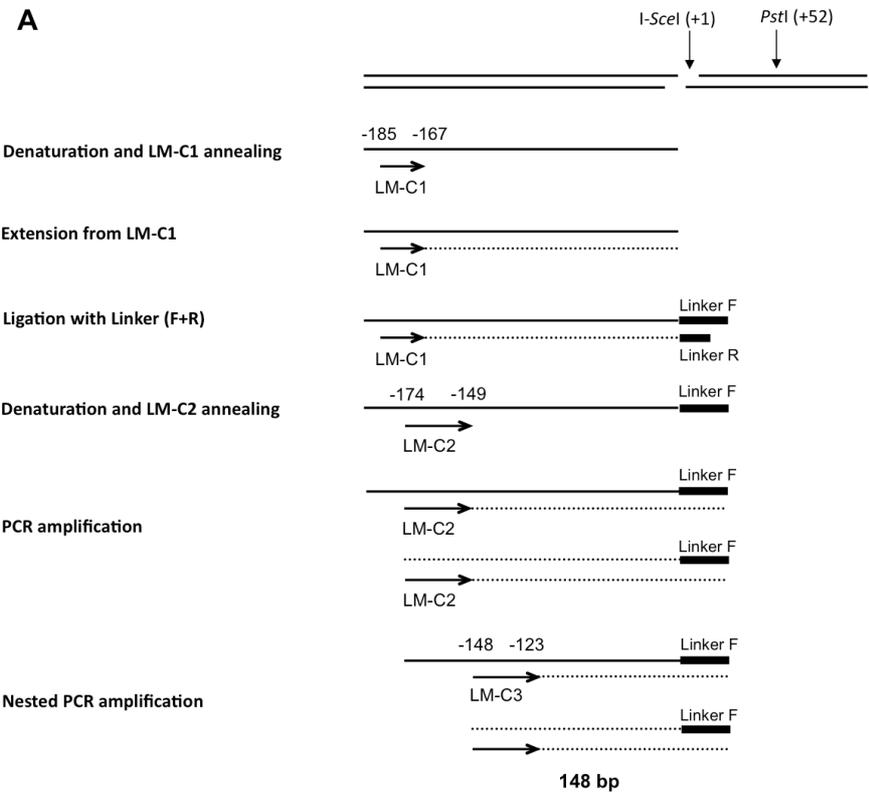
FIG.2.



**FIG.3.**



**FIG.4.**



**FIG.5.**

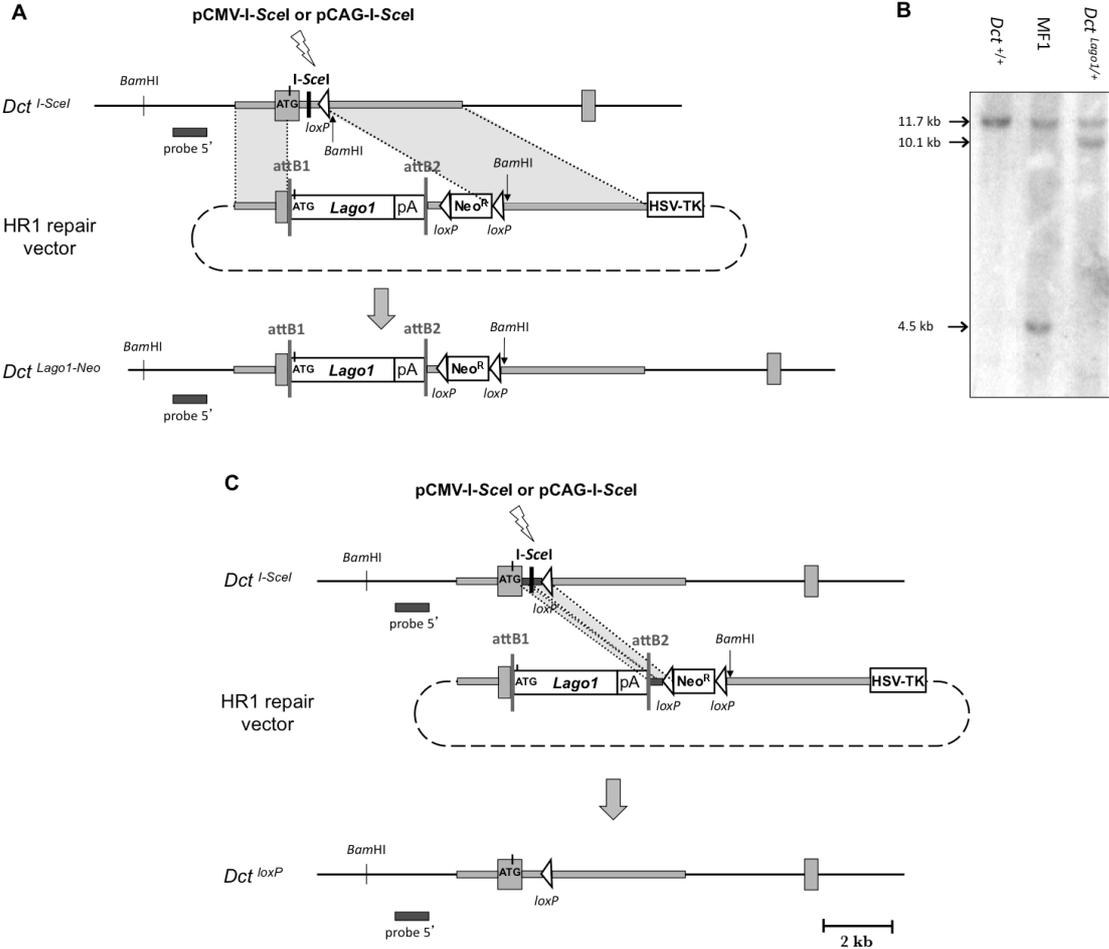
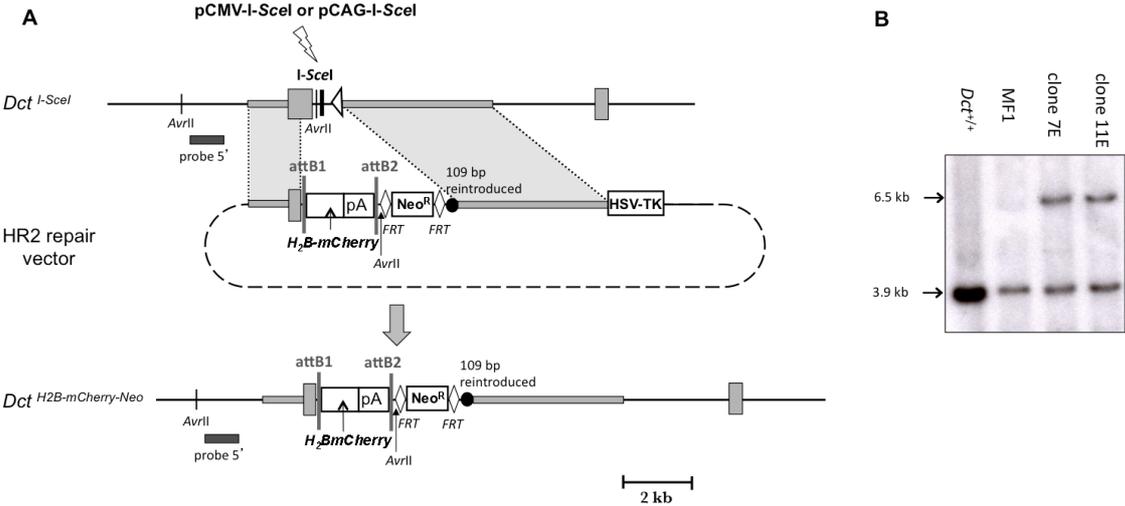


FIG.6.



## 2 Discussion et perspectives

La première étape du développement de l'outil consistait à intégrer un site I-SceI au locus *Dct* par la méthode conventionnelle de recombinaison homologue. Nous avons obtenu des cellules ES *Dct*<sup>I-SceI/+</sup>. La deuxième étape consistait à transfecter les cellules ES *Dct*<sup>I-SceI/+</sup> simultanément avec un vecteur d'expression de la méganucléase I-SceI et un vecteur de réparation qui permettrait l'intégration d'une séquence d'intérêt au locus *Dct*. L'induction d'un DSB par la méganucléase I-SceI dans l'intron 1 de *Dct* devait favoriser le mécanisme de RH et permettre l'intégration de la séquence d'intérêt avec une fréquence supérieure à celle obtenue par une méthode conventionnelle de recombinaison homologue. Dans les stratégies que nous avons développées, la fréquence de la recombinaison homologue en présence de la méganucléase I-SceI et d'un DSB au locus *Dct* n'est pas supérieure à la fréquence de la recombinaison homologue de la méthode conventionnelle.

Les résultats obtenus au fur et à mesure de ce projet nous ont amené à nous poser plusieurs questions : (1) le site I-SceI intégré au locus *Dct* peut-il être coupé par la méganucléase *in vivo* ? (2) Le site I-SceI inséré dans le gène *Dct* est-il placé de façon optimale ? (3) La méganucléase I-SceI est-elle aussi spécifique que nous le pensions ? (4) Les cellules ES se prêtent-elles à une telle stratégie ?

### 2.1 Le site I-SceI intégré au locus *Dct* : peut-il être coupé par la méganucléase *in vivo* ?

Pour induire un DSB au site I-SceI, nous avons testé deux plasmides d'expression de la MN : pCMV-I-SceI et pCAG-I-SceI (fournis par M. Cohen-Tannoudji, Institut Pasteur). Le plasmide pCMV-I-SceI a déjà prouvé son efficacité dans l'induction d'un DSB à un site I-SceI intégré dans le génome murin (Choulika, *et al.*, 1995; Cohen-Tannoudji, *et al.*, 1998). Le plasmide pCAG-I-SceI a été construit à partir du vecteur pCMV-I-SceI dont le promoteur CMV a été remplacé par le promoteur composite CAG (Okabe, *et al.*, 1997). L'efficacité du plasmide pCAG-I-SceI a été testée en culture cellulaire sur une séquence extrachromosomique. Cette séquence comprenait un site I-SceI qui séparait un gène *LacZ* en deux parties, ce qui le rendait non fonctionnel. L'induction d'un DSB au site I-SceI entraînait la RH entre les deux motifs et donc l'expression de la  $\beta$ -galactosidase. pCAG-I-SceI était plus

efficace que pCMV-I-SceI pour digérer cette séquence et induire la RH entre les deux parties du gène *LacZ* (Cohen-Tannoudji, non publié). Nous n'avons pas remis en question la capacité des plasmides pCAG-I-SceI et pCMV-I-SceI à exprimer une méganucléase capable de couper son site de reconnaissance. Néanmoins, pour nous assurer que la protéine I-SceI est exprimée dans les cellules ES transfectées, nous pourrions réaliser un Western blot avec un anticorps dirigé contre I-SceI. Pour cela, des cellules ES *Dct*<sup>I-SceI/+</sup> seraient électroporées avec pCAG-I-SceI ou pCMV-I-SceI. Après quelques heures, les protéines cellulaires seraient extraites et un Western blot réalisé sur ces extraits protéiques.

Nous nous sommes demandés si le site I-SceI inséré dans la séquence du gène *Dct* pouvait être coupé par la méganucléase *in vivo*. En effet, il est envisageable que, pour des raisons de conformation de la chromatine, le site I-SceI ne soit pas accessible à la protéine. Nous avons donc effectué une expérience de *ligation-mediated-PCR* (LM-PCR) pour répondre à cette question. A l'origine, cette technique a été développée pour étudier la structure de l'ADN, analyser les interactions entre une protéine et l'ADN ou les dommages à l'ADN. Elle permet de détecter l'existence d'une cassure à un site donné du génome (Besaratina and Pfeifer, 2009). L'utilisation de cette technique sur les cellules ES *Dct*<sup>I-SceI/+</sup> nous a permis de montrer que les plasmides pCMV-I-SceI et pCAG-I-SceI sont effectivement capables d'induire un DSB au locus *Dct*. La technique que nous avons développée nous donne une réponse qualitative quant à la création de la coupure au site I-SceI. Pour l'instant, nous n'avons pas obtenu de résultat quantitatif. Nous ne pouvons pas évaluer quelle proportion des cellules électroporées est coupée par I-SceI. Il est possible que la coupure dont l'existence est démontrée par LM-PCR ne soit pas assez fréquente pour favoriser une recombinaison homologe. Deux plasmides supplémentaires qui permettent d'exprimer l'enzyme I-SceI ont été testés par LM-PCR : pBL133-I-SceI qui contient un promoteur CMV et une séquence de localisation nucléaire (*nls*) et pBL464-pCBASceI qui contient un promoteur CAG et une séquence *nls* (fournis par B. Lopez, CEA). Des résultats préliminaires en LM-PCR indiquent que les plasmides pBL133-I-SceI et pBL464-pCBASceI coupent plus efficacement l'allèle *Dct*<sup>I-SceI/+</sup> que les plasmides pCMV-I-SceI et pCAG-I-SceI utilisés dans nos expériences. Ces deux groupes de plasmides se distinguent par la présence ou l'absence d'une séquence *nls*. Celle-ci adresse la méganucléase au noyau des cellules. Il est possible que la présence de cette séquence favorise l'induction d'un DSB. L'absence de données

quantitatives nous empêche de conclure sur ce point. Il serait intéressant de compléter ces données avec des résultats quantitatifs de LM-PCR.

L'influence du niveau d'expression de la méganucléase sur la fréquence de recombinaison homologue au locus *Dct* est une question intéressante. Lors des expériences avec les vecteurs HR1 et HR2, le plasmide pCAG-I-*SceI* a donné de meilleurs résultats que le plasmide pCMV-I-*SceI*. Par deux fois, pCAG-I-*SceI* a produit un clone recombinant homologue, tandis que pCMV-I-*SceI* n'a pas produit de clone recombinant homologue. Nous ne pouvons pourtant pas conclure que le plasmide pCAG-I-*SceI* est plus efficace, car un clone recombinant homologue n'est pas significativement différent de zéro clone. Si les plasmides pBL133-I-*SceI* et pBL464-pCBA*SceI* étaient effectivement plus efficaces dans l'induction d'un DSB, il faudrait réaliser de nouvelles expériences avec ces plasmides. Le nombre de clones recombinants homologues obtenus avec ces plasmides serait alors comparé au nombre de clones obtenus avec les plasmides déjà testés. Au vu du temps nécessaire à la réalisation d'une telle expérience, il n'était pas envisageable de faire ce test avant la fin de ma thèse. Récemment, une étude de ciblage de gène a été réalisée dans des cellules d'ovaire de Hamster chinois CHO-K1 en utilisant un vecteur lentiviral qui exprime I-*SceI* (Izmiryan, *et al.*, 2011). Un vecteur lentiviral exprimant I-*SceI* couperait peut-être son site de reconnaissance intégré au locus *Dct* plus efficacement que des plasmides d'expression d'I-*SceI*.

Un autre moyen d'évaluer l'événement de cassure au site I-*SceI* est d'observer le phénomène de NHEJ après expression de la MN (Weinstock, *et al.*, 2006). En effet, il est établi que le mécanisme de NHEJ aboutit à de courtes insertions ou délétions au site de cassure. Dans ce cas, ce n'est pas le phénomène de cassure qui est observé, mais sa conséquence, à savoir la réparation par NHEJ. En conséquence, les cellules qui ont été réparés sans modification du site I-*SceI* ou par recombinaison homologue avec la chromatide sœur ne sont pas prises en compte dans cette méthode d'analyse de la cassure. Dans nos expériences, un grand nombre de clones résistants à la néomycine et au gancyclovir ont été obtenus qui n'étaient pas recombinants homologues. Ces clones résistants ont intégré le vecteur de réparation. Or les cellules avaient été transfectées avec au moins le même nombre de molécules de plasmide d'expression de I-*SceI* que de vecteur de réparation. Environ la moitié des cellules ayant intégré le vecteur de réparation doivent donc avoir également reçu le



plasmide d'expression. Dans ces cellules, I-SceI a normalement induit un DSB qui a été réparé vraisemblablement par NHEJ, puisque le vecteur de réparation n'a pas été intégré au site de coupure. Il serait donc intéressant de séquencer les clones résistants à la néomycine et au gancyclovir au niveau du site I-SceI. Si la séquence est modifiée par rapport à l'allèle  $Dct^{I-SceI/+}$ , cela signifiera que le site a été coupé et réparé par NHEJ. Nous avons réalisé une analyse préliminaire sur 16 clones résistants, non recombinaison homologues. Le site I-SceI n'était pas modifié sur ces 16 clones. Pour chacune des expériences, 144 clones ont été prélevés. Il serait intéressant d'analyser le site I-SceI sur un plus grand nombre de ces clones. Cependant, notre analyse préliminaire suggère que le site I-SceI n'a pas été coupé chez les clones résistants non recombinaison homologues. On peut donc se demander quelle proportion des cellules  $Dct^{I-SceI/+}$  a effectivement subi un DSB induit par I-SceI.

Le plasmide pCMV-I-SceI augmente la fréquence de recombinaison homologue au locus *Villin* d'un facteur 100 (Cohen-Tannoudji, *et al.*, 1998), alors qu'il n'a pas permis d'augmenter la fréquence de recombinaison homologue au locus *Dct*. Même si d'autres systèmes d'expression, comme les plasmides pBL133-I-SceI, pBL464-pCBA-SceI ou un vecteur lentiviral qui exprime I-SceI, améliorent l'efficacité de coupure au site *Dct* (d'un facteur 2 à 10 par exemple), il me semble peu probable que la fréquence de recombinaison homologue à ce locus soit augmentée d'un facteur 100. Les différences de fréquence aux locus *Villin* et *Dct* suggèrent que le DSB a un effet sur la fréquence de recombinaison homologue qui est largement dépendant du locus considéré. Il serait donc intéressant d'analyser d'où peut venir cette différence entre les locus.

Nous avons montré en utilisant la méthode LM-PCR que I-SceI coupe son site de reconnaissance, mais nous ne savons pas si le taux de coupure est suffisant pour induire efficacement une réparation par recombinaison homologue. Il est vraisemblable que la structure tridimensionnelle de la chromatine au locus *Dct* n'est pas favorable à l'accès de la méganucléase I-SceI et/ou des protéines de réparation du DSB. Par ailleurs, l'influence de la méthylation sur l'efficacité de ciblage de gène favorisé par un DSB est encore mal connue. Le statut méthylé ou déméthylé dans lequel se trouve un gène ciblé par une nucléase influence probablement les événements de cassure double-brin, puis de réparation de cette cassure. Un moyen de l'étudier serait de traiter les cellules ES  $Dct^{I-SceI/+}$  par l'azacytidine avant d'induire un DSB par I-SceI, et d'étudier l'influence de ce traitement sur l'événement de cassure par LM-PCR d'abord puis sur la fréquence de recombinaison homologue.

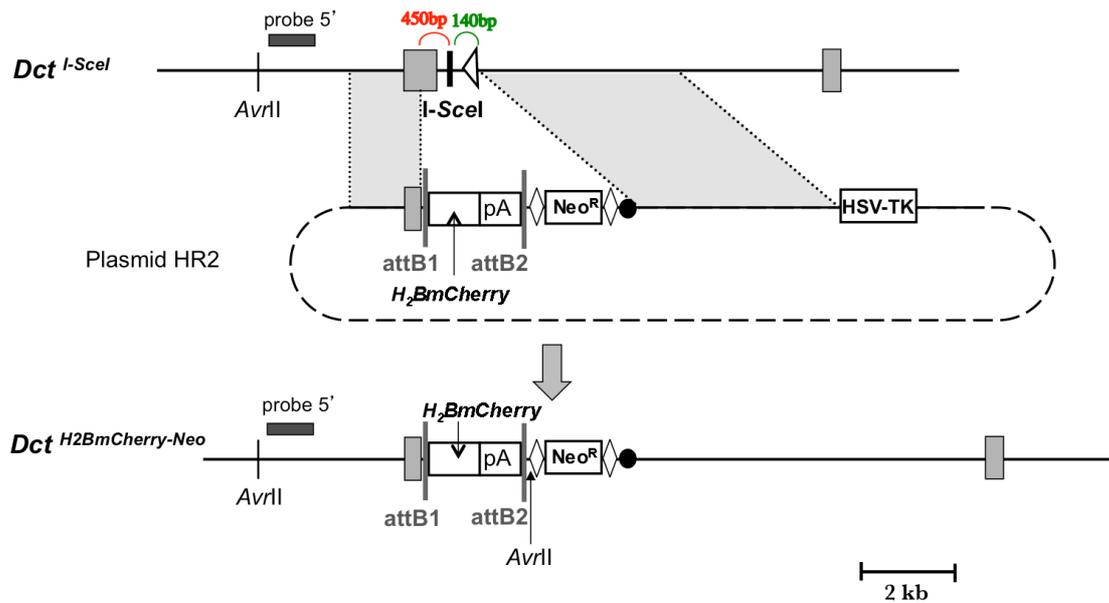


Figure 40 : Schéma de RH favorisée par I-SceI au locus *Dct*

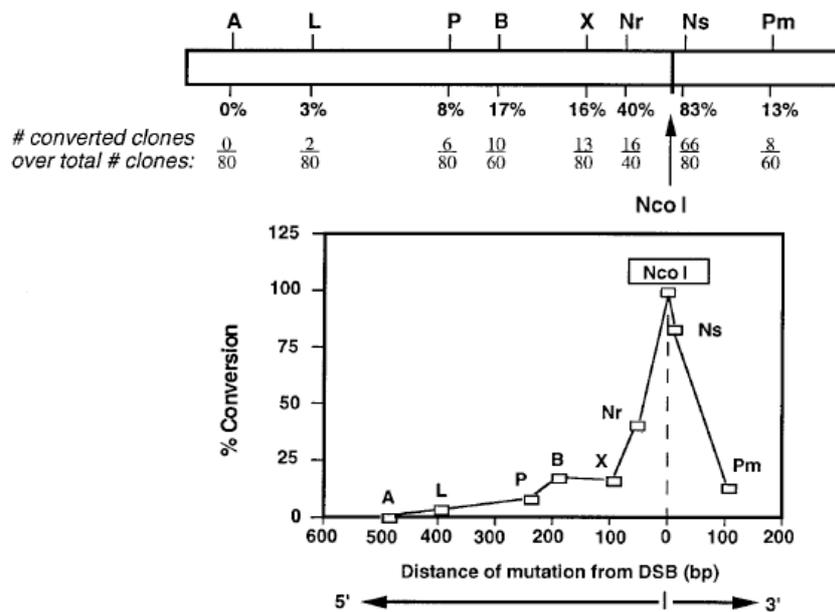


Figure 41 : Fréquence d'insertion de mutations silencieuses selon de la distance à la cassure (Elliott, *et al.*, 1998)

Un gène *Néo* interrompu d'un site *I-SceI* qui le rendait non fonctionnel a été inséré dans le génome de cellules ES. *I-SceI* était situé à la place du site *NcoI* du gène *Néo*. Un plasmide d'expression d'*I-SceI* a ensuite été introduit dans les cellules en même temps qu'un vecteur de réparation. Le vecteur de réparation permettait la restauration du site *NcoI* par RH rendant le gène *Néo* fonctionnel. Il introduisait en même temps des mutations silencieuses dans la séquence du gène *Néo* réparé par RH. Cette figure présente la fréquence d'insertion de chaque mutation silencieuse en fonction de sa localisation par rapport au site *NcoI*, c'est-à-dire au lieu de cassure.

