



**HAL**  
open science

# Fabrication de cellules souches embryonnaires humaines exprimant un marqueur fluorescent de la pluripotence grâce à la technologie CRISPR/Cas9

Antoine Cadiou

## ► To cite this version:

Antoine Cadiou. Fabrication de cellules souches embryonnaires humaines exprimant un marqueur fluorescent de la pluripotence grâce à la technologie CRISPR/Cas9 . Génétique humaine. 2018. hal-01698430

**HAL Id: hal-01698430**

**<https://ephe.hal.science/hal-01698430>**

Submitted on 1 Feb 2018

**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.



École Pratique  
des Hautes Études



**MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA  
RECHERCHE**

**ÉCOLE PRATIQUE DES HAUTES ÉTUDES**

**Science de la Vie et de la Terre**

**MEMOIRE présenté par :**

Cadiou Antoine

pour l'obtention du Diplôme de l'École Pratique des Hautes Études

**Fabrication de cellules souches embryonnaires humaines  
exprimant un marqueur fluorescent de la pluripotence grâce à  
la technologie CRISPR/Cas9**

soutenu le : 15 décembre 2017

devant le jury suivant :

Mme Mireille Rossel	- Présidente
Mr Pierre Savatier	- Tuteur scientifique
Mr Jean-Marie Exbrayat	- Tuteur pédagogique
Mr Pascal Mermillod	- Rapporteur
Mme Suzy Markossian	- Examinatrice

Mémoire préparé sous la direction de :

Mr Pierre Savatier

Laboratoire: Institut de Cellule Souche et Cerveau INSERM U1208 à Bron

Directrice : Mme Colette Dehay

et de Mr. Jean-Marie Exbrayat

UMRS 449 Biologie générale – Reproduction de Développement comparé, UCLy/EPHE,  
Lyon

Directeur : Mr. Jean-Marie Exbrayat

EPHE (Science de la Vie et de la Terre)

Groupe de recherche de d'Enseignement Thématiques de l'EPHE : Les Cellules et  
Leur Environnement : Dynamique et Interactions (CEDI)

# Remerciements

---

Mes premiers remerciements s'adressent à Mme **Colette Dehay**, Directrice, pour m'avoir accueilli pendant presque 3 ans dans l'Institut Cellule Souche et Cerveau.

Un immense merci à **Pierre Savatier**. Tu m'as accueilli dans ton équipe, tout d'abord dans le projet lapin pour le stage de troisième année de l'ESTBB, puis au sein du projet humain dans le cadre de l'EPHE. Tu as toujours été accueillant, disponible, à l'écoute et bienveillant envers moi.

**Jean-Marie Exbrayat**. Merci à vous de m'avoir accompagné durant ces deux années à l'EPHE, nous avons eu des échanges riches. Je pensais être l'ultime étudiant que vous accompagneriez pour le diplôme EPHE. Vous avez une santé de fer et vous êtes passionné, alors vous continuez votre travail de tuteur pour le bonheur des étudiants.

**Pierre-Yves : Le coureur au grand cœur**. Tu m'as accompagné tout au long de mon projet EPHE, merci à toi de m'avoir guidé durant ce parcours, aussi bien intellectuel que sportif ! Tu as passé beaucoup de temps pour corriger mon mémoire. Pour tout cela je te remercie.

**Claire : Motarde malgré toi**. Tu es ma voisine de bureau, ma voisine de hotte on partage une paillasse en Bio mol. C'est mon seul espace de travail bien rangé. Car oui tu es beaucoup plus ordonnée que moi ! Merci d'avoir relu mon mémoire pendant ta convalescence

**Anaïs : La touche à tout**. Comment ne pas parler de toi car tu m'as appris à cultiver durant le stage ESTBB et surtout tu m'as appris à porter une boîte de culture sans la faire tomber ! J'espère que tu vas décrocher un poste dans l'équipe.

**Cloé : Le sourire bienveillant**. Toujours au P2 à chouchouter tes cellules et parfois les miennes. Tu prends le temps de prendre des nouvelles de tout le monde. Merci pour le nombre de week-ends que tu as passé pour mes cellules. C'est un bonheur de travailler avec toi.

**Florence : Le travail c'est la santé**. C'est bien rare quand j'arrive avant toi au labo. Je suis content de t'avoir convertie dans la confrérie des Tomates. Depuis que tu es rentrée dans un stade de foot, tu dis « on a gagné » quand l'équipe de Lyon gagne. Merci d'avoir

égayer mes journées car comme Cloé tu as toujours la pêche. Et que dire des mots que tu nous laisses avec des dessins réalisés à la perfection !

**Synara : La lambada.** Tu as suivi le même chemin qu'Anaïs, Lucas et moi en faisant le diplôme EPHE. C'est un très bon choix ! Dommage que tu ne sois pas aussi bonne pâtissière que je l'espérais ! Tu danses tout le temps, tu es fière de tes origines brésiliennes, et tu souhaites nous faire découvrir ton pays pour notre plus grand plaisir. Dans un an, tu écriras ton mémoire et tu seras diplômée. J'espère être présent pour ta soutenance et déguster des mets brésiliens.

**Marielle : La maman lapin.** Tu as été ma tutrice lors de mon stage ESTBB, tu prends soin du groupe lapin comme tu prends soin de tes enfants.

**Fabrice : Le couteau suisse du labo.** Au début de mon EPHE tu étais la personne que tout le monde venait voir pour avoir une info pratique alors que tu écrivais ta thèse. Maintenant tu es docteur, tu as une petite fille. Merci de ton aide pour les expériences et pour ta bonne humeur.

**Lucas : Le serial killer.** Tu as été mon voisin de chambre lors de nombreux modules à Paris. Tu es diplômé depuis peu et tu as un beau projet. Je te souhaite de réussir dans cette voie.