Nous avons vu dans ce manuscrit de nombreux exemples de ciblage de gène favorisés par un DSB qui illustrent l'efficacité accrue de la recombinaison homologue après induction d'une cassure au locus ciblé. Cependant, le nombre de locus ciblés par cette méthode est encore très faible. Trois locus naturels ont été efficacement ciblés à l'aide de la méganucléase I-SceI dans des cellules ES de souris, les locus *Villin*, *Hprt* et *Dbx1* (Bielle, *et al.*, 2005; Cohen-Tannoudji, *et al.*, 1998; Donoho, *et al.*, 1998).

## 2.2 Le site I-SceI inséré dans le gène *Dct* est-il placé de façon optimale ?

Le site I-SceI a été inséré dans l'intron 1 du gène *Dct*. Nous souhaitons réaliser un *knock in* qui permette d'exprimer une protéine d'intérêt dépourvue d'acides aminés de la protéine DCT. Pour cette raison, le site I-SceI est situé à 450 pb de la région d'homologie en 5' et à 140 pb de l'homologie en 3' (Figure 40). Il a été montré que la recombinaison homologue a lieu dans les 200 pb qui encadrent immédiatement le site de cassure double-brin (Figure 41) (Elliott, *et al.*, 1998). Je ne détaillerai pas ici le défaut du premier vecteur de réparation HR1 (qui porte le gène *Lago1*) ; il est présenté dans l'article. La deuxième expérience avec le vecteur HR2 (qui porte le gène *H2B-mCherry*) montre que les courtes régions d'homologie de HR1 n'expliquent pas la faible efficacité de la recombinaison homologue, même s'il s'agissait d'un défaut important dans la construction. Les résultats obtenus par Elliott et collaborateurs soulignent l'importance potentielle des courtes régions d'homologie. Ils montrent aussi que les régions d'homologie utilisées pour la recombinaison homologue doivent être au plus près de la cassure. Nous avons vu dans l'introduction que le mécanisme de recombinaison homologue commence par une digestion des extrémités de l'ADN. Plus les séquences à digérer pour atteindre les régions d'homologie de part et d'autre de la cassure sont courtes, plus il y a de chance que la recombinaison homologue soit efficace. De fait, la majorité des expériences de recombinaison homologue après induction d'un DSB utilise des régions d'homologie qui jouxtent directement la cassure (Connelly, *et al.*, 2010; Donoho, *et al.*, 1998; Elliott, *et al.*, 1998; Hockemeyer, *et al.*, 2009). Ainsi les séquences de 450 pb et de 140 pb d'un côté de l'autre du site I-SceI réduisent probablement l'efficacité de la recombinaison homologue au locus *Dct* après DSB.

Dans d'autres expériences, des régions d'homologie qui ne jouxtaient pas directement la cassure ont été utilisées avec succès (Cabaniols and Paques, 2008; Cabaniols, *et al.*, 2010;

Cohen-Tannoudji, *et al.*, 1998). Une certaine distance entre le site de cassure et les régions d'homologie n'empêche donc pas qu'une recombinaison homologue soit facilitée par un DSB. Cependant, nous avons vu dans l'Introduction (Chap1, 5.1.2) que la fréquence de la recombinaison homologue semble diminuer lorsque les séquences homologues sont distantes du site de coupure (Cabaniols and Paques, 2008). Il serait intéressant de savoir quelle est l'influence précise de la distance entre le DSB et les régions d'homologie sur la fréquence de recombinaison homologue. D'autres études seraient nécessaires pour répondre à cette question.

En conclusion, la localisation du site I-*SceI* au sein de l'intron 1 de *Dct* réduit probablement l'efficacité de la recombinaison homologue induite par DSB. Au vu des fortes efficacités obtenues dans d'autres études dans lesquelles le site I-*SceI* était distant des régions d'homologie, je ne pense pas que l'emplacement du site de coupure soit la cause de la faible efficacité de la stratégie cGPS au locus *Dct*. S'il avait été possible de modifier l'emplacement du site I-*SceI*, je l'aurais inséré dans l'exon 1, au niveau de la séquence ATG. Cette modification n'était pas envisageable pendant ma thèse car elle aurait demandé beaucoup de temps pour un bénéfice incertain.

### 2.3 La méganucléase I-*SceI* est-elle aussi spécifique que nous le pensions ?

I-*SceI* est considérée comme une méganucléase qui reconnaît spécifiquement une séquence de 18 pb. La longueur de cette séquence fait qu'elle ne doit pas être présente dans le génome des Mammifères. Récemment, le taux de coupures de I-*SceI* à des sites chromosomiques non attendus (coupures *off-target*) a été évalué dans des cellules HT-1080, dérivées d'un fibrosarcome de l'homme (Petek, *et al.*, 2010). Soixante et un pour cent des événements d'intégration d'un virus adéno-associé qui font suite à une coupure par I-*SceI* se produisent à des sites *off-target*. De plus, la caractérisation des sites *off-target* par des méthodes *in silico* ou *in vitro* ne permet pas de prédire les sites de coupure *in vivo* (Petek, *et al.*, 2010). Il est donc possible que la méganucléase I-*SceI* induisent des coupures *off-target* dans les cellules *Dct*<sup>I-*SceI*+</sup>, ce qui diminue les chances d'obtenir un ciblage au locus *Dct*. Cela ne peut pourtant pas expliquer la faible efficacité de la stratégie cGPS au locus *Dct*, puisque

de nombreuses études ont démontré qu'I-SceI favorise la recombinaison homologue à d'autres sites chromosomiques que *Dct* indépendamment des coupures *off-target*.

La capacité des ZFN à entraîner des coupures *off-target* est un sujet qui préoccupe beaucoup les chercheurs qui utilisent ces nucléases (Carroll, 2011). En effet, même si une ZFN est développée pour reconnaître une séquence chromosomique bien spécifique, il n'est pas rare de découvrir qu'elle coupe des séquences différentes. Ce phénomène est encore mal compris et donc peu maîtrisable, or les ZFN sont très prometteuses pour la thérapie génique. Dans ce type d'applications, il n'est pas tolérable que des événements *off-target* se produisent. Pour cette raison, des méthodes visant à prédire et tester des sites de coupure *off-target* des ZFN sont développées (Gabriel, *et al.*, 2011; Pattanayak, *et al.*, 2011).

## 2.4 Les cellules ES se prêtent-elles à une telle stratégie ?

Nous avons vu dans l'Introduction (Chap1, 4.2) que la recombinaison homologue est restreinte aux phases S et G2 du cycle cellulaire (Rothkamm, *et al.*, 2003). La phase S représente 75% du cycle cellulaire chez les cellules ES (Savatier, *et al.*, 2002). Le temps que dure la phase S favorise donc le mécanisme de recombinaison homologue dans les cellules ES (Tichy, *et al.*, 2010). Par ailleurs, le marquage  $\gamma$ -H2AX révèle les DSB présents dans une cellule. Il a été montré que les cellules ES présentent un marquage  $\gamma$ -H2AX fort comparé aux cellules somatiques (Tichy and Stambrook, 2008). Cela signifie qu'en l'absence d'expression d'endonucléases, le génome des cellules ES est la cible de cassures qu'elle doivent réparer. Dans le cas des cellules *Dct<sup>I-SceI/+</sup>*, l'introduction d'un vecteur d'expression de la méganucléase I-SceI crée (parfois) une cassure supplémentaire, ciblée au locus *Dct*. Le locus *Dct* est alors en compétition avec les autres sites de coupure du génome pour recevoir le vecteur de réparation. Ce phénomène ne doit pas favoriser un ciblage de gène efficace au locus *Dct*. Pourtant plusieurs études démontrant une importante efficacité de la recombinaison homologue induite par I-SceI ont été réalisées dans des cellules ES (Cohen-Tannoudji, *et al.*, 1998; Donoho, *et al.*, 1998; Smih, *et al.*, 1995). La faible efficacité de la recombinaison homologue dans notre étude ne peut donc pas s'expliquer uniquement par la survenue de nombreux DSB dans le génome des cellules ES *Dct<sup>I-SceI/+</sup>*.

Il a été montré qu'il existe des variations d'efficacité de la recombinaison homologue selon le type cellulaire utilisé. Lombardo et collaborateurs ont utilisé des vecteurs lentiviraux qui expriment une ZFN dont la cible est le gène *CCR5* (*chemokine (C-C motif) receptor 5*). Ils ont montré que la fréquence des événements de recombinaison homologue au locus *CCR5* favorisée par un DSB atteint 50% dans des lignées immortalisées humaines, la lignée dérivée d'une érythroleucémie myéloïde K-562 et la lignée de cellules lymphoblastoïdes Jurkat, tandis qu'elle n'est que de 5% dans des cellules ES humaines et de 0,1% dans des progéniteurs de sang de cordon CD34+ (Lombardo, *et al.*, 2007). Ces données suggèrent que l'efficacité du ciblage de gène favorisé par un DSB est plus élevée dans les cellules immortalisées que dans les cellules en culture primaire, et qu'elle est plus faible dans les cellules non différenciées (ES et iPS) que dans les cellules différenciées. Les nombreux DSB observés dans les cellules ES explique peut-être en partie ces résultats.

## 2.5 Conclusion

Jusqu'à présent, une seule étude a fait état d'un faible taux de RH suite à l'induction d'un DSB dans les cellules de Mammifères. Il a été observé que le NHEJ est le mécanisme majoritaire, même en présence d'un vecteur de réparation (Lukacsovich, *et al.*, 1994). A travers notre projet, nous avons montré que l'induction d'une cassure double-brin au site I-SceI intégré dans l'intron 1 de *Dct* dans les cellules ES murines n'a pas favorisé la recombinaison homologue à ce locus. Nous venons de passer en revue plusieurs facteurs qui ne favorisent pas une fréquence de recombinaison homologue optimale, comme les coupures *off-target* de I-SceI, les caractéristiques des cellules ES, ou l'emplacement du site I-SceI dans l'intron de *Dct*. Cependant ces facteurs n'expliquent pas la faible fréquence de recombinaison homologue obtenue dans notre étude. A mon avis, la structure tridimensionnelle de la chromatine au locus *Dct* est le facteur principalement responsable de cette faible fréquence. Du fait de la structure de la chromatine, la méganucléase peut avoir difficilement accès à son site de coupure. L'étude plus approfondie de l'événement de coupure au site I-SceI par LM-PCR serait importante pour savoir quel est le pourcentage des cellules qui présentent effectivement une coupure double-brin induite par la méganucléase. Il n'est pas certain non plus que le faible taux de coupure au locus *Dct* soit la cause de la pauvre efficacité de recombinaison homologue. Il faudrait pouvoir comparer le taux de coupure au locus *Dct* à celui obtenu à un autre site chromosomique connu pour avoir une forte efficacité de

recombinaison homologue favorisée par un DSB dans le même type cellulaire, comme le locus *Villin*. Enfin, il est établi que la fréquence de recombinaison homologue dans les cellules ES varie selon le locus ciblé entre  $10^{-5}$  et  $10^{-8}$ . Je pense qu'il peut en être de même pour la fréquence de recombinaison homologue favorisée par un DSB. Le ciblage de gène favorisé par un DSB serait donc locus-dépendant.

L'objectif de notre projet était de réaliser des ciblage de gène de façon répétée à un locus précis chez la souris, via l'utilisation d'I-SceI. En comparaison, l'emploi de ZFN pour favoriser des ciblage de gène chez la souris est une méthode qui est de plus en plus documentée. Il existe au moins deux exemples de ciblage de gène réalisés grâce à des ZFN dans les cellules ES de souris, dans lesquelles une intégration par *knock in* de gènes rapporteurs a été effectuée à des locus endogènes, *Rosa26* et *H3.3* (Connelly, *et al.*, 2010; Goldberg, *et al.*, 2010). Les ZFN, comme les méganucléases à façon, ne nécessitent pas l'insertion préalable d'un site de reconnaissance au locus ciblé. Pour cette raison, ces enzymes peuvent être directement utilisées dans l'embryon ou même dans un tissu particulier de l'animal. Récemment plusieurs études ont utilisé des ZFN dans des ciblage de gène réalisés directement dans l'embryon de souris. Cette technique a d'abord été employée pour promouvoir la réparation par NHEJ (Carbery, *et al.*, 2010), puis l'intégration par *knock in* de transgènes à des locus endogènes (Cui, *et al.*, 2011; Meyer, *et al.*, 2010). L'induction d'un DSB via une ZFN directement dans le foie de l'animal a également favorisé la RH à un locus endogène (Li, *et al.*, 2011). Au-delà du travail que j'ai réalisé au cours de ma thèse, les perspectives offertes par les méganucléases à façon et les nucléases à doigts de zinc sont donc très prometteuses et contribuent à faire évoluer radicalement les techniques de transgénèse chez les Mammifères.



## **Chapitre 2 : Etude du patron d'expression des gènes rapporteurs *Lago1* et *H<sub>2</sub>BmCherry* inséré au locus *Dct***

L'objectif de cette étude était de générer de nouveaux systèmes rapporteurs de l'expression du gène *Dct* chez la souris, via la production de cellules ES *knock in*. Le premier gène rapporteur choisi a été *Lago1*. Ce gène a été construit à partir du gène bactérien *LacZ*. Sa séquence nucléotidique a été modifiée pour s'adapter à la machinerie de traduction des Mammifères. Le gène *Lago1* devrait donc être exprimé plus fortement que le gène *LacZ* si tous deux sont exprimés sous le contrôle d'un même promoteur. Le second gène rapporteur que nous avons testé était le gène *H<sub>2</sub>B-mCherry*. Il n'existait pas de protéine fluorescente rouge permettant d'identifier les mélanocytes et leurs précurseurs chez la souris et nous pensions que mCherry pourrait remplir ce rôle. Ces nouveaux rapporteurs devaient compléter la gamme des outils existants qui permettent de suivre le lignage mélanocytaire. Avec ces gènes rapporteurs, nous avons testé l'efficacité de notre outil de modification génétique au locus *Dct*. Une fois l'outil validé par un gène rapporteur, nous aurions pu introduire n'importe quel gène d'intérêt au locus *Dct* dans les cellules ES de génotype *Dct<sup>L-Scel/+</sup>* et générer la souris *knock in* correspondante.

### **1 Etude du profil d'expression des gènes rapporteurs *Lago1* et *H<sub>2</sub>B-mCherry* dans les mélanocytes en culture**

#### **1.1 Objectif et stratégie**

Après avoir obtenu des cellules ES *Dct<sup>Lago1-Néo/+</sup>* et *Dct<sup>H2BmCherry-Néo/+</sup>* portant respectivement les gènes rapporteurs *Lago1* et *H2B-mCherry*, nous avons voulu caractériser leur expression dans des cellules en culture. L'expression des gènes rapporteurs était attendue dans les mélanocytes et leurs précurseurs. Nous avons donc utilisé une méthode de différenciation des cellules ES en mélanocytes (Yamane, *et al.*, 1999). Des cellules ES portant le transgène *Dct-LacZ* (fournies par L. Larue) ont été utilisées comme témoin positif (Pla, *et al.*, 2004).

Les cellules ES *Dct-LacZ*, *Dct<sup>Lago1-Néo/+</sup>* et *Dct<sup>H2BmCherry-Néo/+</sup>* ont été étalées à J0 et maintenues jusqu'à J41 dans un milieu qui favorise leur différenciation en mélanocytes. A

J41, les cellules pigmentées contenaient de plus en plus de grains de mélanines, signe d'un état sénescence. A différents stades du processus de différenciation, des cellules *Dct-LacZ* et *Dct<sup>Lago1-Néo/+</sup>* ont été colorées en présence de X-gal. Les cellules *Dct<sup>H2BmCherry-Néo/+</sup>* en culture ont été observées au microscope à fluorescence, puis l'expression de la protéine mCherry a été analysée par immunocytofluorescence. *Lago1* possède une séquence de localisation nucléaire (*nls*). H<sub>2</sub>BmCherry est une protéine de fusion entre l'histone (H<sub>2</sub>B) et la protéine mCherry qui est elle-même dérivée de la protéine fluorescente red 1 (mRFP1) des cnidaires. L'expression des protéines *Lago1* et *H<sub>2</sub>B-mCherry*, dans les cellules *Dct<sup>Lago1-Néo/+</sup>* et *Dct<sup>H2BmCherry-Néo/+</sup>* doit être nucléaire, tandis que l'expression de *LacZ* dans les cellules *Dct-LacZ*, doit être cytoplasmique.

## 1.2 Matériels et méthodes

### *Différenciation des cellules ES en mélanocytes*

Des cellules de la lignée stromale ST2 (fournies par L. Larue) ont été amplifiées dans un milieu RPMI supplémenté avec 5% de sérum de veau fœtal, 100 U/mL de pénicilline, 0,1 mg/mL de streptomycine et 50 µM de β-mercaptoéthanol. Les cellules ont été étalées dans des plaques de culture 6 puits à raison de 250.000 cellules/puits. Trois jours plus tard, la confluence était atteinte. Pour l'expérience d'immunocytofluorescence, les cellules ST2 ont été étalées dans des chambres microscopiques (LabTek 0,8cm<sup>2</sup>/puits), à raison de 21.000 cellules/puits. Les cellules ES *Dct-LacZ*, *Dct<sup>Lago1-Néo/+</sup>* et *Dct<sup>H2BmCherry-Néo/+</sup>* ont été préalablement amplifiées (voir article M. Fenina, en préparation). A J0, les cellules ES ont été étalées sur les cellules ST2 confluentes. Pour induire leur différenciation, les cellules ES ont été cultivées dans un milieu α-MEM, supplémenté avec 10% de sérum de veau fœtal, 100 U/mL de pénicilline, 0,1 mg/mL de streptomycine, 10<sup>-7</sup> M dexaméthasone, 20 pM bFGF et 10 pM choléra-toxine. Le milieu était changé deux fois par semaine.

### *Mesure de l'activité β-galactosidase des cellules*

Les cellules *Dct-LacZ* et *Dct<sup>Lago1-Néo/+</sup>* ont été fixées par de la paraformaldéhyde (PFA) à 4%, dans un tampon phosphate (PBS) 1X. Après fixation, l'activité β-galactosidase a été révélée dans une solution X-gal, composée de 5 mM ferricyanure, 5 mM ferrocyanure, 0,01% tween 20, 2 mM MgCl<sub>2</sub> et 10 µg/mL de X-gal. Les cellules ont été incubées dans cette solution à

37°C pendant 12 h puis rincées en PBS 1X. Les cellules ont été observées et prises en photo au microscope en champ clair.

#### *Révélation de l'expression de mCherry par immunocytofluorescence*

L'anticorps primaire utilisé est un anticorps polyclonal de lapin anti-red fluorescent protein (living color Ds Red clontech 632496) dilué au 1/100<sup>ème</sup>. L'anticorps secondaire est un anticorps d'âne anti-lapin (dylight 594), dilué au 1/200<sup>ème</sup>. L'expérience d'immunocytofluorescence a été réalisée sur des cellules ES *Dct<sup>H2BmCherry-Néo/+</sup>* différenciées en chambres microscopiques. Des cellules contenant le rapporteur H<sub>2</sub>B-RFP ont été utilisées comme témoin positif. Les cellules ont été fixées 10 minutes en PFA 4%, rincées abondamment en PBS, perméabilisées en PBS/0,25% Triton X100 10 minutes, lavées à nouveau et saturées avec une solution de PBS/10% de sérum de cheval/0,1% Tween 20 pendant 20 minutes. L'anticorps primaire a été dilué dans une solution de PBS/0,1% Tween 20 et 10% sérum de cheval. Les cellules ont été incubées une nuit à 4°C avec l'anticorps primaire. Le lendemain, les cellules ont été lavées puis incubées 1 heure avec l'anticorps secondaire, dilué dans une solution de PBS/0,1% Tween 20. Les cellules ont été ensuite lavées en PBS, incubées 5 minutes dans une solution de Hoescht (dilué au 1/200<sup>ème</sup>), puis fixées entre lame et lamelle pour observation au microscope à fluorescence.

### 1.3 Résultats

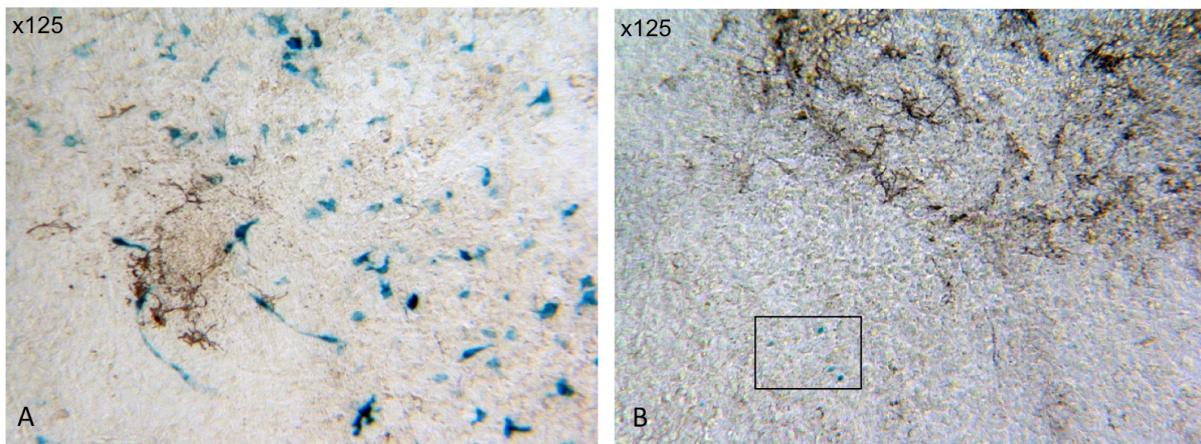
#### *Différenciation des cellules ES en mélanocytes*

Pour induire la différenciation des cellules ES en mélanocytes, les cellules ES *Dct-LacZ*, *Dct<sup>Lago1-Néo/+</sup>* et *Dct<sup>H2BmCherry-Néo/+</sup>* ont été étalées sur un tapis de cellules stromales ST2 et cultivées pendant 41 jours dans un milieu supplémenté en dexaméthasone, toxine cholérique et bFGF. Des nombres différents de cellules ES ont été étalés à J0 : 500, 750 et 1.000 cellules par puits, dans des plaques 6 puits. A J7, les cellules ES ont commencé à former des amas de cellules, identifiés comme des corps embryoides. A partir de J24, des cellules pigmentées ont été observées. Après 38 jours de culture, des cellules devenaient très fortement pigmentées, signe d'une sénescence. La méthode de différenciation des cellules ES en mélanocytes a permis de produire des cellules pigmentées.



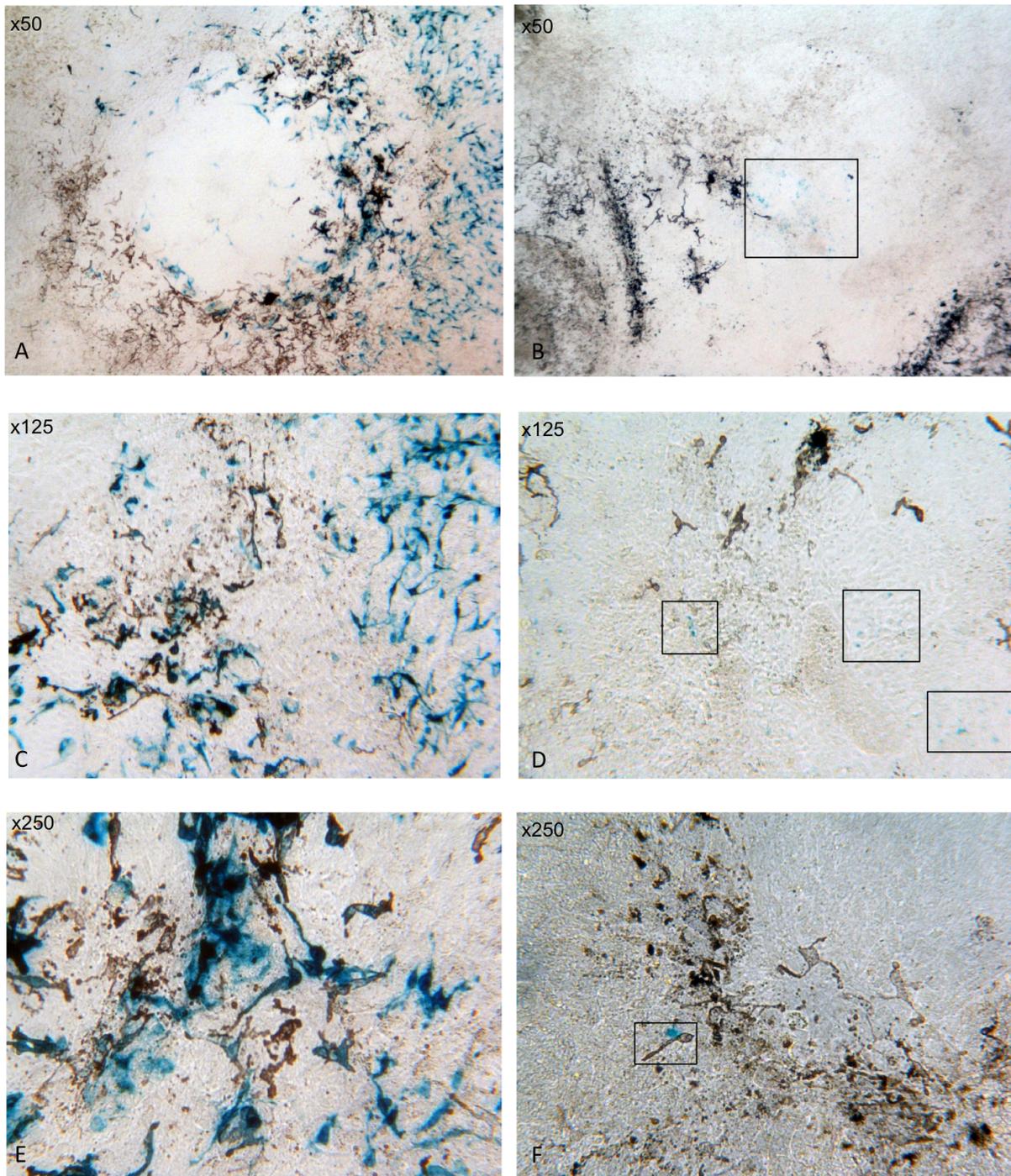
### Étude de l'activité $\beta$ -galactosidase des cellules *Dct-LacZ* et *Dct<sup>Lago1-Néo/+</sup>*

Pour analyser le profil d'expression de *LacZ* et *Lago1* dans les cellules *Dct-LacZ* et *Dct<sup>Lago1-Néo/+</sup>*, les cellules ont été colorées dans une solution de X-gal à différents stades du processus de différenciation. Une activité  $\beta$ -galactosidase a été observée à J12, J25, J38 et J41. A J12, les cellules ES en différenciation formaient des corps embryoïdes. A ce stade, quelques cellules *Dct-LacZ* en différenciation étaient bleues, mais ce n'était pas le cas des cellules *Dct<sup>Lago1-Néo/+</sup>*. A J25, les cellules *Dct-LacZ* marquées par la coloration X-gal, de couleur bleu clair ou bleu foncé, étaient majoritairement non pigmentées mais situées à proximité de cellules pigmentées. Le marquage X-gal était cytoplasmique. Certaines cellules marquées portaient des dendrites typiques des mélanocytes (Figure 42A). A ce même stade, les premières cellules *Dct<sup>Lago1-Néo/+</sup>* marquées par la coloration X-gal ont été observées. Elles présentaient une couleur bleu clair localisée dans leur noyau. L'observation de l'ensemble des puits a montré que les cellules *Dct<sup>Lago1-Néo/+</sup>* X-gal positives étaient peu nombreuses comparées aux cellules *Dct-LacZ* X-gal positives (Figure 42B).



**Figure 42 : Coloration X-gal au stade J25 du processus de différenciation**

- A-** Photo de cellules *Dct-LacZ* (grossissement  $\times 125$ ) colorées au X-gal 25 jours après étalement des cellules ES sur le tapis de cellules ST2. Les cellules marquées par la coloration X-gal sont de couleur bleu clair à bleu foncé. La majorité des cellules colorées ne sont pas pigmentées, mais se trouvent à proximité des cellules pigmentées. Pour un nombre donné de cellules pigmentées, on observe au moins le même nombre de cellules bleues.
- B-** Photo de cellules *Dct<sup>Lago1-Néo/+</sup>* (grossissement  $\times 125$ ) colorées au X-gal 25 jours après étalement des cellules ES sur le tapis de cellules ST2. Les cellules marquées par la coloration X-gal sont de couleur bleu clair. Les cellules colorées ne sont pas pigmentées, mais se trouvent à proximité des cellules pigmentées. Pour un nombre donné de cellules pigmentées, on observe un nombre inférieur de cellules bleues.



**Figure 43 : Coloration X-gal au stade J41 du processus de différenciation**

**A, C et E** : Photos de cellules *Dct-LacZ* (grossissement  $\times 50$  ; 125 et 250) colorées au X-gal 41 jours après étalement des cellules ES sur le tapis de cellules ST2. Les cellules marquées par la coloration X-gal sont de couleur bleu clair à bleu foncé. La majorité des cellules colorées ne sont pas pigmentées, mais se trouvent à proximité des cellules pigmentées. Pour un nombre donné de cellules pigmentées, on observe au moins le même nombre de cellules bleues.

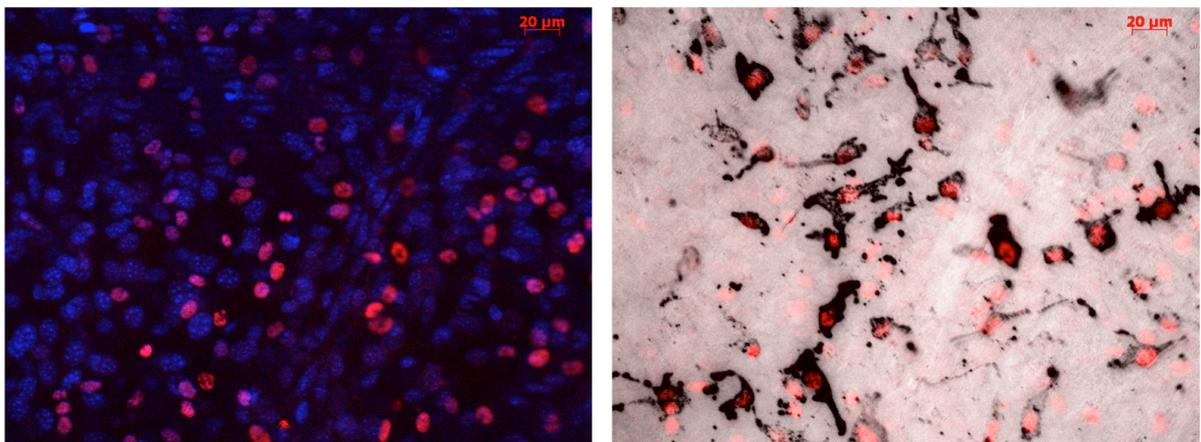
**B, D et F** : Photos de cellules *Dct<sup>Lago1-Néo<sup>+/+</sup></sup>* (grossissement  $\times 50$  ; 125 et 250) colorées au X-gal 41 jours après étalement des cellules ES sur le tapis de cellules ST2. Les cellules marquées par la coloration X-gal sont de couleur bleu clair. Les cellules colorées ne sont pas pigmentées, mais se trouvent à proximité des cellules pigmentées. Pour un nombre donné de cellules pigmentées, on observe un nombre inférieur de cellules bleues.

A J41, la coloration X-gal était présente dans des cellules *Dct-LacZ* pigmentées et non pigmentées (Figure 43 **A, C, E**). La majorité des cellules X-gal positives n'était pas pigmentée. A ce même stade, les cellules *Dct<sup>Lago1-Néo/+</sup>* avaient une activité  $\beta$ -galactosidase plus faible : les cellules bleues étaient de couleur bleu clair et étaient assez difficiles à visualiser. Le nombre de cellules marquées était également plus faible. La coloration X-gal n'a pas révélé de cellule *Dct<sup>Lago1-Néo/+</sup>* pigmentée et bleue (Figure 43 **B, D, F**). Dans les cultures de cellules *Dct-LacZ* et *Dct<sup>Lago1-Néo/+</sup>*, la différenciation des cellules ES en mélanocytes s'est accompagnée d'une expression de la  $\beta$ -galactosidase. L'activité  $\beta$ -galactosidase présentait cependant un profil différent selon le génotype de la cellule.

#### *Étude de l'expression de mCherry dans les cellules *Dct<sup>H2BmCherry-Néo/+</sup>**

Pour étudier le profil d'expression de mCherry dans les cellules *Dct<sup>H2BmCherry-Néo/+</sup>* en culture, les cellules ont été observées en microscopie à fluorescence à différents stades du processus de différenciation. Comme témoin positif, des cellules portant le gène rapporteur H2B-RFP intégré dans leur génome ont été utilisées. Ces cellules expriment effectivement la protéine fluorescente dsRed (RFP, *red fluorescent protein*). A partir de J28, les cellules *Dct<sup>H2BmCherry-Néo/+</sup>* ont été observées au microscope à fluorescence. Quel que soit le stade observé, les cellules n'exprimaient pas de fluorescence rouge détectable en microscopie à fluorescence. A ces stades, des cellules pigmentées étaient visibles, preuve que la différenciation était en cours. L'expression du gène rapporteur *mCherry* sous le contrôle du promoteur *Dct* n'a donc pas été observée au microscope à fluorescence dans les cellules en culture.

Pour tester l'expression de la protéine mCherry dans les cellules *Dct<sup>H2BmCherry-Néo/+</sup>*, nous avons effectué une expérience d'immunocytofluorescence sur ces cellules au stade J38. A J0, des cellules ES *Dct<sup>H2BmCherry-Néo/+</sup>* ont été étalées sur un tapis de cellules ST2 dans des chambres microscopiques, à raison de 42, 62 et 83 cellules par puits. Ces cellules ont été cultivées dans le milieu de différenciation, changé tous les trois jours. En parallèle, des cellules H2B-RFP ont été cultivées et fixées et l'activité RFP a été révélée par immunofluorescence avec un anticorps anti-DsRed. Par immunofluorescence, l'expression de la protéine RFP était visible dans toutes les cellules H2B-RFP. Trente huit jours après l'étalement des cellules ES sur le tapis de cellules stromales ST2, les cellules *Dct<sup>H2BmCherry-Néo/+</sup>* ont été fixées et la présence de la protéine mCherry a été recherchée par immunofluorescence avec le même anticorps anti-DsRed. Un marquage nucléaire de cellules



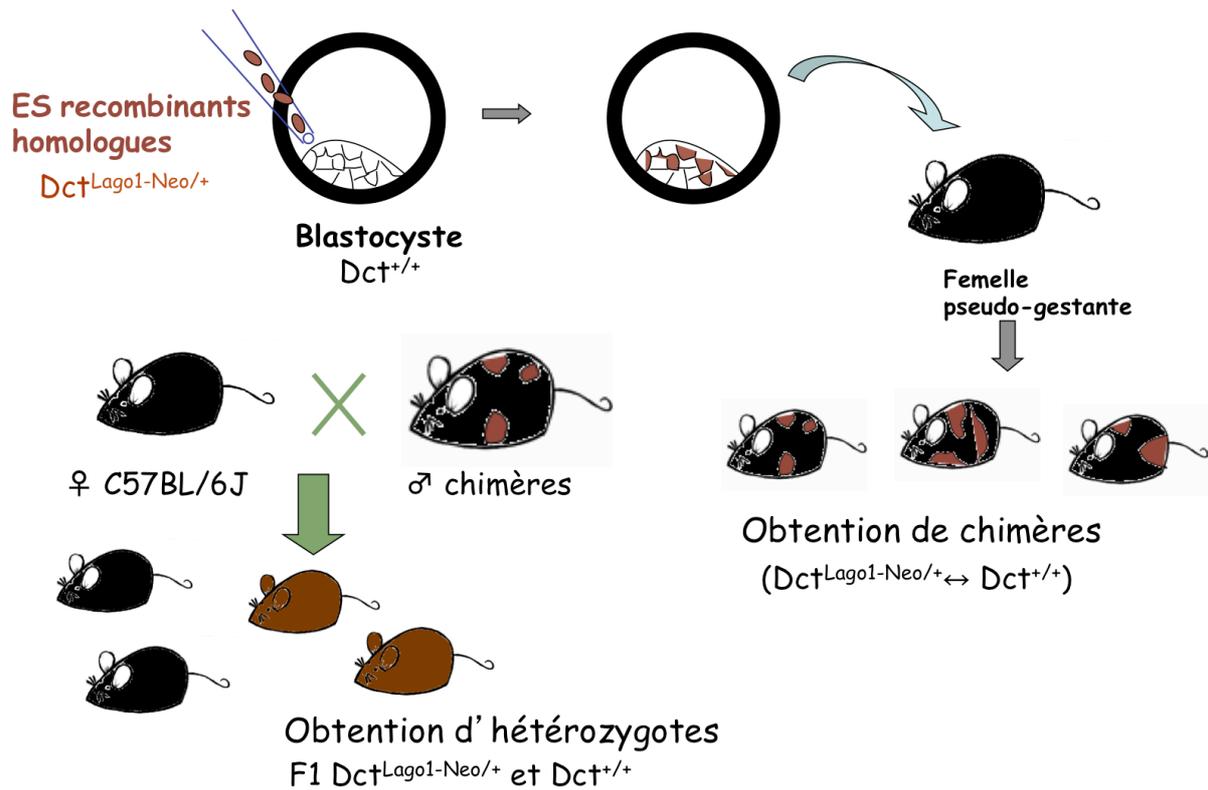
**Figure 44 : Révélation de l'expression de mCherry par immunofluorescence**

A gauche et à droite sont représentées deux vues d'une même photo de cellules *Dct*<sup>H2BmCherry-Néo/+</sup> marquées par un anticorps anti-DsRed. Sur l'image de gauche, tous les noyaux des cellules possèdent un marquage au Hoescht bleu. Les cellules dont le noyau est rouge sont marquées par l'anticorps anti-DsRed qui cible la protéine mCherry. L'image de droite représente une vue en champ clair. La couleur noire correspond aux pigments de mélanine présents dans les cellules pigmentées et les cellules dont le noyau est rouge représentent les cellules qui expriment mCherry, ciblé par l'anticorps anti-DsRed.

*Dct<sup>H2BmCherry-Néo/+</sup>* a été observé. Les cellules marquées par l'anticorps anti-DsRed correspondaient à une partie des cellules fixées. Les cellules pigmentées exprimaient le plus souvent la protéine mCherry, cependant toutes les cellules mCherry positives n'étaient pas pigmentées (Figure 44). L'expression de la protéine fluorescente mCherry dans les cellules *Dct<sup>H2BmCherry-Néo/+</sup>* était donc visible par immunofluorescence dans des cellules pigmentées et non pigmentées.

#### 1.4 Conclusion

Grâce à la technique de différenciation des cellules ES en mélanocytes, des cellules pigmentées ont été générées quel que soit le génotype des cellules ES de départ. Dans le cas des cellules *Dct-LacZ*, le rapporteur *LacZ* était exprimé à partir de J12. Une fois les premières cellules pigmentées apparues, l'activité  $\beta$ -galactosidase a été observée dans des cellules pigmentées et non pigmentées. Dans le cas des cellules *Dct<sup>Lago1-Néo/+</sup>*, le rapporteur *Lago1* était exprimé à partir de J25. Son expression a été observée exclusivement (ou presque) dans des cellules non pigmentées. De plus, les cellules *Dct<sup>Lago1-Néo/+</sup>* marquées par la coloration X-gal étaient moins nombreuses que les cellules *Dct-LacZ* marquées, quel que soit le stade de différenciation observé. Dans le cas des cellules *Dct<sup>H2BmCherry-Néo/+</sup>*, la fluorescence de la protéine mCherry n'était pas observable en culture. Cependant, l'activité mCherry a été révélée par immunofluorescence à J38 dans des cellules pigmentées et non pigmentées.



**Figure 45 : Principe d'obtention des souris hétérozygotes  $Dct^{Lago1-Neo/+}$**

Les cellules ES  $Dct^{Lago1-Neo/+}$  sont insérées dans un blastocyste C57BL/6J, après quoi ce dernier est transféré dans une femelle C57BL/6J pseudo-gestante. Les mâles chimères ainsi obtenus sont croisés avec des femelles C57BL/6J. Les souris agouti issues de ce croisement sont génotypées au locus *Dct*.

## 2 Etude du phénotype des souris $Dct^{Lago1-Néo/Lago1-Néo}$ et $Dct^{Lago1/+}$ : KO et KI de $Dct$

### 2.1 Objectif et stratégie

A partir des cellules ES  $Dct^{Lago1-Néo/+}$ , des chimères ( $Dct^{Lago1-Néo/+} \leftrightarrow Dct^{+/+}$ ) puis une lignée de souris  $Dct^{Lago1-Néo/+}$  ont été générées. Les souris hétérozygotes ont été croisées entre elles pour obtenir des souris homozygotes  $Dct^{Lago1-Néo/Lago1-Néo}$ . En parallèle, des mâles chimères ( $Dct^{Lago1-Néo/+} \leftrightarrow Dct^{+/+}$ ) ont été croisés avec des femelles  $PGK-Cre$  pour produire une lignée de souris  $Dct^{Lago1/+}$ . Le promoteur du gène  $Dct$  a été utilisé dans plusieurs expériences de transgénèse (Chapitre 2, section 3 de l'introduction). Chez ces souris transgéniques, le rapporteur est exprimé dans les mélanocytes et leurs précurseurs. Dans le cas des souris  $Dct^{Lago1/+}$ , l'expression du rapporteur  $Lago1$  était donc attendu chez les mélanoblastes/cytes. Comme témoin, le profil du rapporteur  $LacZ$  chez les souris  $Dct-LacZ$  a été comparé à celui du rapporteur  $Lago1$  chez les souris  $Dct^{Lago1/+}$ .