**Irène : La tempête.** Ton rire communicatif, ta bonne humeur ont rythmé nos repas. J'espère rester ton chouchou. Et encore bravo pour le petit Eliott.

**Christian : L'expert Bio mol et CRISPR.** Dès qu'un nouvel outil sort sur la technologie CRISPR tu le testes. Je suis content que tu restes au SBRI.

**Etienne et Wilhelm : Jean-Michel et Lucien.** C'est très agréable de vous voir intégrer l'équipe car depuis quelques temps je me sentais un peu seul face à toutes ces femmes ! Vous êtes arrivés en master 2 et maintenant vous êtes doctorants avec un article dans Nature qui vous attend.

**Mouna : L'embryologiste.** Tu es arrivée pour écrire ta thèse et tu as réussi avec brio ton doctorat. J'espère que tu continues à manger des framboises !

**Joyce : La pro des PCR.** Les PCR n'ont plus de secret pour toi !

**Diana : Un corps sain dans un esprit sain.** Tu es la nouvelle de l'équipe, tu vas cultiver les cellules que j'ai fabriquées. Ça me fait plaisir de voir ton enthousiasme et ton intérêt pour

ce que nous faisons dans l'équipe. Profite de tes week-ends pour découvrir la ville avec ton compagnon et ta fille.

**Thierry : Labguru manager** Tu es devenu en quelques mois responsable de toute la logistique de l'équipe. Merci de ton aide avec les extractions et les PCR.

**Aurélie :** Je te souhaite de réussir dans la gendarmerie. S'il te plait sois indulgente si tu m'arrêtes un jour !

**Nathalie :** Tu es arrivée en provenance de la campagne de Paris, avec Mouna dans tes bagages. C'est super d'avoir une belge qui fête la St Nicolas.

**Murielle :** Tu puises ton énergie non pas dans les mitochondries mais dans l'aviron et les randonnées en plein air.

Merci aux personnes du P2 Primates (Florence, Rafal, Charlotte et Léo) de votre patience car nous monopolisons hottes et microscopes toute la journée.

Merci Arnaud pour ton aide avec les constructions des plasmides, et bravo pour ton nouveau poste.

Merci à Chandara et Nicolas pour votre bonne humeur et pour le travail que vous fournissez.

Merci à toutes les personnes du SBRI pour la bonne ambiance qui règne dans cette unité.

Et merci à toute ma famille qui m'a supporté durant ces deux années. Et spécialement à ma mère qui m'a préparé de bon petits plats tous les jours.

# Table des matières

---

<b>Remerciements</b> .....	<b>2</b>
<b>Table des matières</b> .....	<b>5</b>
<b>Table des illustrations</b> .....	<b>9</b>
<b>Tableaux</b> .....	<b>12</b>
<b>Abréviations</b> .....	<b>13</b>
<b>Présentation</b> .....	<b>16</b>
<b>Introduction</b> .....	<b>17</b>
I. Les cellules souches embryonnaires .....	17
I. 1. Définition .....	17
I. 2. Propriétés .....	17
I. 2. 1) La pluripotence .....	17
I. 2. 2) L'autorenouvellement .....	17
I. 2. 3) La tumorigénèse .....	17
I. 3. Les ESC des différentes espèces .....	18
I. 3. 1) Le modèle murin .....	18
I. 3. 2) Le modèle humain .....	19
I. 4. Les gènes de régulations de l'état de pluripotence .....	20
I. 4. 1) <i>OCT4</i> (POU5F1) .....	20
I. 4. 2) <i>Nanog</i> .....	21
I. 4. 3) <i>Sox2</i> .....	21
II. Modification du génome .....	22
II. 1. La nucléase en doigt de zinc et les TALENs .....	22
II. 2. CRISPR Cas 9 .....	22
II. 2. 1) Immunité bactérienne .....	23
II. 2. 2) Nucléase Cas 9 .....	24
II. 2. 2. 1) Structure .....	24
II. 2. 2. 2) Interactions avec d'autres molécules .....	24
II. 2. 2. 2. 1) <i>ARN guide</i> .....	25

II. 2. 2. 2) Site PAM .....	25
II. 2. 3) Développement de la technologie .....	26
II. 2. 3. 1) Réponse cellulaire au clivage de l'ADN par la Cas9 .....	26
II. 2. 3. 1. 1) NHEJ .....	26
II. 2. 3. 1. 2) Recombinaison Homologue .....	26
II. 2. 3. 2) Adaptation de CRISPR/Cas9 dans la modification génomique de.....	27
cellules eucaryotes.....	27
II. 2. 3. 3) « Off-targets » .....	28
II. 2. 3. 4) Choix de la protéine Cas à utiliser .....	28
II. 2. 3. 4. 1) Cas9_D10A .....	28
II. 2. 3. 4. 2) CPF1.....	29
II. 2. 3. 4. 3) SpCas9-HF1 .....	30
II. 2. 3. 4. 4) Utilisations des différentes méthodes. ....	30
II. 2. 3. 5) Amélioration de la recombinaison homologue et diminution des « off-targets » ...	31
II. 2. 3. 5. 1) SCR7 et RS-1 .....	31
II. 2. 3. 5. 2) Nocodazole et ABT-751 .....	31
II. 2. 3. 5. 3) Anti-CRISPR .....	31
II. 2. 4) Utilisation de la technologie CRISPR Cas9 dans les hESC.....	32
<b>Projet EPHE .....</b>	<b>34</b>
<b>Matériels et Méthodes.....</b>	<b>35</b>
I. Techniques de biologie cellulaire .....	35
I. 1. Préparation des fibroblastes embryonnaires de souris (MEF) .....	35
I. 1. 1) Dérivation et congélation des MEF .....	35
I. 1. 2) Décongélation des MEFs et traitement à la mytocyamine C .....	35
I. 2. Culture des cellules ES humaines .....	37
I. 2. 1) Préparation du bFGF.....	37
I. 2. 2) Conditions de culture.....	37
I. 2. 3) Passage cellulaire .....	37
I. 2. 4) Congélation et décongélation .....	38
I. 2. 4. 1) Congélation .....	38
I. 2. 4. 2) Décongélation.....	38

I. 2. 5) Préparation du milieu conditionné (MC).....	38
I. 3. Transfection par électroporation .....	38
I. 4. Sélection des clones cellulaires .....	39
I. 5. Culot sec.....	40
II. Techniques de biologie moléculaire .....	40
II. 1. Test PCR contamination Mycoplasme .....	40
II. 2. Purification des plasmides par la méthode du PEG/LiCl .....	41
II. 2. 1) Amplification plasmidique .....	41
II. 2. 2) Extraction plasmidique.....	41
II. 2. 3) Purification .....	42
II. 3. Fabrication des plasmides CRISPR .....	42
II. 3. 1) Design des sgRNA .....	42
II. 3. 2) Construction du plasmide CRISPR ( <a href="#">Sakuma et al, 2014</a> ).....	42
II. 4. Extraction ADN génomique.....	44
<b>Résultats .....</b>	<b>45</b>
I. Outils nécessaires à la mise en place de la technologie CRISPR/Cas9 .....	45
I. 1. Construction des plasmides d'expression des guides et de la Cas9.....	45
I. 1. 1) Design des guides (sgRNA) .....	45
I. 1. 2) Construction des plasmides pX330A_D10A_sgRNA et pX330S_sgRNA.....	46
I. 2. Construction de la matrice de réparation .....	50
I. 2. 1) Mutation des sites PAM dans la matrice de réparation .....	51
I. 2. 2) Fabrication d'une matrice de réparation contenant le gène codant pour la protéine fluorescente Kate2.....	53
II. Fabrication d'une lignée fluorescente révélant l'expression du gène OCT4.....	54
II. 1. Mises au point des conditions de transfections.....	54
II. 2. Fabrication d'une lignée OSCAR OCT4-P2A-eGFP-Puro.....	56
II. 2. 1) Obtention des clones .....	56
II. 2. 2) Caractérisation des clones.....	56
II. 3. Fabrication d'une lignée OSCAR OCT4-P2A-Kate2-Puro.....	59
II. 3. 1) Obtention des clones .....	59
II. 3. 2) . Caractérisation des clones .....	59

II. 3. 3) Analyse de l'expression de la protéine Katushka dans la lignée OSCAR <i>OCT4-Kate2-Puro</i> .....	63
II. 3. 3. 1) Analyse de l'expression de la protéine Katushka dans les cellules ..... indifférenciées. ....	64
II. 3. 3. 2) Analyse de l'expression de la protéine Kate2 dans les cellules ..... différenciées.....	65
II. 4. Fabrication d'une lignée OSCAR <i>OCT4-Kate2</i> avec excision de la cassette pGK-Puro loxée.....	67
II. 4. 1) Analyse de l'excision de la cassette de sélection par PCR sur la population totale OSCAR <i>OCT4-Kate2</i> .....	67
II. 4. 2) Analyse de l'expression du gène Kate2 après excision de la cassette pGK-Puro69	
<b>Discussions et conclusion.....</b>	<b>71</b>
I. Fabrication d'une lignée hESC rapportrice de l'état de pluripotence.....	72
II. Perspectives .....	74
III. Conclusion personnelle .....	75
<b>Bibliographie .....</b>	<b>76</b>
<b>Annexes.....</b>	<b>81</b>
Annexe 1 : Amorces PCR pour la caractérisation des plasmides pX330A_D10A et pX330S2 .....	81
Annexe 2 : Amorces PCR pour la caractérisation des cellules OSCAR <i>OCT4-Kate2-Puro</i>	81

# Table des illustrations

---

Figure 1 : Les propriétés des ESC .....	18
Figure 2 : Noyau de pluripotence.....	20
Figure 3 : Immunité adaptative de la bactérie par le système CRISPR Cas9 .....	24
Figure 4 : Interactions du sgRNA dans le systèmes CRISPR (A) Interaction sgRNA/Cas9 (TracrRNA). (B) Interaction sgRNA/ADN (crRNA). .....	25
Figure 5 : Reconnaissance du site PAM par la Cas9 et activation des sites catalytiques HNH et RuvC .....	26
Figure 6 : Stratégie de la cellule pour réparer son ADN après une cassure double brins causer par le système CRISPR-Cas9 .....	27
Figure 7 : Système CRISPR-Cas9 pour effectuer une cassure double brins par la stratégie de la double nickases (Cas9_D10A). .....	29
Figure 8 : Système CRISPR pour effectuer une cassure double brins par la stratégie la nucléase CPF1 .....	30
Figure 9 : Images de 2 colonies, l'une indifférenciée et l'autre différenciée, pour lesquelles sont observées l'expression du gène <i>OCT4</i> par immunofluorescence, et marquage des noyaux cellulaires au 4,6-diamidino-2-phenylindole (DAPI). (Images de la colonie différencier tirée de l'article de Babaie et collaborateur ( <a href="#">Babaie et al, 2006</a> )).....	34
Figure 10 : Gel d'agarose du profil type d'un test Mycoplasme.....	40
Figure 11 : positionnement des 8 sgRNA conçus autour du codon STOP du gène <i>OCT4</i> .....	46
Figure 12 : Construction des plasmides pX330A_D10A-G1_a et pX330S2-G1_b. (A) Hybridation des amorces codant pour l'ARNg, puis ligation des oligonucléotides double-brun dans les plasmides pX330A_D10A et pX330S2. (B) Analyse par PCR des minipréparations des plasmides pX330A_D10A_G1a et pX330S2_G1b puis électrophorèse sur gel d'agarose.....	48
Figure 13 : Construction des plasmide pX330A_D10A-G1, plasmide pX330A_D10A-G2, plasmide pX330A_D10A-G3 et plasmide pX330A_D10A-G4. (A) Insertion du fragment BsaI contenant la cassette promoteur U6-Guide1b (G1_b) dans le plasmide pX330A_D10A-G1_a grâce à la technologie Golden Gate assembly. (B) Analyse par PCR des minipréparations des plasmides pX330A_D10A_G1 puis électrophorèse sur gel d'agarose. (C) Analyse par PCR des minipréparations des plasmides pX330A_D10A_G2, pX330A_D10A_G3 et pX330A_D10A_G4 puis électrophorèse sur gel d'agarose. ....	49
Figure 14 : Matrice de réparation permettant d'insérer le gène codant pour la GFP à la place du codon STOP du gène <i>OCT4</i> , dans les cellules hESC. ....	50
Figure 15 : Mutation des sites PAM dans la matrice de réparation Oct4-P2A-eGFP-PGK-Puro. (A) Mutations à réaliser dans la matrice Oct4-P2A-eGFP-PGK-Puro (B) Insertion de la séquence contenant les sites PAM mutés dans le plasmide pOCT4-eGFP-2A-Puro (C) Gel d'agarose 0,8% des minipréparations pOCT4-eGFP-2A-Puro-PAM_mutés, après	