### 2.2 Matériels et méthodes

#### *Génération de la lignée de souris $Dct^{Lago1-Néo/+}$*

Ce travail a été réalisé en collaboration avec Francina Langa Vives (Centre d'Ingénierie Génétique Murine, Institut Pasteur). Des cellules ES  $Dct^{Lago1-Néo/+}$  produites dans un fond génétique 129/Sv ont été introduites dans des blastocystes de souris C57BL/6J, après quoi ces derniers ont été transférés dans des femelles C57BL/6J pseudo-gestantes. Les souris chimères issues de ce croisement étaient de génotype ( $129/Sv Dct^{Lago1-Néo/+} \leftrightarrow C57BL/6J Dct^{+/+}$ ). Des mâles chimères ont été croisées avec des femelles C57BL/6J  $Dct^{+/+}$ . Les souris agouti issues de ce croisement ont été génotypées au locus  $Dct$  et nous avons conservé les animaux ( $129/Sv \times C57BL/6J$ ) F1 de génotype  $Dct^{Lago1-Néo/+}$  pour les expériences ultérieures (Figure 45).



### *Obtention de souris homozygotes $Dct^{Lago1-Néo/Lago1-Néo}$*

Les individus hétérozygotes (129/Sv× C57BL/6J)F1  $Dct^{Lago1-Néo/+}$  descendants de mâles chimères ont été croisés entre eux. La couleur du pelage des souris homozygotes  $Dct^{Lago1-Néo/Lago1-Néo}$  dans un fonds génétique mixte (129/Sv× C57BL/6J) a été comparée à celle des souris sauvages  $Dct^{+/+}$ .

### *Génération de la lignée de souris $Dct^{Lago1/+}$*

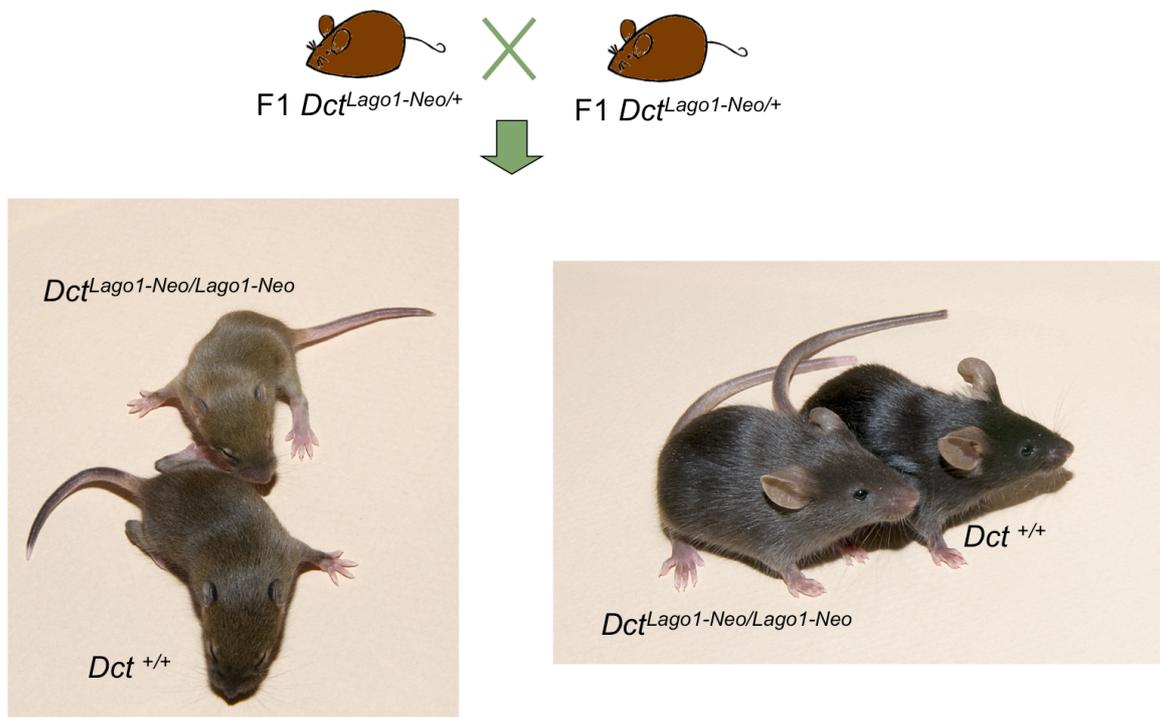
Des mâles chimères ( $Dct^{Lago1-Néo/+} \leftrightarrow Dct^{+/+}$ ) ont été croisés avec des femelles transgéniques *PGK-Cre* dans un fonds génétique 129/Sv. Ces femelles expriment la recombinaison Cre durant l'oogenèse (Lallemand, *et al.*, 1998). Les souris 129/Sv  $Dct^{Lago1/+}$  issues de ce croisement ont été conservées pour les expériences ultérieures. Nous avons conservé celles qui portaient une délétion de la cassette *Néo* au locus *Dct* et ne portaient pas le transgène *PGK-Cre*.

### *Mesure de l'activité $\beta$ -galactosidase chez les embryons $Dct^{Lago1/+}$*

Un mâle 129/Sv  $Dct^{Lago1/+}$  a été croisé avec des femelles (C57BL/6×DBA2). Des embryons de différents stades embryonnaires ont été obtenus. Les embryons ont été fixés dans une solution composée de PFA 4%, 0,1M tampon phosphate (pH= 7,3), 2mM  $MgCl_2$ ,  $H_2O$ . La fixation a été réalisée à 4°C pendant 30 minutes. Après fixation, les embryons ont été lavés 3 fois 10 minutes dans une solution 0,1M tampon phosphate (pH= 7,3), 2mM  $MgCl_2$ , 0,01% DOC, 0,02% NP40,  $H_2O$ . Puis, les embryons ont été colorés dans une solution composée de 5 mM ferricyanure, 5 mM ferrocyanure, 0,1M tampon phosphate, 0,01% DOC, 0,02% NP40, 2 mM  $MgCl_2$ ,  $H_2O$  et 10  $\mu g/mL$  de X-gal pendant une nuit à 32°C. Des embryons *Dct-LacZ* de mêmes stades ont été colorés avec le même protocole.

### *Mesure de l'activité $\beta$ -galactosidase chez les nouveau-nés et les adultes $Dct^{Lago1/+}$*

L'activité  $\beta$ -galactosidase a été observée sur la peau de souris nouveau-nées de 3 jours. Des biopsies de peau de souris  $Dct^{Lago1/+}$  de 3 jours ont été prélevées, fixées en PFA 4%, PBS 1X 20 minutes, rincées en PBS et colorées dans une solution X-gal (cf. ci-dessus pour la composition) pendant une nuit à 37°C. Des fragments d'oreille de souris adultes ont également été fixés en PFA 4% et colorés en solution X-gal à 37°C pendant une nuit. L'activité  $\beta$ -galactosidase a été observée sur la peau d'individus  $Dct^{Lago1/+}$  et *Dct-LacZ* âgés de 2 mois. Les biopsies de peau ont été fixées 45 minutes en PFA 4% tampon phosphate à température ambiante (cf. ci-dessus pour la composition). Après rinçage en PBS, les biopsies



**Figure 46 : Comparaison de souris homozygotes  $Dct^{Lago1-Néo/Lago1-Néo}$  et de souris sauvages  $Dct^{+/+}$**

Les souris  $Dct^{Lago1-Néo/Lago1-Néo}$  ont été obtenues par croisement de souris  $Dct^{Lago1-Néo/+}$  entre elles. Sur la photo de gauche, on observe deux jeunes souris de pelage agouti et de génotypes  $Dct^{Lago1-Néo/Lago1-Néo}$  et  $Dct^{+/+}$ . Sur la photo de droite, on observe deux souris adultes de pelage non agouti et de génotypes  $Dct^{Lago1-Néo/Lago1-Néo}$  et  $Dct^{+/+}$ . Les souris  $Dct^{Lago1-Néo/Lago1-Néo}$  ont, dans les deux cas, un pelage de couleur plus claire que les souris sauvages.

ont été placées dans une solution de sucrose 15% en PBS 1X à 4°C pendant une nuit. Le lendemain, les biopsies ont été incluses dans une solution de sucrose 15%, gélatine 7%. Après inclusion, les biopsies ont été congelées à -20°C, puis coupées au cryostat. Les coupes histologiques congelées ont été post-fixées en PFA 2% tampon phosphate pendant 10 minutes, rincées en PBS, puis colorées avec une solution X-gal à 32°C pendant une nuit.

## 2.3 Résultats

### *Phénotype des souris $Dct^{Lago1-Néo/Lago1-Néo}$*

Des souris homozygotes  $Dct^{Lago1-Néo/Lago1-Néo}$  ont été produites et leur phénotype a été comparé à celui de leurs frères et sœurs de portée de génotypes  $Dct^{Lago1-Néo/+}$  et  $Dct^{+/+}$ . Les souris  $Dct^{Lago1-Néo/Lago1-Néo}$  ont été obtenues en croisant entre eux des hybrides (129/Sv× C57BL/6J) F1 hétérozygotes  $Dct^{Lago1-Néo/+}$ . Les descendants hétérozygotes nés de ce croisement ne pouvaient être distingués des souris sauvages par la couleur de leur pelage. Les homozygotes  $Dct^{Lago1-Néo/Lago1-Néo}$ , au pelage agouti ou non agouti, avaient un pelage de couleur plus claire que celui des individus  $Dct^{+/+}$  et  $Dct^{Lago1-Néo/+}$  (Figure 46). Les homozygotes  $Dct^{Lago1-Néo/Lago1-Néo}$  n'ont aucun autre phénotype évident que celui qui concerne la couleur de leur pelage. Ces résultats sont en accord avec le phénotype rapporté des souris  $Dct^{Cre/Cre}$  qui portent le gène codant la recombinaison Cre inséré au locus *Dct* (Guyonneau, *et al.*, 2004). Les conséquences phénotypiques de l'inactivation du gène *Dct* semblent donc restreintes à la couleur du pelage. Guyonneau et collaborateurs ont également comparé le cerveau, l'iris et le RPE des souris de génotype  $Dct^{Cre/Cre}$  et de génotype sauvage et n'ont observé aucune différence (Guyonneau, *et al.*, 2004).

### *Étude de l'expression du gène *Lago1* chez les embryons $Dct^{Lago1/+}$*

Pour étudier le profil d'expression du rapporteur *Lago1* chez les embryons  $Dct^{Lago1/+}$ , l'activité  $\beta$ -galactosidase a été révélée à différents stades par coloration X-gal. Un mâle 129/Sv  $Dct^{Lago1/+}$  a été croisé avec des femelles (C57BL/6×DBA2). Des embryons ont été prélevés à E9,5 ; E11,5 ; E12,5 et E 15,5. L'activité  $\beta$ -galactosidase a été révélée par coloration d'embryons *in toto* de génotype  $Dct^{Lago1/+}$ . Des embryons *Dct-LacZ* des mêmes stades ont été utilisés comme témoins positifs. Chez les embryons *Dct-LacZ*, l'activité  $\beta$ -galactosidase était visible dans le cerveau, au niveau de l'œil, de cellules en migration à



partir de la crête neurale et dans les ganglions dorsaux. En revanche, la coloration *in toto* des embryons  $Dct^{Lago1/+}$  ne révélait aucune expression du rapporteur *Lago1* visible sur l'ensemble de l'embryon (Figure 47). Cependant, un marquage bleu était visible dans le RPE. Ce marquage discret, très localisé était peu discernable car masqué par la mélanine des cellules du RPE. L'expression de *Lago1* révélée par une coloration X-gal *in toto* de l'embryon semble donc limitée au RPE.

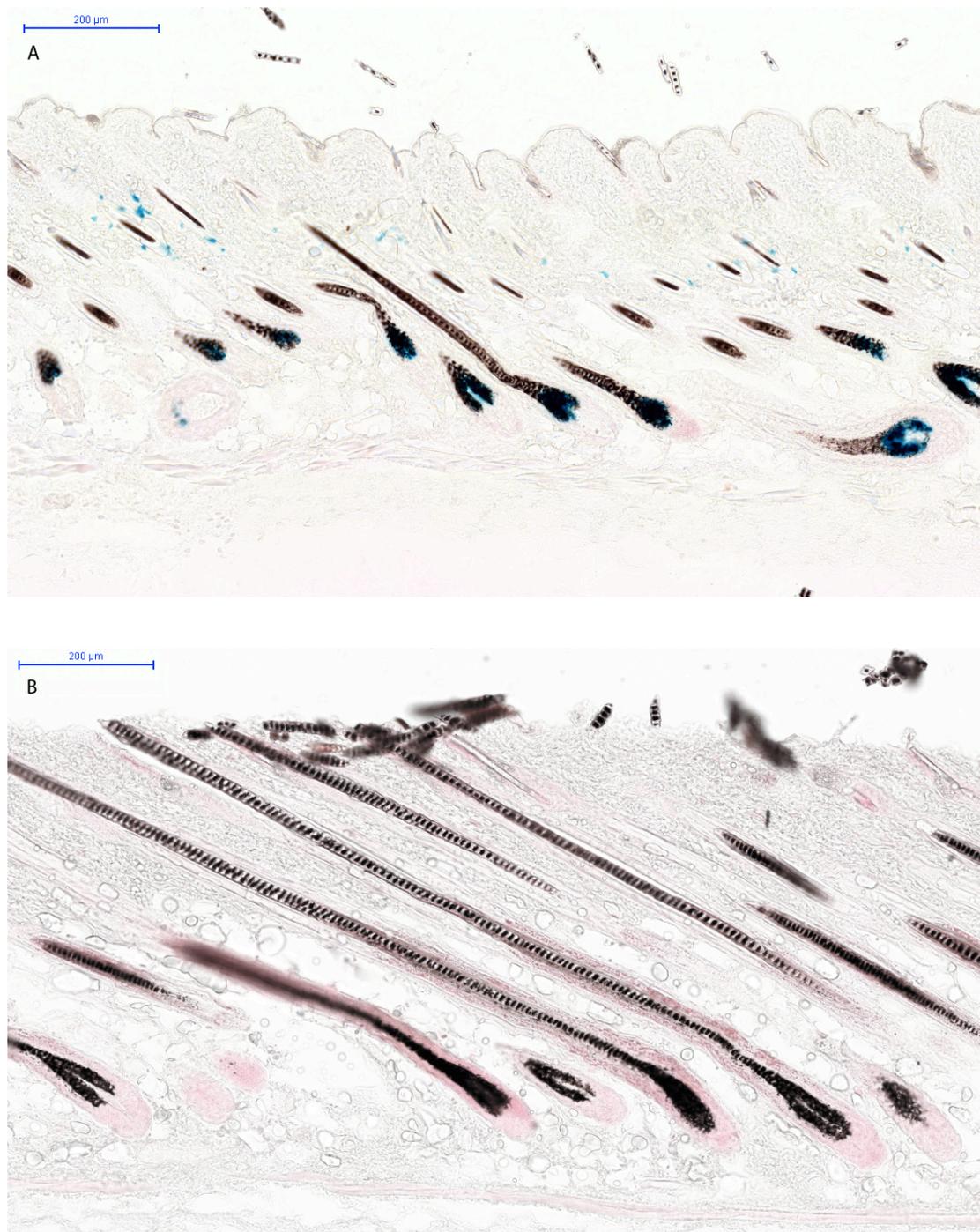


**Figure 47 : Comparaison des profils d'expression de la  $\beta$ -galactosidase chez un embryon  $Dct^{Lago1/+}$  et chez un embryon *Dct-LacZ* de stade E11,5**

A droite est représenté un embryon de stade E11,5 *Dct-LacZ*. Le marquage bleu est observé au niveau du cerveau, de l'œil, de cellules en migration et des ganglions dorsaux. A gauche est représenté un embryon de stade E11,5 de génotype  $Dct^{Lago1/+}$ . Aucun marquage bleu n'est visible sur cette photo. Un marquage bleu est visible au niveau du RPE à fort grossissement, parmi les pigments.

#### *Étude de l'expression de *Lago1* chez les souris $Dct^{Lago1/+}$ nouveau-nées et adultes*

Pour étudier le profil d'expression du rapporteur *Lago1* chez les souris  $Dct^{Lago1/+}$ , l'activité  $\beta$ -galactosidase a été observée sur des biopsies de peau après coloration X-gal. Lors du premier cycle du follicule pileux, des biopsies de peau de nouveau-nés de 3 jours ont été prélevées et colorées au X-gal. Ces nouveau-nés ne présentaient pas de marquage bleu visible dans la peau. Lors du deuxième cycle du follicule pileux, des biopsies de peau de souris  $Dct^{Lago1/+}$  et *Dct-LacZ* âgées de deux mois ont été prélevées. Des coupes histologiques de ces peaux ont été réalisées au cryostat, puis post-colorées. Chez les souris *Dct-LacZ*, l'activité  $\beta$ -galactosidase a été observée dans le bulbe des follicules pileux et au niveau du bulge.



**Figure 48 : Comparaison des profils d'expression de la  $\beta$ -galactosidase sur des coupes de peaux de souris *Dct-LacZ* et *Dct<sup>Lago1</sup>+/+* de 34 jours**

- A-** Coupe de peau de souris *Dct-LacZ* (10  $\mu$ m) colorées au X-gal. La coloration bleue est observée dans tous les bulbes des follicules pileux. Quelques cellules bleues sont également observées dans la partie supérieure des follicules pileux. Ces dernières sont identifiées comme les cellules souches de mélanocytes situées dans le bulge du follicule pileux.
- B-** Coupe de peau de souris *Dct<sup>Lago1-Néo</sup>+/+* (10  $\mu$ m) colorées au X-gal. On n'observe pas de coloration bleue.

Aucun marquage bleu n'a été observé sur la peau des souris  $Dct^{Lago1/+}$  (Figure 48). Par ailleurs, des biopsies d'oreille de souris  $Dct^{Lago1/+}$  adultes ont été colorés en solution X-gal. L'observation au microscope de ces biopsies a révélé la présence d'un marquage X-gal sous forme de points bleus, visibles entre les poils. Il n'a pas été possible d'identifier le type cellulaire responsable de ce marquage. Les souris  $Dct^{Lago1/+}$  ne présentent donc pas d'activité  $\beta$ -galactosidase dans la peau des nouveau-nés et des adultes. Un marquage bleu est cependant observable au niveau des oreilles des souris adultes, sans que le type cellulaire soit connu.

## 2.4 Conclusion

Les souris homozygotes  $Dct^{Lago1-Néo/Lago1-Néo}$  sont déficientes pour le gène  $Dct$ . Les conséquences de l'inactivation du gène  $Dct$  sur le phénotype semblent restreintes à la couleur du pelage, plus clair que celui de souris de génotype sauvage. Ces résultats sont en accord avec les résultats publiés sur la souris  $Dct^{Cre/Cre}$ . Une lignée de souris  $Dct^{Lago1/+}$  dans le fond génétique 129/Sv a également été produite. Chez l'embryon, l'expression du rapporteur  $Lago1$  a été détectée seulement au niveau du RPE. Chez le nouveau-né et l'adulte, l'expression de  $Lago1$  n'a pas été détectée dans la peau, où s'insèrent les follicules pileux. L'expression de  $Lago1$  a été détectée uniquement dans les oreilles des souris  $Dct^{Lago1/+}$  adultes.

## 3 Discussion et perspectives

Comme nous l'avons rapporté dans l'introduction, plusieurs souris transgéniques expriment un gène rapporteur sous contrôle du promoteur  $Dct$ . La lignée transgénique la plus communément utilisée est celle qui porte le transgène  $Dct-LacZ$ . Mais des souris BAC  $Dct-LacZ$ , des souris *knock in*  $Dct^{Cre/+}$  et des souris transgéniques inductibles  $iDct-GFP$  ont également été produites. La souris  $Dct-LacZ$  s'est avérée être un bon modèle rapporteur du lignage mélanocytaire. La question de savoir si l'expression de  $LacZ$  dans les ganglions dorsaux de l'embryon rapporte l'expression fidèle du gène  $Dct$  ou si elle est due à un effet de position du transgène est pour l'instant sans réponse. Nous souhaitons y répondre en caractérisant le profil d'expression du gène  $Lago1$  dans la lignée  $Dct^{Lago1/+}$ .

L'étude du patron d'expression des gènes *Lago1* et *H2B-mCherry* intégrés au locus *Dct* nous amène à nous poser deux questions: (1) le promoteur du gène *Dct* est-il un promoteur trop faible pour contrôler l'expression d'un transgène ? (2) la construction utilisée pour insérer les gènes rapporteurs *Lago1* et *H2B-mCherry* est-elle responsable de la faible expression des gènes insérés au locus *Dct* ?

### 3.1 Le promoteur *Dct* est-il un promoteur trop faible pour exprimer des gènes d'intérêt dans les mélanocytes et leurs précurseurs ?

La différenciation des cellules ES *Dct<sup>Lago1-Néo/+</sup>* et *Dct<sup>H2BmCherry-Néo/+</sup>* en mélanocytes a permis d'obtenir des cellules pigmentées. J'ai observé des cellules *Dct-LacZ* qui étaient pigmentées et X-gal-positives. En revanche, je n'ai pas observé de cellules *Dct<sup>Lago1-Néo/+</sup>* qui étaient pigmentées et X-gal-positives. En outre, les premières cellules *Dct-LacZ* X-gal-positives sont apparues bien avant que la mélanine ne soit visible (à J12). Au contraire, les premières cellules *Dct<sup>Lago1-Néo/+</sup>* X-gal positives n'ont été remarquées qu'après l'apparition des grains de mélanine (à J25). Cette observation a été faite sur deux expériences successives. Enfin, de nombreuses cellules X-gal-positives, *Dct<sup>Lago1-Néo/+</sup>* ou *Dct-LacZ*, ne sont pas pigmentées. Ces cellules X-gal-positives sont vraisemblablement des mélanoblastes, puisque elles sont situées à proximité de cellules pigmentées. Pour confirmer cette hypothèse, il faudrait effectuer une immunofluorescence avec un anticorps contre une protéine exprimée dans les mélanoblastes, comme par exemple un anticorps anti-DCT ou anti-MITF.