linéarisation. (D) Séquençage du clone 4 montrant la présence des mutations C926T, C953T, C974A et C989T. .... 52

Figure 16 : Remplacement de la GFP par Kate 2 dans la matrice de réparation. (A) Insertion de la séquence contenant le gène Kate2 dans le plasmide pOCT4-2A-eGFP-Puro-PAM\_mutés (B) Gel d'agarose 0,8% d'une minipréparation pOCT4-2A-Kate2-Puro-PAM\_mutés, après linéarisation. .... 53

Figure 17 : Analyse de l'expression de la GFP dans les cellules OSCAR transfectées avec le plasmide pGAE-GFP. (A) Cellules hESC OSCAR non transfectées (B) Dans des cellules hESC OSCAR transfectées avec l'appareil Gene Pulser II ou Neon, en amas et en suspension cellulaire. (C) Histogramme montrant le pourcentage d'hESC OSCAR exprimant la GFP en fonction de la condition de transfection. .... 55

Figure 18 : Criblage des clones puromycine résistants obtenus après transfection de la matrice pOCT4-eGFP-2A-Puro. (A) Schéma représentant la structure d'un allèle *OCT4* après recombinaison homologue avec la matrice pOCT4-2A-eGFP-Puro. Les flèches noires représentent les amorces OCT4g\_For et GFP\_Rev. Les flèches rouges représentent les amorces Mat-3'\_For et OCT4g-3'\_Rev. (B) Analyse par PCR de l'ADN génomique des clones résistants à la puromycine avec les amorces OCT4g\_For et GFP\_Rev. Le gène de la GAPDH est utilisé comme contrôle interne, (C) Analyse par PCR de l'ADN génomique du clone n°3 avec les amorces Mat-3'\_For et OCT4g-3'\_Rev. Le gène de la GAPDH est utilisé comme contrôle interne..... 57

Figure 19 : Etude de l'hétérogénéité de l'insertion du pOCT4-eGFP-2A-Puro-PAM\_mutés dans le clone n°3. (A) Schéma représentant la structure d'un allèle *OCT4* après recombinaison homologue avec la matrice pOCT4-eGFP-2A-Puro et un allèle sauvage du gène *OCT4*. Les flèches noires représentent les amorces OCT4g\_For et OCT4\_Rev. (B) Analyse par PCR de l'ADN génomique du clone n°3 et des cellules OSCAR. Le gène de la GAPDH est utilisé comme contrôle interne..... 58

Figure 20 : Criblage des clones puromycine résistants, obtenus après transfection de la matrice pOCT4-2A-Kate2-Puro-PAM\_mutés. (A) Schéma représentant la structure d'un allèle *OCT4* après recombinaison homologue avec la matrice pOCT4-2A-Kate2-Puro-PAM\_mutés. Les flèches noires représentent les amorces OCT4g\_For et Kate2\_Rev. Les flèches rouges représentent les amorces Mat-3'\_For et OCT4g-3'\_Rev. (B) Analyse par PCR de l'ADN génomique des clones résistants à la puromycine, avec les amorces OCT4g\_For et Kate2\_Rev. Le gène de la GAPDH est utilisé comme contrôle interne. (C) Séquençage de la région du codon STOP du gène *OCT4* dans les cellules hESC OSCAR et les clones ayant intégré la matrice de réparation dans le locus du gène *OCT4*. (D) Analyse par PCR de l'ADN génomique des clones ayant intégré la matrice de réparation dans le locus du gène *OCT4*, avec les amorces Mat-3'\_For et OCT4g-3'\_Rev..... 60

Figure 21 : Etude de l'hétérogénéité de l'insertion du pOCT4-eGFP-2A-Puro dans les clones positifs. (A) Schéma représentant la structure d'un allèle *OCT4* après recombinaison homologue avec la matrice pOCT4-eGFP-2A-Puro et un allèle sauvage du gène *OCT4*. Les flèches noires représentent les amorces OCT4g\_For et OCT4\_Rev. (B) Analyse par PCR de l'ADN génomique des clones OSCAR *OCT4*-Kate2-Puro résistants à la puromycine et des cellules OSCAR. (C) Séquençage de la région du codon STOP du gène *OCT4* pour les cellules hESC OSCAR et du clone 46 OSCAR *OCT4*-Kate2-Puro. (D) Séquences insérées la région du codon STOP du gène *OCT4* pour les clones 5 et 26 OSCAR *OCT4*-Kate2-Puro. .... 62

- Figure 22 : Analyse de l'expression de la protéine Kate2 dans les cellules OSCAR *OCT4-Kate2-Puro\_46* (A) Spectre d'excitation et d'émission de la protéine Kate2. (B) Observation de l'expression de la protéine Kate2 dans des colonies de cellules OSCAR et de cellules OSCAR *OCT4-Kate2-Puro#46*, par microscopie à fluorescence. (C) Analyse en cytométrie en flux de l'expression de Kate2 dans les cellules OSCAR et de cellules OSCAR *OCT4-Kate2-Puro#46*. ..... 64
- Figure 23 : Analyse de l'expression de la protéine Kate2 dans les cellules OSCAR *OCT4-Kate2-Puro\_46* indifférenciées et différenciées. Observation au microscope à fluorescence et analyse par cytométrie en flux de l'expression de Kate2 dans les cellules OSCAR (A) et dans les cellules OSCAR *OCT4-Kate2-Puro\_46* indifférenciées (B) et différenciées (C). ..... 66
- Figure 24 : Représentation schématique de l'excision de la cassette PGK-Puro par la recombinaison Cre..... 67
- Figure 25 : Excision de la cassette pGK-Puro dans les cellules OSCAR *OCT4-Kate2-Puro\_46* (A) Schéma représentant la structure de l'allèle *OCT4-P2A-Kate2-Puro*, avant et après excision de la cassette pGK-Puro. Les flèches orange représentent les amorces Kate2\_For et OCT4\_Rev. (B) Analyse par PCR de l'ADN génomique des cellules OSCAR-*OCT4-Kate2-Puro\_46* avant et après transfection de la Cre Recombinase, à l'aide des oligonucléotides Kate2\_For et OCT4\_Rev. Le gène de la GAPDH est utilisé comme contrôle interne. .... 68
- Figure 26 : Analyse de l'expression de Kate2 dans les cellules OSCAR *OCT4-Kate2-Puro* transfectées avec la recombinaison Cre par cytométrie en flux avant (B) et après (C) transfection. Les cellules OSCAR sont utilisées comme contrôle négatif (A). ..... 69

# Tableaux

---

Tableau 1 : Exemples de Knock-In dans des hESC à l'aide de la technologie CRISPR/Cas9.	33
Tableau 2 : Composition du milieu de culture des MEF.....	36
Tableau 3 : Conditions d'électroporation de Gene Pulser II et de Neon.....	39
Tableau 4 : (A) séquence des sgRNA. (B) Oligonucléotides sens et anti-sens contenant la séquence du sgRNA (minuscules) et la séquence nécessaire au clonage dans le site BsbI (majuscules). (C) Oligonucléotides double-brin obtenus après hybridation des oligonucléotides sens et anti-sens, insérés dans le site BsbI des plasmides pX330A_D10A et pX330S2. ....	47
Tableau 5 : Récapitulatif des résultats de la création de lignées OSCAR <i>OCT4</i> -P2A-Kate2-Puro .....	63

# Abréviations

---

ADN : Acide DésoxyriboNucléique

ARN : Acide RiboNucléique

ARNm : Acide RiboNucléique messenger

dFGF : basic Fibroblast Growth Factor

BSA : Bovine Sérum Albumine

CPF1 : CRISPR-associated endonuclease in Prevotella and Francisella 1

CRISPR : Clustered Regularly Interspaced Short Palindromic Repeats

crRNA : CRISPR RiboNucleic Acid

DMEM : Dulbecco's Modified Eagle's Medium

DMSO : Dimethylsulfoxide

DSB : Double Strand Breaks

EB : Embryoïd Body

EPHE : Ecole Pratique des Hautes Etudes

EpiSC : Epiblast Stem Cell ou cellule souche de l'épiblaste

ESC : Embryonic Stem Cell ou cellule souche embryonnaire

FACS : cytomètre en flux ou Fluorescence-Activated Cell Sorting

FBS : Foetal Bovine Serum ou sérum de veau foetal

GFP : Green Fluorescent Protein

h : Heure

HDR : Homology Directed Repair

hESC : human Embryonic Stem Cell ou cellule souche embryonnaire humaine

hiPSC : human induced Pluripotent Stem Cell ou cellule souche pluripotente induite humaine

i : inhibiteur

ICM : Inner Cell Mass ou masse cellulaire interne

J : Jour

Kate2 : Katushka 2

KI : Knock-In

KO : Knock-Out

KO-DMEM : Knock Out Dulbecco's Modified Eagle's Medium ou milieu de base des hESC