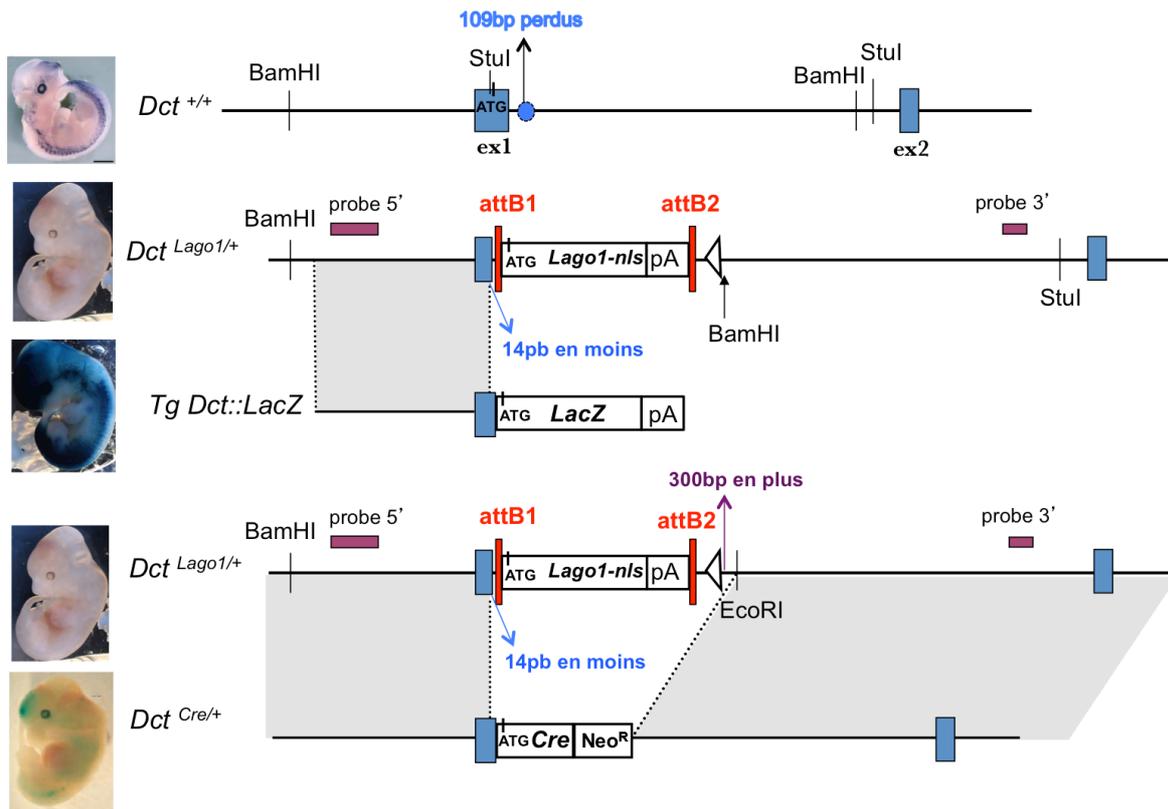
L'apparition tardive des cellules X-gal-positives dans les cellules *Dct<sup>Lago1-Néo/+</sup>* pourrait être due à une très faible activité du promoteur du gène *Dct*. L'activité  $\beta$ -galactosidase n'est détectée que lorsque la quantité de protéines présentes dans le noyau atteint une valeur seuil. Cela pourrait expliquer le faible marquage X-gal observé, de couleur bleu pâle. L'absence de coloration bleue dans les cellules pigmentées pourrait s'expliquer aussi par un niveau d'expression de la  $\beta$ -galactosidase très faible. Le marquage pourrait être masqué par les pigments. Cependant, la coloration bleue devrait alors être visible dans des cellules pigmentées qui contiennent peu de mélanine.

Nous avons noté que l'expression du gène rapporteur *mCherry* n'est pas détectée dans des cellules *Dct<sup>H2BmCherry-Néo/+</sup>* en microscopie à fluorescence. Un anticorps anti-mCherry a cependant révélé que la protéine mCherry est exprimée dans les cultures de cellules ES se

différenciant en mélanocytes. La protéine mCherry a été observée dans des cellules non pigmentées et pigmentées. Il semble que la quantité de protéine mCherry soit trop faible pour que sa fluorescence directe soit visible. En comparaison, la méthode de coloration par le X-gal est très sensible. Il serait néanmoins intéressant d'utiliser un anticorps anti- $\beta$ -galactosidase sur des cellules  $Dct^{Lago1-N\acute{e}o/+}$  pour rechercher la présence de la protéine plutôt que son activité. Nous avons combiné des expériences qui utilisent des anticorps anti- $\beta$ -galactosidase et anti-DCT sur des cultures des cellules ES  $Dct^{Lago1-N\acute{e}o/+}$ , sans succès. Une mise au point des conditions expérimentales de cette double immunofluorescence permettrait d'évaluer la quantité de  $\beta$ -galactosidase dans les mélanoblastes/cytes se différenciant à partir des cellules ES  $Dct^{Lago1-N\acute{e}o/+}$ .

Le profil d'expression du gène *Lago1* chez les souris  $Dct^{Lago1/+}$  nous amène également à nous demander si le promoteur *Dct* n'est pas trop faible pour exprimer un gène rapporteur. Contrairement aux embryons *Dct-LacZ* colorés *in toto* qui présentent un marquage bleu très fort, le seul marquage X-gal observé chez les embryons  $Dct^{Lago1/+}$  colorés *in toto* est localisé dans le RPE. Si le marquage X-gal chez les embryons  $Dct^{Lago1/+}$  est aussi faible que celui qui est observé dans les cellules ES en différenciation de même génotype, il est possible que le marquage X-gal ne soit pas visible sur l'embryon *in toto*, parce que trop faible. Pour préciser le profil d'expression du gène *Lago1*, il serait utile d'éclaircir les embryons  $Dct^{Lago1/+}$  au glycérol, ce qui rendrait les tissus « translucides ». Nous pourrions peut-être ainsi déterminer s'il s'agit d'un problème d'intensité du marquage.

Nous avons réalisé des coupes de peau de souris adultes *Dct-LacZ* et  $Dct^{Lago1/+}$  puis nous les avons post-colorées. Les follicules pileux de souris adultes sont très pigmentés. Malgré la pigmentation intense dans le bulbe, l'activité  $\beta$ -galactosidase est très forte dans les follicules pileux des souris *Dct-LacZ*. Chez les souris  $Dct^{Lago1/+}$ , le marquage X-gal n'est pas visible. Il est peut-être masqué par la pigmentation. Là encore, il semble que l'expression du gène *Lago1*, sous le contrôle du promoteur *Dct*, soit trop faible pour être détectée. La peau de l'oreille est beaucoup plus fine et la pigmentation y est moins intense. Nous avons observé un marquage X-gal au niveau de la peau de l'oreille, sans que nous ayons pu identifier le type cellulaire marqué. Une immunofluorescence avec un anticorps anti-DCT sur des coupes d'oreille, pourrait permettre de savoir si des mélanoblastes/cytes sont marqués. Pour contourner le problème de l'intense pigmentation des bulbes dans les follicules pileux, nous pourrions analyser le profil d'expression de l'allèle  $Dct^{Lago1}$  dans un fond génétique albinos



**Figure 49 : Comparaison des constructions des souris *knock in Dct*<sup>Lago1/+</sup>, *Dct*<sup>Cre/+</sup> et de la souris transgénique *Dct-LacZ***

De haut en bas sont représentées l'allèle *Dct* endogène, l'allèle *Dct*<sup>Lago1/+</sup>, le transgène *Dct-LacZ*, de nouveau l'allèle *Dct*<sup>Lago1/+</sup> et l'allèle *Dct*<sup>Cre/+</sup>. A gauche sont représentés les photos d'embryons de stade E11,5 colorés en X-gal en fonction de la construction associée. Pour l'allèle endogène, il s'agit de la photo de l'hybridation *in situ*. Les rectangles bleus figurent les exons du gène *Dct*. Les séquences absentes dans l'allèle *Dct*<sup>Lago1/+</sup> sont indiquées en bleu. La séquence de 300 pb supplémentaire chez le *knock in Dct*<sup>Lago1/+</sup> par rapport au *knock in Dct*<sup>Cre/+</sup> est indiquée en violet.

(*Tyr<sup>c/c</sup>*). Nous pourrions établir le rôle de la pigmentation dans notre difficulté à visualiser le marquage X-gal.

### 3.2 La construction utilisée est-elle responsable de la faible expression des gènes insérés au locus *Dct* ?

Nous avons choisi d'intégrer le gène d'intérêt à la place du premier exon du gène *Dct*. La lignée *knock in Dct<sup>Cre/+</sup>* a été produite en utilisant un vecteur de recombinaison homologue très similaire (Guyonneau, *et al.*, 2004). Le profil d'expression du gène *Cre* inséré au locus *Dct* a été étudié en croisant des individus *Dct<sup>Cre/+</sup>* et R26R (Introduction, Chapitre 2, section 3.2). Les embryons *Dct<sup>Cre/+</sup>* ; R26R présentent un marquage X-gal différent des souris *Dct-LacZ* : l'activité  $\beta$ -galactosidase est visible au niveau du télencéphale, du RPE et faiblement dans des ganglions dorsaux (Figure 36). Notons que les protéines rapportrices exprimées chez les individus *Dct<sup>Cre/+</sup>* ; R26R et *Dct<sup>Lago1/+</sup>* ne sont pas localisées dans le même compartiment cellulaire. La  $\beta$ -galactosidase est exprimée dans le cytoplasme des cellules *Dct<sup>Cre/+</sup>* ; R26R, tandis qu'elle est exprimée dans le noyau des cellules *Dct<sup>Lago1/+</sup>*. Cette différence pourrait expliquer pourquoi on observe un léger marquage X-gal chez les embryons *Dct<sup>Cre/+</sup>* ; R26R et pas chez les embryons *Dct<sup>Lago1/+</sup>*. Dans ces deux cas, l'expression des gènes rapporteurs ne reflètent pas l'expression du gène *Dct* endogène, telle qu'elle peut être observée par hybridation *in situ* (Figure 34).

Comparons les trois constructions réalisées au locus *Dct* : *Dct<sup>Lago1/+</sup>*, *Dct<sup>Cre/+</sup>* et *Dct-LacZ* (Figure 49). Le promoteur utilisé pour exprimer le transgène est identique, à 14 pb près. Ces 14 pb sont présents dans les constructions *Dct<sup>Cre/+</sup>* et *Dct-LacZ*, mais pas dans la construction *Dct<sup>Lago1/+</sup>* (Figure 49). Les souris *Dct<sup>Cre/+</sup>* et *Dct-LacZ* portent toutes les deux ces 14 pb ; le profil d'expression du rapporteur est cependant différents. Il semble donc peu probable que l'absence de ces 14 pb explique le profil du gène *Lago1* chez les souris *Dct<sup>Lago1/+</sup>*. Cent neuf paires de bases de l'intron 1 du gène *Dct* ont été supprimés lors de la construction du vecteur de recombinaison homologue qui a permis d'obtenir l'allèle *Dct<sup>Lago1</sup>* (Figure 49). Une comparaison des constructions à l'origine des allèles *Dct<sup>Cre</sup>* et *Dct<sup>Lago1</sup>* montre qu'il manque 300 pb (en plus des 109 pb) de l'intron 1 chez la souris *Dct<sup>Cre/+</sup>* (Figure 49). Nous ne pouvons pas exclure l'hypothèse selon laquelle la séquence intronique de 109 pb contient un élément de régulation essentiel à l'expression correcte du gène *Dct*. Cependant, la



souris transgénique *Dct-LacZ* qui ne porte pas ces 109 pb exprime le rapporteur *LacZ* dans le lignage mélanocytaire. En conclusion, il est peu probable que ces 109 pb soient à l'origine du profil d'expression des gènes *Lagol* et *Cre* chez les individus *Dct<sup>Lagol/+</sup>* et *Dct<sup>Cre/+</sup>*.

Les embryons *Dct<sup>Cre/+</sup>*; R26R expriment une activité  $\beta$ -galactosidase faible mais détectable, contrairement aux embryons *Dct<sup>Lagol/+</sup>*. Chez les embryons *Dct<sup>Cre/+</sup>*; R26R, le marquage X-gal résulte de la combinaison des expressions du gène *Cre*, sous contrôle du promoteur *Dct*, et du gène *LacZ* au locus *ROSA*. La combinaison des expressions ne reflète pas exactement l'expression du gène *Dct* endogène. Notre hypothèse est que les gènes *Cre* et *Lagol* insérés au locus *Dct* sont exprimés très faiblement. Dans le cas des embryons *Dct<sup>Cre/+</sup>*; R26R, cette faible expression suffit pour provoquer l'excision de la cassette STOP et permettre une expression du gène rapporteur *LacZ* au locus *ROSA*, de sorte que le marquage soit visible. Par comparaison, une expression faible du rapporteur *Lagol* ne suffit pas pour détecter le marquage nucléaire dans l'embryon entier. En conclusion, il est difficile de comparer les deux modèles, mais dans chacun des cas, les gènes rapporteurs sont exprimés faiblement et leur expression ne reflète pas le profil du gène *Dct* endogène, révélé par hybridation *in situ*.

Les souris transgéniques qui possèdent un transgène sous contrôle du promoteur *Dct*, comme la souris *iDct-GFP* ou le BAC *Dct-LacZ*, ont donné des résultats comparables à la souris *Dct-LacZ*. Ce promoteur donne donc de bons résultats en transgénèse, technique qui s'accompagne généralement d'une insertion de multi-copies du transgène. Par comparaison, les embryons et adultes *Dct<sup>Cre/+</sup>* et *Dct<sup>Lagol/+</sup>* présentent une expression très faible des gènes rapporteurs. Notre conclusion est que le promoteur *Dct* n'est pas suffisamment fort pour être utilisé en *knock in* et permettre l'expression d'un gène d'intérêt dans le lignage mélanocytaire.



## REFERENCES BIBLIOGRAPHIQUES

- Adameyko, I., Lallemand, F., Aquino, J.B., Pereira, J.A., Topilko, P., Muller, T., Fritz, N., Beljajeva, A., Mochii, M., Liste, I., Usoskin, D., Suter, U., Birchmeier, C., and Ernfors, P.** (2009). Schwann cell precursors from nerve innervation are a cellular origin of melanocytes in skin. *Cell* 139, 366-379.
- Adameyko, I., and Lallemand, F.** (2010). Glial versus melanocyte cell fate choice: Schwann cell precursors as a cellular origin of melanocytes. *Cell Mol Life Sci* 67, 3037-3055.
- Aksan, I., and Goding, C.R.** (1998). Targeting the microphthalmia basic helix-loop-helix-leucine zipper transcription factor to a subset of E-box elements in vitro and in vivo. *Mol Cell Biol* 18, 6930-6938.
- Arnould, S., Chames, P., Perez, C., Lacroix, E., Duclert, A., Epinat, J.C., Stricher, F., Petit, A.S., Patin, A., Guillier, S., Rolland, S., Prieto, J., Blanco, F.J., Bravo, J., Montoya, G., Serrano, L., Duchateau, P., and Paques, F.** (2006). Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets. *J Mol Biol* 355, 443-458.
- Arnould, S., Perez, C., Cabaniols, J.P., Smith, J., Gouble, A., Grizot, S., Epinat, J.C., Duclert, A., Duchateau, P., and Paques, F.** (2007). Engineered I-CreI derivatives cleaving sequences from the human XPC gene can induce highly efficient gene correction in mammalian cells. *J Mol Biol* 371, 49-65.
- Arnould, S., Delenda, C., Grizot, S., Desseaux, C., Paques, F., Silva, G.H., and Smith, J.** (2011). The I-CreI meganuclease and its engineered derivatives: applications from cell modification to gene therapy. *Protein Eng Des Sel* 24, 27-31.
- Aubin-Houzelstein, G., Bernex, F., Elbaz, C., and Panthier, J.J.** (1998). Survival of patchwork melanoblasts is dependent upon their number in the hair follicle at the end of embryogenesis. *Dev Biol* 198, 266-276.
- Aubin-Houzelstein, G., Djian-Zaouche, J., Bernex, F., Gadin, S., Delmas, V., Larue, L., and Panthier, J.J.** (2008). Melanoblasts' proper location and timed differentiation depend on Notch/RBP-J signaling in postnatal hair follicles. *J Invest Dermatol* 128, 2686-2695.
- Babinet, C., and Cohen-Tannoudji, M.** (2000). Vingt ans d'interventions délibérées sur le génome de la souris. *Med Sci (Paris)* 16, 31-42.
- Babinet, C.** (2003). La transgénèse chez la souris. In *De la transgénèse animale à la biothérapie chez l'homme*, Académie des sciences, ed. (Paris, Tec&Doc), pp. 2-26.
- Barnes, D.E., Stamp, G., Rosewell, I., Denzel, A., and Lindahl, T.** (1998). Targeted disruption of the gene encoding DNA ligase IV leads to lethality in embryonic mice. *Curr Biol* 8, 1395-1398.
- Baxter, L.L., and Pavan, W.J.** (2002). The oculocutaneous albinism type IV gene *Matp* is a new marker of pigment cell precursors during mouse embryonic development. *Mech Dev* 116, 209-212.
- Belfort, M., and Roberts, R.J.** (1997). Homing endonucleases: keeping the house in order. *Nucleic Acids Res* 25, 3379-3388.
- Bentley, N.J., Eisen, T., and Goding, C.R.** (1994). Melanocyte-specific expression of the human tyrosinase promoter: activation by the microphthalmia gene product and role of the initiator. *Mol Cell Biol* 14, 7996-8006.
- Berger, M.F., Levin, J.Z., Vijayendran, K., Sivachenko, A., Adiconis, X., Maguire, J., Johnson, L.A., Robinson, J., Verhaak, R.G., Sougnez, C., Onofrio, R.C., Ziaugra, L., Cibulskis, K., Laine, E., Barretina, J., Winckler, W., Fisher, D.E., Getz, G., Meyerson, M., Jaffe, D.B., Gabriel, S.B., Lander, E.S., Dummer, R., Gnirke, A., Nusbaum, C., and Garraway, L.A.** (2010). Integrative analysis of the melanoma transcriptome. *Genome Res* 20, 413-427.

- Bermingham-McDonogh, O., Xu, Y.T., Marchionni, M.A., and Scherer, S.S.** (1997). Neuregulin expression in PNS neurons: isoforms and regulation by target interactions. *Mol Cell Neurosci* 10, 184-195.
- Besaratinia, A., and Pfeifer, G.P.** (2009). DNA-lesion mapping in mammalian cells. *Methods* 48, 35-39.
- Beumer, K.J., Trautman, J.K., Bozas, A., Liu, J.L., Rutter, J., Gall, J.G., and Carroll, D.** (2008). Efficient gene targeting in *Drosophila* by direct embryo injection with zinc-finger nucleases. *Proc Natl Acad Sci U S A* 105, 19821-19826.
- Bibikova, M., Carroll, D., Segal, D.J., Trautman, J.K., Smith, J., Kim, Y.G., and Chandrasegaran, S.** (2001). Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol Cell Biol* 21, 289-297.
- Bibikova, M., Golic, M., Golic, K.G., and Carroll, D.** (2002). Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics* 161, 1169-1175.
- Bibikova, M., Beumer, K., Trautman, J.K., and Carroll, D.** (2003). Enhancing gene targeting with designed zinc finger nucleases. *Science* 300, 764.
- Bielle, F., Griveau, A., Narboux-Neme, N., Vigneau, S., Sigrist, M., Arber, S., Wassef, M., and Pierani, A.** (2005). Multiple origins of Cajal-Retzius cells at the borders of the developing pallium. *Nat Neurosci* 8, 1002-1012.
- Blanpain, C., Lowry, W.E., Geoghegan, A., Polak, L., and Fuchs, E.** (2004). Self-renewal, multipotency, and the existence of two cell populations within an epithelial stem cell niche. *Cell* 118, 635-648.
- Boissy, R.E., Sakai, C., Zhao, H., Kobayashi, T., and Hearing, V.J.** (1998). Human tyrosinase related protein-1 (TRP-1) does not function as a DHICA oxidase activity in contrast to murine TRP-1. *Exp Dermatol* 7, 198-204.
- Bollag, R.J., Waldman, A.S., and Liskay, R.M.** (1989). Homologous recombination in mammalian cells. *Annu Rev Genet* 23, 199-225.
- Botchkareva, N.V., Khlgtian, M., Longley, B.J., Botchkarev, V.A., and Gilchrist, B.A.** (2001). SCF/c-kit signaling is required for cyclic regeneration of the hair pigmentation unit. *Faseb J* 15, 645-658.
- Brinster, R.L., Chen, H.Y., Trumbauer, M., Senechal, A.W., Warren, R., and Palmiter, R.D.** (1981). Somatic expression of herpes thymidine kinase in mice following injection of a fusion gene into eggs. *Cell* 27, 223-231.
- Britsch, S., Goerich, D.E., Riethmacher, D., Peirano, R.I., Rossner, M., Nave, K.A., Birchmeier, C., and Wegner, M.** (2001). The transcription factor Sox10 is a key regulator of peripheral glial development. *Genes Dev* 15, 66-78.
- Budd, P.S., and Jackson, I.J.** (1995). Structure of the mouse tyrosinase-related protein-2/dopachrome tautomerase (Tyrp2/Dct) gene and sequence of two novel slaty alleles. *Genomics* 29, 35-43.
- Cabaniols, J.P., and Paques, F.** (2008). Robust cell line development using meganucleases. *Methods Mol Biol* 435, 31-45.
- Cabaniols, J.P., Ouvry, C., Lamamy, V., Fery, I., Craplet, M.L., Moulharat, N., Guenin, S.P., Bedut, S., Nosjean, O., Ferry, G., Devavry, S., Jacqmarcq, C., Lebuhotel, C., Mathis, L., Delenda, C., Boutin, J.A., Duchateau, P., Coge, F., and Paques, F.** (2010). Meganuclease-driven targeted integration in CHO-K1 cells for the fast generation of HTS-compatible cell-based assays. *J Biomol Screen* 15, 956-967.
- Cairns, K., and Finklestein, S.P.** (2003). Growth factors and stem cells as treatments for stroke recovery. *Phys Med Rehabil Clin N Am* 14, S135-142.
- Carbery, I.D., Ji, D., Harrington, A., Brown, V., Weinstein, E.J., Liaw, L., and Cui, X.** (2010). Targeted genome modification in mice using zinc-finger nucleases. *Genetics* 186, 451-459.
- Carroll, D.** (2011). Genome engineering with zinc-finger nucleases. *Genetics* 188, 773-782.
- Chase, H.B.** (1954). Growth of the hair. *Physiol Rev* 34, 113-126.
- Chen, L., Nievera, C.J., Lee, A.Y., and Wu, X.** (2008). Cell cycle-dependent complex formation of BRCA1.CtIP.MRN is important for DNA double-strand break repair. *J Biol Chem* 283, 7713-7720.