KOSR : Knock Out Serum Replacement

LIF: Leukemia Inhibitory Factor

MC : Milieu Conditionné sur MEF inactivés

MEF : Fibroblastes embryonnaires murins

mESC : mouse Embryonic Stem Cell ou cellule souche embryonnaire de souris

min: minutes

MMC : Mitomycine C

Nb : Nombre

NEAA : Non Essential Amino-Acide

NHEJ : Non Homologous End Joining

nhESC : non human Embryonic Stem Cell ou cellule souche embryonnaire non humaine

P : Passage cellulaire

PAM : Protospacer Adjacent Motif

pb : paire de bases

PBS : Phosphate Phosphate-Buffered Saline

PCR : Polymerase Chain Reaction

POU : Pict Oct Unc

PSG : Pénicilline Streptomycine Glutamine

PSMII : Poste de Sécurité Microbiologique de type II

Puro : Puromycine

rbESC : rabbit Embryonic Stem Cell ou cellule souche embryonnaire de lapin

ROCK : Rho-associated Coiled-coil Kinase

SCR7 : SCR7 pyrazine

sec : seconde

sgRNA : single strand guide RiboNucleic Acid

TALEN : Transcription Activator Like Effector Nuclease

TE: Tris EDTA

tracrRNA : trans-activating CRISPR RiboNucleic Acid

# Présentation

---

J'ai été accueilli dans l'Institut Cellule souche et Cerveau (INSERM U1208), dirigé par Mme Colette Dehay. J'ai préparé et réalisé le diplôme EPHE dans l'équipe « Cellules souches pluripotentes chez les mammifères » dirigé par Mr Pierre Savatier, qui a été mon tuteur scientifique, avec l'aide de Mr Pierre-Yves Bourillot. Mr Jean-Marie Exbrayat a été mon tuteur pédagogique. Le travail de recherche que j'ai effectué lors de ces deux années porte sur la création d'une lignée rapportrice de l'état de pluripotence dans des cellules souches embryonnaires humaines (hESC) en utilisant la technologie du CRISPR/Cas9 pour insérer un gène de fluorescence. Cet outil cellulaire permettra de suivre l'évolution de la pluripotence des cellules en temps réel, en fonction des cocktails moléculaires testés. Les projets de l'équipe de Pierre Savatier visent à étudier la pluripotence des cellules souches pluripotentes dans plusieurs espèces tels que le macaque (rhECS), le lapin (RbECS) et la souris (mECS). Dans ce mémoire, je présenterai de brèves notions sur les cellules souches embryonnaires (ESC), je résumerai les connaissances sur la technologie CRISPR/Cas9. Puis j'exposerai plus précisément mon projet EPHE. Ensuite, j'exposerai les méthodes utilisées et les résultats de mon projet. Enfin, je discuterai de mon travail et je décrirai les perspectives.

# Introduction

---

## **I. Les cellules souches embryonnaires**

### **I. 1. Définition**

Les cellules souches embryonnaires (ESC) dérivent de la masse cellulaire interne ICM (en anglais Inner Cell Mass) à partir d'un embryon au stade blastocyste préimplantatoire. Elles ont été isolées, cultivées, et maintenues en culture *in vitro* pour la première fois chez la souris en 1981 ([Evans et Kaufman 1981](#)). Contrairement aux cellules somatiques, les ESC sont capables de se multiplier indéfiniment en préservant la propriété de pluripotence.

### **I. 2. Propriétés**

Les ESC sont des cellules indifférenciées qui comportent deux propriétés fondamentales : la pluripotence, et l'autorenouvellement. Elles possèdent également un potentiel tumoral qui les distingue des cellules somatiques.

#### **I. 2. 1) La pluripotence**

La pluripotence est la capacité des ESC à se différencier en cellules des trois feuillets embryonnaires : l'ectoderme, le mésoderme et l'endoderme. Il existe plusieurs techniques qui permettent de mettre en évidence la pluripotence des ESC. La première consiste à cultiver les ESC dans des conditions favorisant la différenciation et ainsi former des corps embryoides (EB). Les EB sont des structures cellulaires tridimensionnelles. Après section des EB, on peut observer des cellules hématopoïétiques, endothéliales, musculaires et neuronales issues des trois feuillets ([Keller 1995](#)).

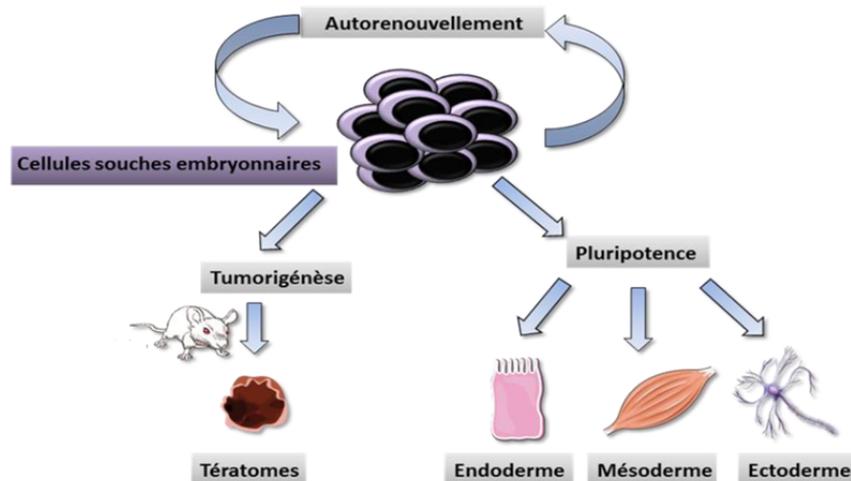
#### **I. 2. 2) L'autorenouvellement**

L'autorenouvellement se caractérise par la capacité des ESC à se multiplier indéfiniment. La phase G1 du cycle cellulaire est réduite, ce qui permet une vitesse de prolifération élevée ([Savatier et al, 1996](#) ; [Savatier et al, 1994](#) ; [Coronado et al, 2013](#)). Les ESC ont une forte activité télomérase conservant la taille des télomères qui leur confère une grande stabilité génétique ([Armstrong et al, 2000](#)).

#### **I. 2. 3) La tumorigénèse**

La tumorigénèse est la propriété des cellules ESC à générer des tumeurs lorsqu'elles sont placées en dehors de leur contexte embryonnaire. Ces tumeurs sont non cancéreuses car les

cellules n'ont pas de modifications génétiques. Cette propriété est vérifiée par l'injection de cellules sous la capsule des glandes surrénales de souris immunodéprimées qui entraîne la formation de tératomes possédant des cellules différenciées issues des trois feuillets embryonnaires. Cette expérience permet de vérifier également la propriété de pluripotence des ESC (*Martin and Evans 1974*) (Figure 1).