- Chevalier, B.S., Kortemme, T., Chadsey, M.S., Baker, D., Monnat, R.J., and Stoddard, B.L.** (2002). Design, activity, and structure of a highly specific artificial endonuclease. *Mol Cell* 10, 895-905.
- Chevalier-Mariette, C., Henry, I., Montfort, L., Capgras, S., Forlani, S., Muschler, J., and Nicolas, J.F.** (2003). CpG content affects gene silencing in mice: evidence from novel transgenes. *Genome Biol* 4, R53.
- Choulika, A., Perrin, A., Dujon, B., and Nicolas, J.F.** (1995). Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccharomyces cerevisiae*. *Mol Cell Biol* 15, 1968-1973.
- Cohen-Tannoudji, M., Robine, S., Choulika, A., Pinto, D., El Marjou, F., Babinet, C., Louvard, D., and Jaissier, F.** (1998). I-SceI-induced gene replacement at a natural locus in embryonic stem cells. *Mol Cell Biol* 18, 1444-1448.
- Cole, F., Keeney, S., and Jasin, M.** (2010). Evolutionary conservation of meiotic DSB proteins: more than just Spo11. *Genes Dev* 24, 1201-1207.
- Colleaux, L., d'Auriol, L., Betermier, M., Cottarel, G., Jacquier, A., Galibert, F., and Dujon, B.** (1986). Universal code equivalent of a yeast mitochondrial intron reading frame is expressed into *E. coli* as a specific double strand endonuclease. *Cell* 44, 521-533.
- Commo, S., Gaillard, O., Thibaut, S., and Bernard, B.A.** (2004). Absence of TRP-2 in melanogenic melanocytes of human hair. *Pigment Cell Res* 17, 488-497.
- Connelly, J.P., Barker, J.C., Pruett-Miller, S., and Porteus, M.H.** (2010). Gene correction by homologous recombination with zinc finger nucleases in primary cells from a mouse model of a generic recessive genetic disease. *Mol Ther* 18, 1103-1110.
- Cotsarelis, G., Cheng, S.Z., Dong, G., Sun, T.T., and Lavker, R.M.** (1989). Existence of slow-cycling limbal epithelial basal cells that can be preferentially stimulated to proliferate: implications on epithelial stem cells. *Cell* 57, 201-209.
- Cui, X., Ji, D., Fisher, D.A., Wu, Y., Briner, D.M., and Weinstein, E.J.** (2011). Targeted integration in rat and mouse embryos with zinc-finger nucleases. *Nat Biotechnol* 29, 64-67.
- Dalgaard, J.Z., Garrett, R.A., and Belfort, M.** (1993). A site-specific endonuclease encoded by a typical archaeal intron. *Proc Natl Acad Sci U S A* 90, 5414-5417.
- Deleyrolle, L., Marchal-Victorion, S., Dromard, C., Fritz, V., Saunier, M., Sabourin, J.C., Tran Van Ba, C., Privat, A., and Hugnot, J.P.** (2006). Exogenous and fibroblast growth factor 2/epidermal growth factor-regulated endogenous cytokines regulate neural precursor cell growth and differentiation. *Stem Cells* 24, 748-762.
- Delmas, V., Beermann, F., Martinozzi, S., Carreira, S., Ackermann, J., Kumasaka, M., Denat, L., Goodall, J., Luciani, F., Viros, A., Demirkan, N., Bastian, B.C., Goding, C.R., and Larue, L.** (2007). Beta-catenin induces immortalization of melanocytes by suppressing p16INK4a expression and cooperates with N-Ras in melanoma development. *Genes Dev* 21, 2923-2935.
- Deng, C., and Capecchi, M.R.** (1992). Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. *Mol Cell Biol* 12, 3365-3371.
- Dinkelmann, M., Spehalski, E., Stoneham, T., Buis, J., Wu, Y., Sekiguchi, J.M., and Ferguson, D.O.** (2009). Multiple functions of MRN in end-joining pathways during isotype class switching. *Nat Struct Mol Biol* 16, 808-813.
- Dion, S., Demattei, M.V., and Renault, S.** (2007). [Zinc finger proteins: tools for site-specific correction or modification of the genome]. *Med Sci (Paris)* 23, 834-839.
- Doetschman, T., Gregg, R.G., Maeda, N., Hooper, M.L., Melton, D.W., Thompson, S., and Smithies, O.** (1987). Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. *Nature* 330, 576-578.
- Donoho, G., Jasin, M., and Berg, P.** (1998). Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. *Mol Cell Biol* 18, 4070-4078.
- Dujon, B., Bolotin-Fukuhara, M., Coen, D., Deutsch, J., Netter, P., Slonimski, P.P., and Weill, L.** (1976). Mitochondrial genetics. XI. Mutations at the mitochondrial locus omega affecting the

- recombination of mitochondrial genes in *Saccharomyces cerevisiae*. *Mol Gen Genet* 143, 131-165.
- Egidy, G., Jule, S., Bosse, P., Bernex, F., Geffrotin, C., Vincent-Naulleau, S., Horak, V., Sastre-Garau, X., and Panthier, J.J.** (2008). Transcription analysis in the MeLiM swine model identifies RACK1 as a potential marker of malignancy for human melanocytic proliferation. *Mol Cancer* 7, 34.
- Elliott, B., Richardson, C., Winderbaum, J., Nickoloff, J.A., and Jasin, M.** (1998). Gene conversion tracts from double-strand break repair in mammalian cells. *Mol Cell Biol* 18, 93-101.
- Epinat, J.C., Arnould, S., Chames, P., Rochaix, P., Desfontaines, D., Puzin, C., Patin, A., Zanghellini, A., Paques, F., and Lacroix, E.** (2003). A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. *Nucleic Acids Res* 31, 2952-2962.
- Erickson, C.A., and Goins, T.L.** (1995). Avian neural crest cells can migrate in the dorsolateral path only if they are specified as melanocytes. *Development* 121, 915-924.
- Ernfors, P.** (2010). Cellular origin and developmental mechanisms during the formation of skin melanocytes. *Exp Cell Res* 316, 1397-1407.
- Evans, M.J., and Kaufman, M.H.** (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292, 154-156.
- Feil, R., Brocard, J., Mascrez, B., LeMeur, M., Metzger, D., and Chambon, P.** (1996). Ligand-activated site-specific recombination in mice. *Proc Natl Acad Sci U S A* 93, 10887-10890.
- Fiering, S., Bender, M.A., and Groudine, M.** (1999). Analysis of mammalian cis-regulatory DNA elements by homologous recombination. *Methods Enzymol* 306, 42-66.
- Fishman-Lobell, J., and Haber, J.E.** (1992). Removal of nonhomologous DNA ends in double-strand break recombination: the role of the yeast ultraviolet repair gene RAD1. *Science* 258, 480-484.
- Foley, J.E., Yeh, J.R., Maeder, M.L., Reyon, D., Sander, J.D., Peterson, R.T., and Joung, J.K.** (2009). Rapid mutation of endogenous zebrafish genes using zinc finger nucleases made by Oligomerized Pool ENgineering (OPEN). *PLoS One* 4, e4348.
- Frank, K.M., Sharpless, N.E., Gao, Y., Sekiguchi, J.M., Ferguson, D.O., Zhu, C., Manis, J.P., Horner, J., DePinho, R.A., and Alt, F.W.** (2000). DNA ligase IV deficiency in mice leads to defective neurogenesis and embryonic lethality via the p53 pathway. *Mol Cell* 5, 993-1002.
- Fuchs, E., Merrill, B.J., Jamora, C., and DasGupta, R.** (2001). At the roots of a never-ending cycle. *Dev Cell* 1, 13-25.
- Fuchs, E.** (2007). Scratching the surface of skin development. *Nature* 445, 834-842.
- Gabriel, R., Lombardo, A., Arens, A., Miller, J.C., Genovese, P., Kaepffel, C., Nowrouzi, A., Bartholomae, C.C., Wang, J., Friedman, G., Holmes, M.C., Gregory, P.D., Glimm, H., Schmidt, M., Naldini, L., and von Kalle, C.** (2011). An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol* 29, 816-823.
- Gao, Y., Sun, Y., Frank, K.M., Dikkes, P., Fujiwara, Y., Seidl, K.J., Sekiguchi, J.M., Rathbun, G.A., Swat, W., Wang, J., Bronson, R.T., Malynn, B.A., Bryans, M., Zhu, C., Chaudhuri, J., Davidson, L., Ferrini, R., Stamato, T., Orkin, S.H., Greenberg, M.E., and Alt, F.W.** (1998). A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. *Cell* 95, 891-902.
- Geurts, A.M., Cost, G.J., Freyvert, Y., Zeitler, B., Miller, J.C., Choi, V.M., Jenkins, S.S., Wood, A., Cui, X., Meng, X., Vincent, A., Lam, S., Michalkiewicz, M., Schilling, R., Foeckler, J., Kalloway, S., Weiler, H., Menoret, S., Anegon, I., Davis, G.D., Zhang, L., Rebar, E.J., Gregory, P.D., Urnov, F.D., Jacob, H.J., and Buelow, R.** (2009). Knockout rats via embryo microinjection of zinc-finger nucleases. *Science* 325, 433.
- Geurts, A.M., and Moreno, C.** (2010). Zinc-finger nucleases: new strategies to target the rat genome. *Clin Sci (Lond)* 119, 303-311.
- Giraldo, P., Martinez, A., Regales, L., Lavado, A., Garcia-Diaz, A., Alonso, A., Busturia, A., and Montoliu, L.** (2003). Functional dissection of the mouse tyrosinase locus control region identifies a new putative boundary activity. *Nucleic Acids Res* 31, 6290-6305.
- Goldberg, A.D., Banaszynski, L.A., Noh, K.M., Lewis, P.W., Elsaesser, S.J., Stadler, S., Dewell, S., Law, M., Guo, X., Li, X., Wen, D., Chagpieri, A., DeKaveler, R.C., Miller, J.C., Lee, Y.L.,**

- Boydston, E.A., Holmes, M.C., Gregory, P.D., Grealley, J.M., Rafii, S., Yang, C., Scambler, P.J., Garrick, D., Gibbons, R.J., Higgs, D.R., Cristea, I.M., Urnov, F.D., Zheng, D., and Allis, C.D.** (2010). Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* 140, 678-691.
- Gossler, A., Doetschman, T., Korn, R., Serfling, E., and Kemler, R.** (1986). Transgenesis by means of blastocyst-derived embryonic stem cell lines. *Proc Natl Acad Sci U S A* 83, 9065-9069.
- Grabher, C., Joly, J.S., and Wittbrodt, J.** (2004). Highly efficient zebrafish transgenesis mediated by the meganuclease I-SceI. *Methods Cell Biol* 77, 381-401.
- Grizot, S., Smith, J., Daboussi, F., Prieto, J., Redondo, P., Merino, N., Villate, M., Thomas, S., Lemaire, L., Montoya, G., Blanco, F.J., Paques, F., and Duchateau, P.** (2009). Efficient targeting of a SCID gene by an engineered single-chain homing endonuclease. *Nucleic Acids Res* 37, 5405-5419.
- Gu, H., Marth, J.D., Orban, P.C., Mossmann, H., and Rajewsky, K.** (1994). Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. *Science* 265, 103-106.
- Guyonneau, L., Rossier, A., Richard, C., Hummler, E., and Beermann, F.** (2002). Expression of Cre recombinase in pigment cells. *Pigment Cell Res* 15, 305-309.
- Guyonneau, L., Murisier, F., Rossier, A., Moulin, A., and Beermann, F.** (2004). Melanocytes and pigmentation are affected in dopachrome tautomerase knockout mice. *Mol Cell Biol* 24, 3396-3403.
- Hardy, M.H.** (1992). The secret life of the hair follicle. *Trends Genet* 8, 55-61.
- Heyting, C., Meijlink, F.C., Verbeet, M.P., Sanders, J.P., Bos, J.L., and Borst, P.** (1979). Fine structure of the 21S ribosomal RNA region on yeast mitochondria DNA. I. Construction of the physical map and localization of the cistron for the 21S mitochondrial ribosomal RNA. *Mol Gen Genet* 168, 231-246.
- Heyting, C., and Menke, H.H.** (1979). Fine structure of the 21S ribosomal RNA region on yeast mitochondrial DNA. III. Physical location of mitochondrial genetic markers and the molecular nature of omega. *Mol Gen Genet* 168, 279-291.
- Higashi, A.Y., Ikawa, T., Muramatsu, M., Economides, A.N., Niwa, A., Okuda, T., Murphy, A.J., Rojas, J., Heike, T., Nakahata, T., Kawamoto, H., Kita, T., and Yanagita, M.** (2009). Direct hematological toxicity and illegitimate chromosomal recombination caused by the systemic activation of CreERT2. *J Immunol* 182, 5633-5640.
- Hirobe, T.** (1992). Control of melanocyte proliferation and differentiation in the mouse epidermis. *Pigment Cell Res* 5, 1-11.
- Hirobe, T.** (1995). Structure and function of melanocytes: microscopic morphology and cell biology of mouse melanocytes in the epidermis and hair follicle. *Histol Histopathol* 10, 223-237.
- Ho, W.H., Armanini, M.P., Nuijens, A., Phillips, H.S., and Osheroff, P.L.** (1995). Sensory and motor neuron-derived factor. A novel heregulin variant highly expressed in sensory and motor neurons. *J Biol Chem* 270, 14523-14532.
- Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., DeKolver, R.C., Katibah, G.E., Amora, R., Boydston, E.A., Zeitler, B., Meng, X., Miller, J.C., Zhang, L., Rebar, E.J., Gregory, P.D., Urnov, F.D., and Jaenisch, R.** (2009). Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat Biotechnol* 27, 851-857.
- Hodgkinson, C.A., Moore, K.J., Nakayama, A., Steingrimsson, E., Copeland, N.G., Jenkins, N.A., and Arnheiter, H.** (1993). Mutations at the mouse microphthalmia locus are associated with defects in a gene encoding a novel basic-helix-loop-helix-zipper protein. *Cell* 74, 395-404.
- Hoek, K., Rimm, D.L., Williams, K.R., Zhao, H., Ariyan, S., Lin, A., Kluger, H.M., Berger, A.J., Cheng, E., Trombetta, E.S., Wu, T., Niinobe, M., Yoshikawa, K., Hannigan, G.E., and Halaban, R.** (2004). Expression profiling reveals novel pathways in the transformation of melanocytes to melanomas. *Cancer Res* 64, 5270-5282.
- Huh, W.J., Mysorekar, I.U., and Mills, J.C.** (2010). Inducible activation of Cre recombinase in adult mice causes gastric epithelial atrophy, metaplasia, and regenerative changes in the absence of "floxed" alleles. *Am J Physiol Gastrointest Liver Physiol* 299, G368-380.

- Hultman, K.A., Budi, E.H., Teasley, D.C., Gottlieb, A.Y., Parichy, D.M., and Johnson, S.L.** (2009). Defects in ErbB-dependent establishment of adult melanocyte stem cells reveal independent origins for embryonic and regeneration melanocytes. *PLoS Genet* 5, e1000544.
- Ip, S.C., Rass, U., Blanco, M.G., Flynn, H.R., Skehel, J.M., and West, S.C.** (2008). Identification of Holliday junction resolvases from humans and yeast. *Nature* 456, 357-361.
- Isalan, M., Klug, A., and Choo, Y.** (1998). Comprehensive DNA recognition through concerted interactions from adjacent zinc fingers. *Biochemistry* 37, 12026-12033.
- Izmiryan, A., Basmaciogullari, S., Henry, A., Paques, F., and Danos, O.** (2011). Efficient gene targeting mediated by a lentiviral vector-associated meganuclease. *Nucleic Acids Res* 39, 7610-7619.
- Jackson, I.J.** (1988). A cDNA encoding tyrosinase-related protein maps to the brown locus in mouse. *Proc Natl Acad Sci U S A* 85, 4392-4396.
- Jackson, I.J., Chambers, D.M., Tsukamoto, K., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., and Hearing, V.** (1992). A second tyrosinase-related protein, TRP-2, maps to and is mutated at the mouse slaty locus. *Embo J* 11, 527-535.
- Jacquier, A., and Dujon, B.** (1985). An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell* 41, 383-394.
- Jasin, M.** (1996). Genetic manipulation of genomes with rare-cutting endonucleases. *Trends Genet* 12, 224-228.
- Jessen, K.R., Brennan, A., Morgan, L., Mirsky, R., Kent, A., Hashimoto, Y., and Gavrilovic, J.** (1994). The Schwann cell precursor and its fate: a study of cell death and differentiation during gliogenesis in rat embryonic nerves. *Neuron* 12, 509-527.
- Jessen, K.R., and Mirsky, R.** (2005). The origin and development of glial cells in peripheral nerves. *Nat Rev Neurosci* 6, 671-682.
- Jiao, Z., Mollaaghababa, R., Pavan, W.J., Antonellis, A., Green, E.D., and Hornyak, T.J.** (2004). Direct interaction of Sox10 with the promoter of murine Dopachrome Tautomerase (Dct) and synergistic activation of Dct expression with Mitf. *Pigment Cell Res* 17, 352-362.
- Jiao, Z., Zhang, Z.G., Hornyak, T.J., Hozeska, A., Zhang, R.L., Wang, Y., Wang, L., Roberts, C., Strickland, F.M., and Chopp, M.** (2006). Dopachrome tautomerase (Dct) regulates neural progenitor cell proliferation. *Dev Biol* 296, 396-408.
- Jordan, S.A., and Jackson, I.J.** (2000). MGF (KIT ligand) is a chemokinetic factor for melanoblast migration into hair follicles. *Dev Biol* 225, 424-436.
- Jurica, M.S., Monnat, R.J., Jr., and Stoddard, B.L.** (1998). DNA recognition and cleavage by the LAGLIDADG homing endonuclease I-CreI. *Mol Cell* 2, 469-476.
- Kabbarah, O., Nogueira, C., Feng, B., Nazarian, R.M., Bosenberg, M., Wu, M., Scott, K.L., Kwong, L.N., Xiao, Y., Cordon-Cardo, C., Granter, S.R., Ramaswamy, S., Golub, T., Duncan, L.M., Wagner, S.N., Brennan, C., and Chin, L.** (2010). Integrative genome comparison of primary and metastatic melanomas. *PLoS One* 5, e10770.
- Kane, P.M., Yamashiro, C.T., Wolczyk, D.F., Neff, N., Goebel, M., and Stevens, T.H.** (1990). Protein splicing converts the yeast TFP1 gene product to the 69-kD subunit of the vacuolar H(+)-adenosine triphosphatase. *Science* 250, 651-657.
- Kass, E.M., and Jasin, M.** (2010). Collaboration and competition between DNA double-strand break repair pathways. *FEBS Lett* 584, 3703-3708.
- Kellendonk, C., Tronche, F., Monaghan, A.P., Angrand, P.O., Stewart, F., and Schutz, G.** (1996). Regulation of Cre recombinase activity by the synthetic steroid RU 486. *Nucleic Acids Res* 24, 1404-1411.
- Kerem, B., Rommens, J.M., Buchanan, J.A., Markiewicz, D., Cox, T.K., Chakravarti, A., Buchwald, M., and Tsui, L.C.** (1989). Identification of the cystic fibrosis gene: genetic analysis. *Science* 245, 1073-1080.
- Kilby, N.J., Snaith, M.R., and Murray, J.A.** (1993). Site-specific recombinases: tools for genome engineering. *Trends Genet* 9, 413-421.
- Kim, J., Lo, L., Dormand, E., and Anderson, D.J.** (2003). SOX10 maintains multipotency and inhibits neuronal differentiation of neural crest stem cells. *Neuron* 38, 17-31.