**Figure 1 : Les propriétés des ESC**

### **I. 3. Les ESC des différentes espèces**

#### **I. 3. 1) Le modèle murin**

Les mESC sont les premières ESC à avoir été isolées en 1981 par Evans et Kaufman. Au départ, les mESC sont cultivées en présence de cellules nourricières (fibroblaste de souris). Le milieu de culture était composé de 10% de sérum de veau fœtal et 10% de sérum de veau nouveau-né. Aujourd'hui, le Leukemia Inhibitory Factor (LIF) a remplacé le sérum de veau nouveau-né et les cellules nourricières a permis d'améliorer la culture des mESC en permettant de les cultiver sur gélatine. En culture, elles forment des petites colonies compactes en forme de dôme. Elles sontensemencées de manière unicellulaire après avoir été dissociées à la trypsine EDTA (*Williams et al. 1998, Smith et al. 1988*). De plus les mESC sont capables de former des chimères germinales. C'est d'ailleurs la seule espèce, avec le rat, dont les cellules ESC possèdent cette capacité.

En 2007, des lignées de cellules ont été dérivées à partir de l'épiblaste d'un embryon au stade post-implantatoire. Ces nouvelles lignées sont appelées Epiblast Stem cells (EpiSCs). Contrairement aux mESC, les mEpiSC sont cultivées en présence d'activine et du facteur de croissance bFGF (en anglais basic Fibroblast Growth Factor) (*Brons et al, 2007*). En culture, les mEpiSC forment des colonies étendues. De plus elles tendent à se différencier spontanément lorsqu'elles sont ensemencées de façon unicellulaire. Cela nécessite de les repiquer manuellement en petits amas.

### **I. 3. 2) Le modèle humain**

Obtenues pour la première fois en 1998, les cellules ES humaines (hESC) sont obtenues à partir de l'ICM d'un blastocyste au stade préimplantatoire (*Thomson, 1998*). Les embryons utilisés dans la recherche sont des embryons surnuméraires issus de fécondation *in vitro* (FIV) qui ne serviront pas à des fins médicales. C'est dans ce cadre que Thomson et collaborateurs ont pu obtenir des embryons humains et dériver les cinq premières lignées de cellules ES humaines.

Les hESC se multiplient en formant des colonies plates comme les EpiSCs. Elles sont amplifiées par dissociation mécanique en amas car elles rentrent en différenciation sous l'effet de la dissociation enzymatique à la trypsine. Elles possèdent une forte activité télomérase et phosphatase alcaline. Les hESC sont stables génétiquement. En effet elles peuvent être cultivées au moins un an sans présenter d'anomalie chromosomique (*Amit et al, 2000*). Comme toutes les cellules pluripotentes, les hESC sont capables de se différencier *in vitro et in vivo* en formant des EBs ainsi que des tératomes. Une étude a montré que la culture en condition hypoxique (un taux d'oxygène de 5%) permettait d'obtenir des hESC possédant deux chromosome X actifs (*Lengner et al, 2010*). Enfin, pour des questions éthiques, la capacité de ces cellules à coloniser l'embryon humain ne peut être testée, du moins en France.

Les premières lignées de hESC ont été dérivées en présence de 20% de FBS sur MEF (*Thomson, 1998*). Très rapidement, le milieu de culture a été remplacé par le KO-DMEM + 20% de sérum de remplacement (KOSR) (*Xu et al, 2001*) et est encore utilisé aujourd'hui. En utilisant ces conditions de culture, les auteurs ont pu cultiver les hESC sans cellules nourricières, sur matrigel ou laminine, avec du milieu préalablement conditionné par des MEF. Ceci suggère que les MEF produisent des facteurs de croissance nécessaires à l'autorenouvellement des cellules ES humaines. Différentes équipes ont alors cherché à caractériser les facteurs de croissance soutenant la prolifération des hESC.

## I. 4. Les gènes de régulations de l'état de pluripotence

Les cellules souches indifférenciées ne sont pas stables *in vitro* comme *in vivo*, et ont tendance à se différencier rapidement. Pour les aider à se maintenir dans l'état de pluripotence *in vitro*, l'étude des gènes concernés a été nécessaire. Aujourd'hui, beaucoup de gènes ont été identifiés en particulier chez la souris. Ils sont responsables du maintien de l'autorenouvellement et de la pluripotence des cellules et ils ne s'expriment plus quand elles sont différenciées.

Trois gènes majeurs impliqués dans le maintien de la pluripotence ont été identifiés. Il s'agit du gène *OCT4* (OCTamer-binding transcription factor 4), *Sox2* (SRY-related HMG-box (High Mobility Group) gene 2) et *Nanog*. Ils s'auto-activent mutuellement, activent les facteurs de pluripotence, et répriment les facteurs de différenciation (Boyer et al., 2005 ; Boyer et al., 2006). Ils sont situés au centre des mécanismes du maintien de la pluripotence et sont retrouvés exprimés dans toutes les cellules pluripotentes étudiées jusqu'à aujourd'hui (Figure 2).

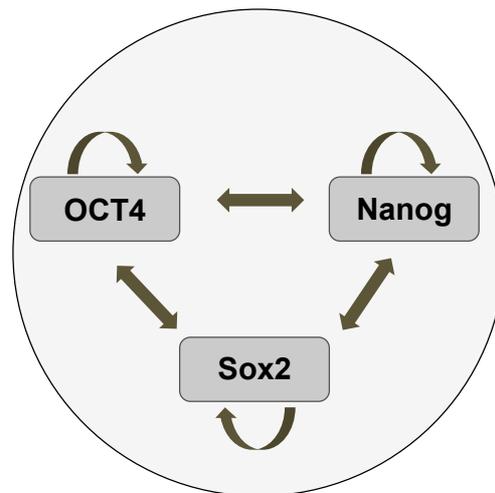


Figure 2 : Noyau de pluripotence

### I. 4. 1) *OCT4* (POU5F1)

Le gène *OCT4* appartient aux facteurs de transcriptions POU (Pict, Oct, Unc) ayant un domaine de fixation à l'ADN nommé POU. Ce gène comprend cinq exons et se trouve sur le chromosome 6. (Takeda et al, 1992). Le gène *OCT4* se manifeste dans tous les stades de la segmentation embryonnaire ; de l'ovocyte non fertilisé jusqu'au blastocyste. Dans les blastocystes, l'expression du gène *OCT4* a été observée dans l'ICM et le TE, mais les études