- Kim, Y.G., and Chandrasegaran, S.** (1994). Chimeric restriction endonuclease. *Proc Natl Acad Sci U S A* 91, 883-887.
- Kim, Y.G., Cha, J., and Chandrasegaran, S.** (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci U S A* 93, 1156-1160.
- Kobayashi, T., Urabe, K., Winder, A., Jimenez-Cervantes, C., Imokawa, G., Brewington, T., Solano, F., Garcia-Borrón, J.C., and Hearing, V.J.** (1994). Tyrosinase related protein 1 (TRP1) functions as a DHICA oxidase in melanin biosynthesis. *Embo J* 13, 5818-5825.
- Koller, B.H., Hagemann, L.J., Doetschman, T., Hageman, J.R., Huang, S., Williams, P.J., First, N.L., Maeda, N., and Smithies, O.** (1989). Germ-line transmission of a planned alteration made in a hypoxanthine phosphoribosyltransferase gene by homologous recombination in embryonic stem cells. *Proc Natl Acad Sci U S A* 86, 8927-8931.
- Kwon, B.S., Haq, A.K., Pomerantz, S.H., and Halaban, R.** (1987). Isolation and sequence of a cDNA clone for human tyrosinase that maps at the mouse c-albino locus. *Proc Natl Acad Sci U S A* 84, 7473-7477.
- Lallemand, Y., Luria, V., Haffner-Krausz, R., and Lonai, P.** (1998). Maternally expressed PGK-Cre transgene as a tool for early and uniform activation of the Cre site-specific recombinase. *Transgenic Res* 7, 105-112.
- Latt, S.A.** (1981). Sister chromatid exchange formation. *Annu Rev Genet* 15, 11-55.
- Le Douarin, N.M., Ziller, C., and Couly, G.F.** (1993). Patterning of neural crest derivatives in the avian embryo: in vivo and in vitro studies. *Dev Biol* 159, 24-49.
- Le Douarin, N.M., Creuzet, S., Couly, G., and Dupin, E.** (2004). Neural crest cell plasticity and its limits. *Development* 131, 4637-4650.
- Lee, K.J., Huang, J., Takeda, Y., and Dynan, W.S.** (2000a). DNA ligase IV and XRCC4 form a stable mixed tetramer that functions synergistically with other repair factors in a cell-free end-joining system. *J Biol Chem* 275, 34787-34796.
- Lee, M., Goodall, J., Verastegui, C., Ballotti, R., and Goding, C.R.** (2000b). Direct regulation of the Microphthalmia promoter by Sox10 links Waardenburg-Shah syndrome (WS4)-associated hypopigmentation and deafness to WS2. *J Biol Chem* 275, 37978-37983.
- Legue, E., and Nicolas, J.F.** (2005). Hair follicle renewal: organization of stem cells in the matrix and the role of stereotyped lineages and behaviors. *Development* 132, 4143-4154.
- Levy, C., Khaled, M., and Fisher, D.E.** (2006). MITF: master regulator of melanocyte development and melanoma oncogene. *Trends Mol Med* 12, 406-414.
- Li, C.Y., Gao, T.W., Wang, G., Han, Z.Y., Shen, Z., Li, T.H., and Liu, Y.F.** (2004). The effect of antisense tyrosinase-related protein 1 on melanocytes and malignant melanoma cells. *Br J Dermatol* 150, 1081-1090.
- Li, H., Haurigot, V., Doyon, Y., Li, T., Wong, S.Y., Bhagwat, A.S., Malani, N., Anguela, X.M., Sharma, R., Ivanciu, L., Murphy, S.L., Finn, J.D., Khazi, F.R., Zhou, S., Paschon, D.E., Rebar, E.J., Bushman, F.D., Gregory, P.D., Holmes, M.C., and High, K.A.** (2011). In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature* 475, 217-221.
- Li, L., Wu, L.P., and Chandrasegaran, S.** (1992). Functional domains in Fok I restriction endonuclease. *Proc Natl Acad Sci U S A* 89, 4275-4279.
- Lieber, M.R.** (2010). The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem* 79, 181-211.
- Lim, D.S., and Hasty, P.** (1996). A mutation in mouse rad51 results in an early embryonic lethal that is suppressed by a mutation in p53. *Mol Cell Biol* 16, 7133-7143.
- Lombardo, A., Genovese, P., Beausejour, C.M., Colleoni, S., Lee, Y.L., Kim, K.A., Ando, D., Urnov, F.D., Galli, C., Gregory, P.D., Holmes, M.C., and Naldini, L.** (2007). Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat Biotechnol* 25, 1298-1306.
- Ludwig, A., Rehberg, S., and Wegner, M.** (2004). Melanocyte-specific expression of dopachrome tautomerase is dependent on synergistic gene activation by the Sox10 and Mitf transcription factors. *FEBS Lett* 556, 236-244.

- Lukacsovich, T., Yang, D., and Waldman, A.S.** (1994). Repair of a specific double-strand break generated within a mammalian chromosome by yeast endonuclease I-SceI. *Nucleic Acids Res* 22, 5649-5657.
- Luo, Y., Robinson, S., Fujita, J., Siconolfi, L., Magidson, J., Edwards, C.K., Wassmann, K., Storm, K., Norris, D.A., Bankaitis-Davis, D., Robinson, W.A., and Fujita, M.** (2011). Transcriptome profiling of whole blood cells identifies PLEK2 and C1QB in human melanoma. *PLoS One* 6, e20971.
- Maacke, H., Jost, K., Opitz, S., Miska, S., Yuan, Y., Hasselbach, L., Luttgies, J., Kalthoff, H., and Sturzbecher, H.W.** (2000). DNA repair and recombination factor Rad51 is over-expressed in human pancreatic adenocarcinoma. *Oncogene* 19, 2791-2795.
- Mackenzie, M.A., Jordan, S.A., Budd, P.S., and Jackson, I.J.** (1997). Activation of the receptor tyrosine kinase Kit is required for the proliferation of melanoblasts in the mouse embryo. *Dev Biol* 192, 99-107.
- Maeder, M.L., Thibodeau-Beganny, S., Osiaik, A., Wright, D.A., Anthony, R.M., Eichinger, M., Jiang, T., Foley, J.E., Winfrey, R.J., Townsend, J.A., Unger-Wallace, E., Sander, J.D., Muller-Lerch, F., Fu, F., Pearlberg, J., Gobel, C., Dassie, J.P., Pruett-Miller, S.M., Porteus, M.H., Sgroi, D.C., Iafrate, A.J., Dobbs, D., McCray, P.B., Jr., Cathomen, T., Voytas, D.F., and Joung, J.K.** (2008). Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. *Mol Cell* 31, 294-301.
- Mansour, S.L., Thomas, K.R., and Capecchi, M.R.** (1988). Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature* 336, 348-352.
- Martin, G.R.** (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A* 78, 7634-7638.
- Mashimo, T., Takizawa, A., Voigt, B., Yoshimi, K., Hiai, H., Kuramoto, T., and Serikawa, T.** (2010). Generation of knockout rats with X-linked severe combined immunodeficiency (X-SCID) using zinc-finger nucleases. *PLoS One* 5, e8870.
- Massari, M.E., and Murre, C.** (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol Cell Biol* 20, 429-440.
- Meier, C., Parmantier, E., Brennan, A., Mirsky, R., and Jessen, K.R.** (1999). Developing Schwann cells acquire the ability to survive without axons by establishing an autocrine circuit involving insulin-like growth factor, neurotrophin-3, and platelet-derived growth factor-BB. *J Neurosci* 19, 3847-3859.
- Meyer, M., de Angelis, M.H., Wurst, W., and Kuhn, R.** (2010). Gene targeting by homologous recombination in mouse zygotes mediated by zinc-finger nucleases. *Proc Natl Acad Sci U S A* 107, 15022-15026.
- Miller, J.C., Holmes, M.C., Wang, J., Guschin, D.Y., Lee, Y.L., Rupniewski, I., Beausejour, C.M., Waite, A.J., Wang, N.S., Kim, K.A., Gregory, P.D., Pabo, C.O., and Rebar, E.J.** (2007). An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol* 25, 778-785.
- Mimitou, E.P., and Symington, L.S.** (2009). Nucleases and helicases take center stage in homologous recombination. *Trends Biochem Sci* 34, 264-272.
- Mintz, B.** (1967). Gene control of mammalian pigmentary differentiation. I. Clonal origin of melanocytes. *Proc Natl Acad Sci U S A* 58, 344-351.
- Monteilhet, C., Perrin, A., Thierry, A., Colleaux, L., and Dujon, B.** (1990). Purification and characterization of the in vitro activity of I-Sce I, a novel and highly specific endonuclease encoded by a group I intron. *Nucleic Acids Res* 18, 1407-1413.
- Morrison, S.J., Perez, S.E., Qiao, Z., Verdi, J.M., Hicks, C., Weinmaster, G., and Anderson, D.J.** (2000). Transient Notch activation initiates an irreversible switch from neurogenesis to gliogenesis by neural crest stem cells. *Cell* 101, 499-510.
- Mortensen, U.H., Bendixen, C., Sunjevaric, I., and Rothstein, R.** (1996). DNA strand annealing is promoted by the yeast Rad52 protein. *Proc Natl Acad Sci U S A* 93, 10729-10734.

- Moure, C.M., Gimble, F.S., and Quioco, F.A.** (2003). The crystal structure of the gene targeting homing endonuclease I-SceI reveals the origins of its target site specificity. *J Mol Biol* 334, 685-695.
- Moury, J.D., and Jacobson, A.G.** (1989). Neural fold formation at newly created boundaries between neural plate and epidermis in the axolotl. *Dev Biol* 133, 44-57.
- Moury, J.D., and Jacobson, A.G.** (1990). The origins of neural crest cells in the axolotl. *Dev Biol* 141, 243-253.
- Moynahan, M.E., Pierce, A.J., and Jasin, M.** (2001). BRCA2 is required for homology-directed repair of chromosomal breaks. *Mol Cell* 7, 263-272.
- Moynahan, M.E., and Jasin, M.** (2010). Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nat Rev Mol Cell Biol* 11, 196-207.
- Muller-Rover, S., Handjiski, B., van der Veen, C., Eichmuller, S., Foitzik, K., McKay, I.A., Stenn, K.S., and Paus, R.** (2001). A comprehensive guide for the accurate classification of murine hair follicles in distinct hair cycle stages. *J Invest Dermatol* 117, 3-15.
- Murisier, F., Guichard, S., and Beermann, F.** (2007). Distinct distal regulatory elements control tyrosinase expression in melanocytes and the retinal pigment epithelium. *Dev Biol* 303, 838-847.
- Murisier, F., Aydin, I.T., Guichard, S., Brunschwiler, S., and Beermann, F.** (2008). Expression from a bacterial artificial chromosome containing the Dct gene locus. *Pigment Cell Melanoma Res* 21, 212-215.
- Nishimura, E.K., Jordan, S.A., Oshima, H., Yoshida, H., Osawa, M., Moriyama, M., Jackson, I.J., Barrandon, Y., Miyachi, Y., and Nishikawa, S.** (2002). Dominant role of the niche in melanocyte stem-cell fate determination. *Nature* 416, 854-860.
- Okabe, M., Ikawa, M., Kominami, K., Nakanishi, T., and Nishimune, Y.** (1997). 'Green mice' as a source of ubiquitous green cells. *FEBS Lett* 407, 313-319.
- Opdecamp, K., Nakayama, A., Nguyen, M.T., Hodgkinson, C.A., Pavan, W.J., and Arnheiter, H.** (1997). Melanocyte development in vivo and in neural crest cell cultures: crucial dependence on the Mitf basic-helix-loop-helix-zipper transcription factor. *Development* 124, 2377-2386.
- Orlow, S.J.** (1995). Melanosomes are specialized members of the lysosomal lineage of organelles. *J Invest Dermatol* 105, 3-7.
- Oro, A.E., and Scott, M.P.** (1998). Splitting hairs: dissecting roles of signaling systems in epidermal development. *Cell* 95, 575-578.
- Oshima, H., Rochat, A., Kedzia, C., Kobayashi, K., and Barrandon, Y.** (2001). Morphogenesis and renewal of hair follicles from adult multipotent stem cells. *Cell* 104, 233-245.
- Pak, B.J., Chu, W., Lu, S.J., Kerbel, R.S., and Ben-David, Y.** (2001). Lineage-specific mechanism of drug and radiation resistance in melanoma mediated by tyrosinase-related protein 2. *Cancer Metastasis Rev* 20, 27-32.
- Pattanayak, V., Ramirez, C.L., Joung, J.K., and Liu, D.R.** (2011). Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat Methods* 8, 765-770.
- Paus, R., and Cotsarelis, G.** (1999). The biology of hair follicles. *N Engl J Med* 341, 491-497.
- Pavan, W.J., and Tilghman, S.M.** (1994). Piebald lethal (sl) acts early to disrupt the development of neural crest-derived melanocytes. *Proc Natl Acad Sci U S A* 91, 7159-7163.
- Pavletich, N.P., and Pabo, C.O.** (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* 252, 809-817.
- Perrin, A., Buckle, M., and Dujon, B.** (1993). Asymmetrical recognition and activity of the I-SceI endonuclease on its site and on intron-exon junctions. *Embo J* 12, 2939-2947.
- Petek, L.M., Russell, D.W., and Miller, D.G.** (2010). Frequent endonuclease cleavage at off-target locations in vivo. *Mol Ther* 18, 983-986.
- Pierce, A.J., Johnson, R.D., Thompson, L.H., and Jasin, M.** (1999). XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes Dev* 13, 2633-2638.
- Pla, P., Solov'eva, O., Moore, R., Alberti, C., Kunisada, T., and Larue, L.** (2004). Dct::lacZ ES cells: a novel cellular model to study melanocyte determination and differentiation. *Pigment Cell Res* 17, 142-149.
- Porter, R.M.** (2003). Mouse models for human hair loss disorders. *J Anat* 202, 125-131.

- Porteus, M.H., and Baltimore, D.** (2003). Chimeric nucleases stimulate gene targeting in human cells. *Science* 300, 763.
- Potterf, S.B., Mollaaghababa, R., Hou, L., Southard-Smith, E.M., Hornyak, T.J., Arnheiter, H., and Pavan, W.J.** (2001). Analysis of SOX10 function in neural crest-derived melanocyte development: SOX10-dependent transcriptional control of dopachrome tautomerase. *Dev Biol* 237, 245-257.
- Raderschall, E., Stout, K., Freier, S., Suckow, V., Schweiger, S., and Haaf, T.** (2002). Elevated levels of Rad51 recombination protein in tumor cells. *Cancer Res* 62, 219-225.
- Rass, E., Grabarz, A., Plo, I., Gautier, J., Bertrand, P., and Lopez, B.S.** (2009). Role of Mre11 in chromosomal nonhomologous end joining in mammalian cells. *Nat Struct Mol Biol* 16, 819-824.
- Rawles, M.E.** (1947). Origin of pigment cells from the neural crest in the mouse embryo. *Physiol Zool* 20, 248-266.
- Redondo, P., Prieto, J., Munoz, I.G., Alibes, A., Stricher, F., Serrano, L., Cabaniols, J.P., Daboussi, F., Arnould, S., Perez, C., Duchateau, P., Paques, F., Blanco, F.J., and Montoya, G.** (2008). Molecular basis of xeroderma pigmentosum group C DNA recognition by engineered meganucleases. *Nature* 456, 107-111.
- Reedy, M.V., Faraco, C.D., and Erickson, C.A.** (1998). The delayed entry of thoracic neural crest cells into the dorsolateral path is a consequence of the late emigration of melanogenic neural crest cells from the neural tube. *Dev Biol* 200, 234-246.
- Riballo, E., Woodbine, L., Stiff, T., Walker, S.A., Goodarzi, A.A., and Jeggo, P.A.** (2009). XLF-Cernunnos promotes DNA ligase IV-XRCC4 re-adenylation following ligation. *Nucleic Acids Res* 37, 482-492.
- Riethmacher, D., Sonnenberg-Riethmacher, E., Brinkmann, V., Yamaai, T., Lewin, G.R., and Birchmeier, C.** (1997). Severe neuropathies in mice with targeted mutations in the ErbB3 receptor. *Nature* 389, 725-730.
- Robertson, E., Bradley, A., Kuehn, M., and Evans, M.** (1986). Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. *Nature* 323, 445-448.
- Rogers, G.E.** (2004). Hair follicle differentiation and regulation. *Int J Dev Biol* 48, 163-170.
- Rothkamm, K., Kruger, I., Thompson, L.H., and Lobrich, M.** (2003). Pathways of DNA double-strand break repair during the mammalian cell cycle. *Mol Cell Biol* 23, 5706-5715.
- Rouet, P., Smih, F., and Jasin, M.** (1994a). Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Mol Cell Biol* 14, 8096-8106.
- Rouet, P., Smih, F., and Jasin, M.** (1994b). Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. *Proc Natl Acad Sci U S A* 91, 6064-6068.
- Saleh-Gohari, N., and Helleday, T.** (2004). Conservative homologous recombination preferentially repairs DNA double-strand breaks in the S phase of the cell cycle in human cells. *Nucleic Acids Res* 32, 3683-3688.
- San Filippo, J., Sung, P., and Klein, H.** (2008). Mechanism of eukaryotic homologous recombination. *Annu Rev Biochem* 77, 229-257.
- Savatier, P., Lapillonne, H., Jirmanova, L., Vitelli, L., and Samarut, J.** (2002). Analysis of the cell cycle in mouse embryonic stem cells. *Methods Mol Biol* 185, 27-33.
- Scheerer, J.B., and Adair, G.M.** (1994). Homology dependence of targeted recombination at the Chinese hamster APRT locus. *Mol Cell Biol* 14, 6663-6673.
- Schlegel, B.P., Jodelka, F.M., and Nunez, R.** (2006). BRCA1 promotes induction of ssDNA by ionizing radiation. *Cancer Res* 66, 5181-5189.
- Schmidt-Ullrich, R., and Paus, R.** (2005). Molecular principles of hair follicle induction and morphogenesis. *Bioessays* 27, 247-261.
- Schouwey, K., Delmas, V., Larue, L., Zimmer-Strobl, U., Strobl, L.J., Radtke, F., and Beermann, F.** (2007). Notch1 and Notch2 receptors influence progressive hair graying in a dose-dependent manner. *Dev Dyn* 236, 282-289.
- Selleck, M.A., and Bronner-Fraser, M.** (1995). Origins of the avian neural crest: the role of neural plate-epidermal interactions. *Development* 121, 525-538.
- Serbedzija, G.N., Fraser, S.E., and Bronner-Fraser, M.** (1990). Pathways of trunk neural crest cell migration in the mouse embryo as revealed by vital dye labelling. *Development* 108, 605-612.

- Serbedzija, G.N., Bronner-Fraser, M., and Fraser, S.E.** (1994). Developmental potential of trunk neural crest cells in the mouse. *Development* 120, 1709-1718.
- Shah, N.M., Marchionni, M.A., Isaacs, I., Stroobant, P., and Anderson, D.J.** (1994). Glial growth factor restricts mammalian neural crest stem cells to a glial fate. *Cell* 77, 349-360.
- Shah, N.M., Groves, A.K., and Anderson, D.J.** (1996). Alternative neural crest cell fates are instructively promoted by TGFbeta superfamily members. *Cell* 85, 331-343.
- Shibahara, S., Tomita, Y., Sakakura, T., Nager, C., Chaudhuri, B., and Muller, R.** (1986). Cloning and expression of cDNA encoding mouse tyrosinase. *Nucleic Acids Res* 14, 2413-2427.
- Simsek, D., Brunet, E., Wong, S.Y., Katyal, S., Gao, Y., McKinnon, P.J., Lou, J., Zhang, L., Li, J., Rebar, E.J., Gregory, P.D., Holmes, M.C., and Jasin, M.** (2011). DNA ligase III promotes alternative nonhomologous end-joining during chromosomal translocation formation. *PLoS Genet* 7, e1002080.
- Singh, T.R., Ali, A.M., Busygina, V., Raynard, S., Fan, Q., Du, C.H., Andreassen, P.R., Sung, P., and Meetei, A.R.** (2008). BLAP18/RMI2, a novel OB-fold-containing protein, is an essential component of the Bloom helicase-double Holliday junction dissolvosome. *Genes Dev* 22, 2856-2868.
- Smih, F., Rouet, P., Romanienko, P.J., and Jasin, M.** (1995). Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. *Nucleic Acids Res* 23, 5012-5019.
- Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A., and Kucherlapati, R.S.** (1985). Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature* 317, 230-234.
- Stark, J.M., and Jasin, M.** (2003). Extensive loss of heterozygosity is suppressed during homologous repair of chromosomal breaks. *Mol Cell Biol* 23, 733-743.
- Stark, J.M., Pierce, A.J., Oh, J., Pastink, A., and Jasin, M.** (2004). Genetic steps of mammalian homologous repair with distinct mutagenic consequences. *Mol Cell Biol* 24, 9305-9316.
- Steel, K.P., and Barkway, C.** (1989). Another role for melanocytes: their importance for normal stria vascularis development in the mammalian inner ear. *Development* 107, 453-463.
- Steel, K.P., Davidson, D.R., and Jackson, I.J.** (1992). TRP-2/DT, a new early melanoblast marker, shows that steel growth factor (c-kit ligand) is a survival factor. *Development* 115, 1111-1119.
- Steingrimsson, E., Copeland, N.G., and Jenkins, N.A.** (2004). Melanocytes and the microphthalmia transcription factor network. *Annu Rev Genet* 38, 365-411.
- Stemple, D.L., and Anderson, D.J.** (1993). Lineage diversification of the neural crest: in vitro investigations. *Dev Biol* 159, 12-23.
- Stenn, K.S., and Paus, R.** (2001). Controls of hair follicle cycling. *Physiol Rev* 81, 449-494.
- Sundberg, J.P., Peters, E.M., and Paus, R.** (2005). Analysis of hair follicles in mutant laboratory mice. *J Invest Dermatol Symp Proc* 10, 264-270.
- Szcepek, M., Brondani, V., Buchel, J., Serrano, L., Segal, D.J., and Cathomen, T.** (2007). Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat Biotechnol* 25, 786-793.
- Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J., and Stahl, F.W.** (1983). The double-strand-break repair model for recombination. *Cell* 33, 25-35.
- Takemoto, Y., Keighren, M., Jackson, I.J., and Yamamoto, H.** (2006). Genomic localization of a Dct-LacZ transgene locus: a simple assay for transgene status. *Pigment Cell Res* 19, 644-645.
- Takeuchi, R., Lambert, A.R., Mak, A.N., Jacoby, K., Dickson, R.J., Gloor, G.B., Scharenberg, A.M., Edgell, D.R., and Stoddard, B.L.** (2011). Tapping natural reservoirs of homing endonucleases for targeted gene modification. *Proc Natl Acad Sci U S A* 108, 13077-13082.
- Thomas, K.R., and Capecchi, M.R.** (1987). Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* 51, 503-512.
- Thompson, S., Clarke, A.R., Pow, A.M., Hooper, M.L., and Melton, D.W.** (1989). Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. *Cell* 56, 313-321.
- Tichy, E.D., and Stambrook, P.J.** (2008). DNA repair in murine embryonic stem cells and differentiated cells. *Exp Cell Res* 314, 1929-1936.

- Tichy, E.D., Pillai, R., Deng, L., Liang, L., Tischfield, J., Schwemberger, S.J., Babcock, G.F., and Stambrook, P.J.** (2010). Mouse embryonic stem cells, but not somatic cells, predominantly use homologous recombination to repair double-strand DNA breaks. *Stem Cells Dev* 19, 1699-1711.
- Tobin, D.J., Foitzik, K., Reinheckel, T., Mecklenburg, L., Botchkarev, V.A., Peters, C., and Paus, R.** (2002). The lysosomal protease cathepsin L is an important regulator of keratinocyte and melanocyte differentiation during hair follicle morphogenesis and cycling. *Am J Pathol* 160, 1807-1821.
- Tobin, D.J.** (2008). Human hair pigmentation--biological aspects. *Int J Cosmet Sci* 30, 233-257.
- Tsukamoto, K., Jackson, I.J., Urabe, K., Montague, P.M., and Hearing, V.J.** (1992). A second tyrosinase-related protein, TRP-2, is a melanogenic enzyme termed DOPACHrome tautomerase. *Embo J* 11, 519-526.
- Tsuzuki, T., Fujii, Y., Sakumi, K., Tominaga, Y., Nakao, K., Sekiguchi, M., Matsushiro, A., Yoshimura, Y., and Morita T** (1996). Targeted disruption of the Rad51 gene leads to lethality in embryonic mice. *Proc Natl Acad Sci U S A* 93, 6236-6240.
- Tumbar, T., Guasch, G., Greco, V., Blanpain, C., Lowry, W.E., Rendl, M., and Fuchs, E.** (2004). Defining the epithelial stem cell niche in skin. *Science* 303, 359-363.
- Urnov, F.D., Miller, J.C., Lee, Y.L., Beausejour, C.M., Rock, J.M., Augustus, S., Jamieson, A.C., Porteus, M.H., Gregory, P.D., and Holmes, M.C.** (2005). Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* 435, 646-651.
- Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S., and Gregory, P.D.** (2010). Genome editing with engineered zinc finger nucleases. *Nat Rev Genet* 11, 636-646.
- Vasquez, K.M., Marburger, K., Intody, Z., and Wilson, J.H.** (2001). Manipulating the mammalian genome by homologous recombination. *Proc Natl Acad Sci U S A* 98, 8403-8410.
- Verastegui, C., Bille, K., Ortonne, J.P., and Ballotti, R.** (2000). Regulation of the microphthalmia-associated transcription factor gene by the Waardenburg syndrome type 4 gene, SOX10. *J Biol Chem* 275, 30757-30760.
- Wang, K., Wang, J.J., Wang, Y., He, Q.H., Wang, X., and Wang, X.M.** (2004). Infusion of epidermal growth factor and basic fibroblast growth factor into the striatum of parkinsonian rats leads to in vitro proliferation and differentiation of adult neural progenitor cells. *Neurosci Lett* 364, 154-158.
- Wang, Y., Lamarche, B.J., and Tsai, M.D.** (2007). Human DNA ligase IV and the ligase IV/XRCC4 complex: analysis of nick ligation fidelity. *Biochemistry* 46, 4962-4976.
- Weeraratna, A.T., Becker, D., Carr, K.M., Duray, P.H., Rosenblatt, K.P., Yang, S., Chen, Y., Bittner, M., Strausberg, R.L., Riggins, G.J., Wagner, U., Kallioniemi, O.P., Trent, J.M., Morin, P.J., and Meltzer, P.S.** (2004). Generation and analysis of melanoma SAGE libraries: SAGE advice on the melanoma transcriptome. *Oncogene* 23, 2264-2274.
- Weinstock, D.M., Nakanishi, K., Helgadottir, H.R., and Jasin, M.** (2006). Assaying double-strand break repair pathway choice in mammalian cells using a targeted endonuclease or the RAG recombinase. *Methods Enzymol* 409, 524-540.
- Wilkie, A.L., Jordan, S.A., and Jackson, I.J.** (2002). Neural crest progenitors of the melanocyte lineage: coat colour patterns revisited. *Development* 129, 3349-3357.
- Winder, A.J., Wittbjer, A., Rosengren, E., and Rorsman, H.** (1993). The mouse brown (b) locus protein has dopachrome tautomerase activity and is located in lysosomes in transfected fibroblasts. *J Cell Sci* 106 (Pt 1), 153-166.
- Wong, E.A., and Capecchi, M.R.** (1986). Analysis of homologous recombination in cultured mammalian cells in transient expression and stable transformation assays. *Somat Cell Mol Genet* 12, 63-72.
- Woods, S.L., and Bishop, J.M.** (2011). A new transgenic mouse line for tetracycline inducible transgene expression in mature melanocytes and the melanocyte stem cells using the Dopachrome tautomerase promoter. *Transgenic Res* 20, 421-428.
- Wu, L., and Hickson, I.D.** (2003). The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* 426, 870-874.
- Xie, A., Kwok, A., and Scully, R.** (2009). Role of mammalian Mre11 in classical and alternative nonhomologous end joining. *Nat Struct Mol Biol* 16, 814-818.

- Xu, D., Guo, R., Soback, A., Bachrati, C.Z., Yang, J., Enomoto, T., Brown, G.W., Hoatlin, M.E., Hickson, I.D., and Wang, W.** (2008). RMI, a new OB-fold complex essential for Bloom syndrome protein to maintain genome stability. *Genes Dev* 22, 2843-2855.
- Yamane, T., Hayashi, S., Mizoguchi, M., Yamazaki, H., and Kunisada, T.** (1999). Derivation of melanocytes from embryonic stem cells in culture. *Dev Dyn* 216, 450-458.
- Yanez, R.J., and Porter, A.C.** (1999). Gene targeting is enhanced in human cells overexpressing hRAD51. *Gene Ther* 6, 1282-1290.
- Yang, C.T., and Johnson, S.L.** (2006). Small molecule-induced ablation and subsequent regeneration of larval zebrafish melanocytes. *Development* 133, 3563-3573.
- Yuan, S.S., Lee, S.Y., Chen, G., Song, M., Tomlinson, G.E., and Lee, E.Y.** (1999). BRCA2 is required for ionizing radiation-induced assembly of Rad51 complex in vivo. *Cancer Res* 59, 3547-3551.
- Zaidi, M.R., Davis, S., Noonan, F.P., Graff-Cherry, C., Hawley, T.S., Walker, R.L., Feigenbaum, L., Fuchs, E., Lyakh, L., Young, H.A., Hornyak, T.J., Arnheiter, H., Trinchieri, G., Meltzer, P.S., De Fabo, E.C., and Merlino, G.** (2011a). Interferon-gamma links ultraviolet radiation to melanomagenesis in mice. *Nature* 469, 548-553.
- Zaidi, M.R., Hornyak, T.J., and Merlino, G.** (2011b). A genetically engineered mouse model with inducible GFP expression in melanocytes. *Pigment Cell Melanoma Res* 24, 393-394.
- Zhang, Y., Andl, T., Yang, S.H., Teta, M., Liu, F., Seykora, J.T., Tobias, J.W., Piccolo, S., Schmidt-Ullrich, R., Nagy, A., Taketo, M.M., Dlugosz, A.A., and Millar, S.E.** (2008). Activation of beta-catenin signaling programs embryonic epidermis to hair follicle fate. *Development* 135, 2161-2172.
- Zijlstra, M., Li, E., Sajjadi, F., Subramani, S., and Jaenisch, R.** (1989). Germ-line transmission of a disrupted beta 2-microglobulin gene produced by homologous recombination in embryonic stem cells. *Nature* 342, 435-438.



# ANNEXE

## Communications à des congrès

- 2010                    13<sup>ème</sup> congrès « Young Researchers in Life Sciences » - poster  
Atelier Inserm « Modification contrôlée du génome à l'aide  
d'endonucléases dirigées à façon » - poster
- 2009                    15th Meeting of the European Society of Pigment Cell Research  
(ESPCR) – poster : **prix du meilleur poster**  
Journées départementales de Biologie du Développement de l'Institut  
Pasteur – poster  
Journées Boris Ephrussi de l'école doctorale « Complexité du Vivant »  
- poster



# A landing pad in mouse embryonic stem (ES) cells : Does double-strand break at the *Dct* locus increase homologous recombination frequency?

Myriam Fenina<sup>1</sup>, Dominique Simon<sup>1</sup>, Sandrine Vandormael-Pournin<sup>1</sup>, Francina Langa-Vives<sup>2</sup>, Bruno Bernard<sup>3</sup>, Michel Cohen-Tannoudji<sup>1</sup>, Jean-Jacques Panthier<sup>1</sup>

<sup>1</sup>Mouse Functional Genetics Unit, Institut Pasteur, URA CNRS 2578, Paris;

<sup>2</sup>Mouse Genetics Engineering Center, Institut Pasteur, Paris;

<sup>3</sup>"Soin, couleur et qualité du cheveu" group, L'Oreal Research, Paris



## Introduction

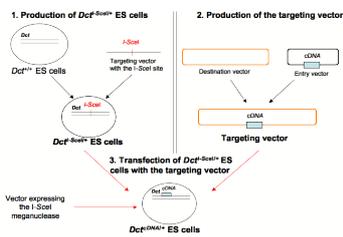
Our project aims at developing a **new strategy** to produce mice carrying virtually any sequence of interest inserted at the *Dct* locus. *Dct* encodes dopachrome tautomerase, a melanogenic enzyme expressed specifically in melanocytes and their precursors. We plan to take advantage of the system to express genes in the **melanocyte lineage** both in the embryo and in the adult.

Indeed, many recently identified genes are involved in the self-renewal and differentiation of melanocyte stem cells or in the melanoma progression. Little is known concerning their precise functions. Genetic tools exist to study their effects *in vivo*. For example, transgenic mice using a *Dct* promoter are currently produced by a number of laboratories. However, these tools are often inaccurate. Actually *Dct::lacZ* transgenic mice express  $\beta$ -galactosidase in ectopic anatomical sites, such as caudal spinal nerves and dorsal root ganglia (MacKenzie *et al.*, 1997). By contrast, our system should allow the accurate expression of genes of interest in pigment cells and their precursors.

## Aim

Insertion by **homologous recombination (HR)** ensures that no ectopic expression, due to position effects, is observed in **knock-in mice**. A limit to this approach is that the frequency of HR, i.e. the percentage of **embryonic stem cells -ES-** clones carrying the transgene at the targeted locus, may be low, depending on the targeted locus. At the *Dct* locus the ratio of HR colonies to the total number of colonies obtained following selection with Neomycin and counter-selection with Gancyclovir is 1/40.

My PhD project relies on the enhanced HR observed following a **double-strand break (DSB)** in ES cells (Cohen-Tannoudji *et al.*, 1998). Our project implies into three steps:



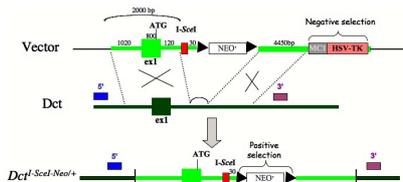
Scheme of the strategy

1. Production of genetically modified ES cells in which the site of recognition of the meganuclease I-SceI, an endonuclease that induces a DSB, is inserted by knock-in at the *Dct* locus.
2. Production of the targeting vector to introduce any gene of interest at the *Dct* locus. The targeting vector was further designed to be used in an iterative way.
3. The genetically modified ES cells will be transfected simultaneously with the targeting vector carrying the gene of interest and a vector expressing the I-SceI meganuclease.

## Results

### Production of *Dct<sup>f</sup>-SceII<sup>+</sup>* ES cells

To introduce the recognition site of the I-SceI meganuclease at the *Dct* locus, we used homologous recombination.



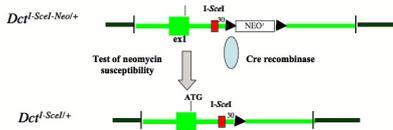
Scheme of the gene targeting: knock-in between ES cells (clone CK35) and a targeting vector containing the recognition site of the I-SceI meganuclease. The external probes 5' and 3' are indicated in blue and violet bars.

After transfection of  $1.6 \times 10^7$  ES cells, 107 Neo<sup>R</sup> Ganc<sup>R</sup> colonies were subcultured. Three HR colonies were identified by PCR analysis. One colony was further studied by Southern blot analysis using probes specific for the 5' and 3' homologous regions.

The frequency of HR (3 HR colonies out of 107 subcultured colonies (3/107)) does not differ significantly from the expected ratio (1/40).

### Deletion of the Neo<sup>R</sup> cassette in *Dct<sup>f</sup>-SceII<sup>+</sup>* ES cells

To delete the neomycin resistance cassette, the Cre recombinase was expressed in *Dct<sup>f</sup>-SceII<sup>+</sup>* ES cells.

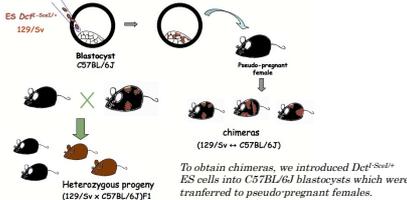


Scheme of gene modification: *Dct<sup>f</sup>-SceII<sup>+</sup>* ES cells were transfected with a vector expressing the Cre recombinase. Clones were isolated in the absence of neomycin. Thirty-six per cent of them died in presence of neomycin (Neo<sup>R</sup>).

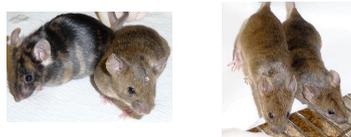
Eight out of 24 subcultured colonies were Neo<sup>R</sup>. Two were further selected and the organization of the *Dct* locus was studied by Southern blot; deletion of the Neo<sup>R</sup> cassette was confirmed.

### Does the *Dct<sup>f</sup>-SceII<sup>+</sup>* ES cells colonize the germ line?

To test whether *Dct<sup>f</sup>-SceII<sup>+</sup>* ES cells are able to colonize the germ line, we introduced cells into C57BL/6J host blastocysts.



Sixteen out of 19 chimeras were more than 95% chimeric, based on their coat colour pattern. Several male chimeras were mated to C57BL/6J females. Half of their progeny was *Dct<sup>f</sup>-SceII<sup>+</sup>*, indicating that the genome of ES cells was transmitted via the germ line.

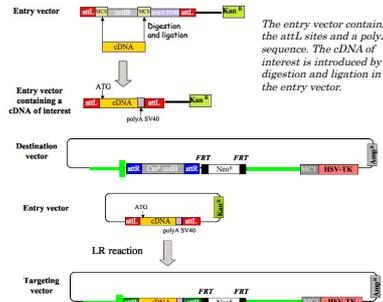


On the left, a mouse that exhibits a 20% coat color chimerism, while another one exhibits a 98% chimerism. On the right, two chimeras exhibit 95-98% coat color chimerisms.

In conclusion, when transferred to blastocysts the *Dct<sup>f</sup>-SceII<sup>+</sup>* ES cells contributed significantly to the coat color of chimeras. Moreover, they are able to colonize the germ line.

### A readily modified targeting vector

To introduce any gene of interest at the *Dct* locus, we generated a targeting vector. We took advantage of the **Gateway technology** which implies the use of 2 vectors: entry and destination vectors.

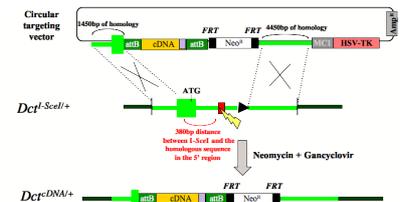


The destination vector contains the region homologous to the *Dct* gene, as well as the negative and positive selection cassettes. The LR reaction allows the insertion of the cDNA of interest within the destination vector though producing the targeting vector.

To test our strategy we introduced H<sub>2</sub>B-mCherry reporter gene at the *Dct* locus.

### Does targeted double-strand break at the *Dct* locus increase the frequency of homologous recombination in mouse ES cells?

*Dct<sup>f</sup>-SceII<sup>+</sup>* ES cells were transfected simultaneously with the targeting vector carrying a cDNA of interest and a vector expressing the I-SceI meganuclease.



Scheme of the gene targeting:  $1.6 \times 10^7$  *Dct<sup>f</sup>-SceII<sup>+</sup>* ES cells were transfected by electroporation with either the linear or the circular targeting vector with or without an expression vector for I-SceI meganuclease.

Experiments	1	2	3	4
	cDNA	cDNA	cDNA	cDNA
	GFP	CMV-F-actin	CAG-F-actin	
Number of Neo <sup>R</sup> Ganc <sup>R</sup> colonies	653	397	370	635
Number of picked colonies	144	144	144	143
PCR analysis in the 5' homologous region	No HR colony	No HR colony	1 HR colony	1 HR colony

Four experiments were performed in parallel. In each of these experiments,  $1.6 \times 10^7$  *Dct<sup>f</sup>-SceII<sup>+</sup>* ES cells were transfected with :

- Exp1. The circular targeting vector (TV) with a mock vector expressing GFP: control experiment
- Exp2. The circular TV with a vector expressing I-SceI under the control of the cytomegalovirus (CMV) promoter
- Exp3. The circular TV with a vector expressing I-SceI under the control of the chicken beta actin promoter fused to the CMV promoter, the so-called CAG promoter - CAG is believed to be a stronger promoter than CMV -.
- Exp4. The linear TV alone: this is the experiment classically performed to obtain knock-in mutants.

These data suggest that:

1. DSB may promote HR at the *Dct* locus (exp1 compared to exp3)
2. HR frequency depends on the I-SceI expression level (exp2 compared to exp3)
3. Interestingly, the number of HR colonies is not significantly different compared with a classical experiment (exp3 compared to exp4).

## Conclusion

According to our data, DSB at the *Dct* locus may promote homologous recombination. It seems that the I-SceI expression level might be a limiting factor. Landing pad in mouse ES cells might be an alternative to classical HR. However, the frequency of HR clones obtained to date is not higher than the frequency obtained following transfection with a linear vector. Further improvements of the method are needed.

### Bibliography:

- Chevaleyre-Marietto, C., Henry I., Montfort, L., Capgras, S., Forlani, S., Muschler, J., and Nicolas, J. F. (2003). CpG content affects gene silencing in mice: evidence from novel transgenes. *Genome Biol* 4, R53.
- Cohen-Tannoudji, M., Robine, S., Choulika, A., Pinto, D., El Marjoui, F., Bahinet, C., Louvard, D., and Jaissier, F. (1998). I-SceI-induced gene replacement at a natural locus in embryonic stem cells. *Mol Cell Biol* 18, 1444-1448.
- Egidij, G., Jule, S., Bosse, P., Bernex, F., Geoffroin, C., Vincent-Naulleau, S., Horak, V., Sastre-Garau, X., and Panthier, J. J. (2008). Transcription analysis in the McLM swine model identifies RACK1 as a potential marker of malignancy for human melanocytic proliferation. *Mol Cancer* 7, 34.
- MacKenzie, M. A., Jordan, S. A., Budd, P. S., and Jackson, I. J. (1997). Activation of the receptor tyrosine kinase Kit is required for the proliferation of melanoblasts in the mouse embryo. *Dev Biol* 192, 99-10

## Acknowledgements

Myriam Fenina is supported by a CIFRE L'Oreal grant.

## **Elaboration d'un outil de modification génétique au locus *Dct* dans les cellules souches embryonnaires de souris à l'aide de la méganucléase I-SceI**

De nombreux gènes sont surexprimés au cours de la progression du mélanome malin cutané. Pour mieux comprendre le rôle de ces gènes, des études fonctionnelles sont nécessaires. Mon projet visait à développer une nouvelle stratégie pour produire des souris portant n'importe quel ADN d'intérêt inséré à la place du premier exon du gène *Dct*. *Dct* code la dopachrome tautomérase, une enzyme de la voie de biosynthèse de la mélanine, exprimée de façon spécifique dans les mélanocytes et leurs précurseurs. Notre stratégie visait à augmenter la fréquence de recombinaison homologue au locus *Dct* dans des cellules souches embryonnaires (ES) en induisant une cassure double-brin à ce locus grâce aux propriétés endonucléases de la méganucléase I-SceI de levure.

Nous avons créé une lignée de cellule ES qui portent un site unique de reconnaissance de la méganucléase I-SceI inséré dans l'intron 1 du gène *Dct*. Notre hypothèse était que l'induction d'une cassure double-brin au locus *Dct* serait recombinogène, et qu'en présence d'une matrice de réparation, la fréquence des recombinaisons homologues à ce locus serait augmentée de façon significative. Nous avons testé cette hypothèse et intégré successivement les gènes rapporteurs *Lago1* et *H<sub>2</sub>B-mCherry* au locus *Dct*. Nos résultats indiquent que, de façon surprenante, l'induction d'une cassure double-brin ne favorise pas la recombinaison homologue au locus *Dct*. Nous avons analysé l'expression des gènes rapporteurs *Lago1* et *H<sub>2</sub>B-mCherry* insérés au locus *Dct* dans des cellules en culture, l'embryon ou chez l'adulte.

## **Development of a tool for the genetic modification at the *Dct* locus in mouse embryonic stem cells using the meganuclease I-SceI**

Many genes are upregulated during the progression of cutaneous malignant melanoma. To better understand the role of these genes, functional studies are needed. Our project aimed at developing a new strategy to produce mice carrying virtually any DNA sequence of interest inserted in place of the first exon of the *Dct* gene. *Dct* encodes dopachrome tautomerase, a melanogenic enzyme expressed in melanocytes and their precursors. Our goal was to increase the frequency of homologous recombination at the *Dct* locus in embryonic stem cells (ES) by inducing double-strand break (DSB) at this locus using the endonuclease properties of the yeast I-SceI meganuclease.

We produced an ES cell line in which a single I-SceI recognition site was inserted into the first intron of the *Dct* gene. Our hypothesis was that inducing a double-strand break at the *Dct* locus would be recombinogenic, and that, in presence of a repair template, the frequency of homologous recombination would increase significantly. We tested this hypothesis and integrated successively the reporter genes, *Lago1* and the *H<sub>2</sub>B-mCherry*, at the *Dct* locus. Our data indicate that, surprisingly, induction of DSB does not promote homologous recombination at the *Dct* locus. We analyzed the expression pattern of the reporter genes, *Lago1* and the *H<sub>2</sub>B-mCherry*, inserted at the *Dct* locus in cultured cells, embryos or adults.