de titrage ont révélé un niveau d'expression plus élevé dans l'ICM ([Hansis et al, 2000](#)). Le gène *OCT4* permet le maintien des cellules à l'état indifférencié par l'inhibition des gènes de différenciations. En effet, l'inhibition du gène *OCT4* entraîne la différenciation des cellules dans les lignages embryonnaires, alors que la surexpression du gène *OCT4* induit la différenciation en endoderme et mésoderme ([Niwa et al, 2000](#)). L'autorégulation du gène *OCT4* est donc importante pour le maintien de la pluripotence et de l'état indifférencié des cellules pluripotentes ([Pesce et al, 1998](#)).

#### **I. 4. 2) *Nanog***

*Nanog* code une protéine contenant un homéodomaine permettant la fixation à l'ADN. L'expression constitutive de *Nanog* permet l'autorenouveaulement uniquement des cellules indifférenciées. Mais *Nanog* n'est pas capable de soutenir la pluripotence si le gène *OCT4* ne s'exprime plus. *Nanog* participe également à la régulation de l'expression du gène *OCT4* ([Chambers et al, 2003](#)).

#### **I. 4. 3) *Sox2***

L'expression de *Sox2* est nécessaire à l'autorenouveaulement des cellules pluripotentes. Cependant la surexpression provoque la diminution de l'expression de *Nanog* pouvant induire la dérégulation du gène *OCT4*. La réduction de l'expression de *Sox2* provoque la perte de l'état indifférencié et la baisse d'expression de *Nanog* et du gène *OCT4* ([Fong et al, 2008](#)).

## **II. Modification du génome**

Grace au séquençage du génome de plusieurs espèces, aussi bien animales que végétales et bactériennes, réalisé ces dernières années, la recherche sur le génome a explosé. Chaque jour permet la découverte de plusieurs gènes. Plus important encore, il est possible de réaliser des modifications génétiques ciblées. Cela signifie que l'on est capable de modifier une région précise de l'ADN. Il existe trois technologies principales permettant de réaliser ces mutations : La nucléase en doigt de zinc, les TALENs et CRISPR/Cas9.

### **II. 1. La nucléase à doigts de zinc et les TALENs**

Au début des années 2000, les chercheurs ont développé des nucléases à doigts de zinc qui sont des protéines synthétiques dont les domaines de liaison à l'ADN leur permettent de créer des cassures double brins de l'ADN dans des régions génomiques ciblées (*Smith et al, 2000*).

En 2010, les nucléases synthétiques appelées activateurs de la transcription des nucléases effectrices (TALEN en anglais *Transcription Activators- Like Effector Nucleases*) ont fourni un moyen plus simple de cibler l'ADN pour effectuer une rupture à double brin (*Christian et al, 2010*).

Les nucléases en doigt de zinc et les TALEN nécessitent la création d'une protéine personnalisée pour chaque séquence d'ADN ciblée. Ces deux technologies sont coûteuses et ne sont pas efficaces.

La technologie des séquences CRISPR (en anglais *Clustered Regularly Interspaced Short Palindromic Repeats*) associée à l'endonucléase Cas 9 (*CRISPR associated protein 9*) est, quant à elle, beaucoup moins coûteuse et plus facile d'utilisation.

### **II. 2. CRISPR Cas 9**

En quelques années, la technologie CRISPR/Cas9 s'est développée pour devenir la référence dans le domaine de la correction génétique.

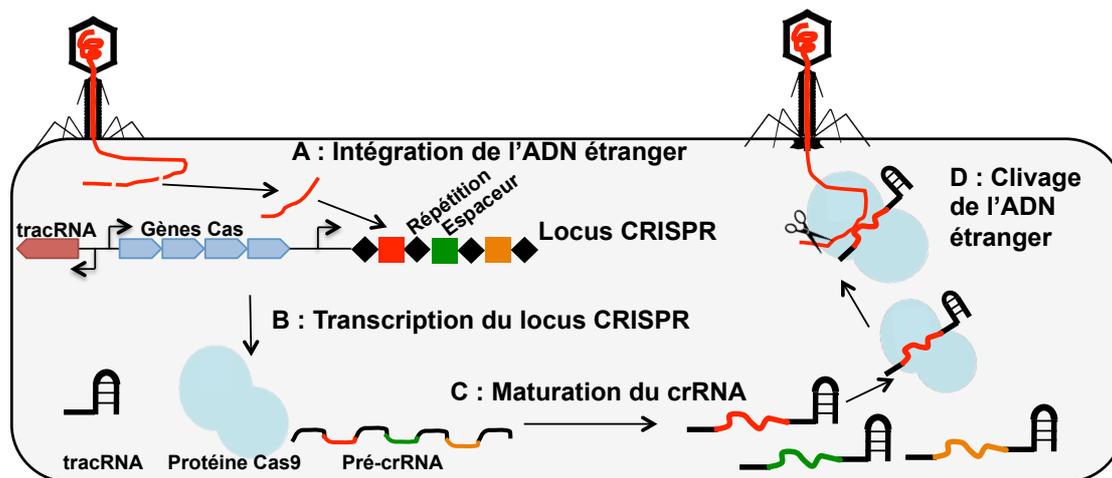
A l'origine, le CRISPR a été observé en 1987 dans une région inter-génique de la bactérie *Escherichia coli* (*Ishino et al, 1987*). Ces régions répétées ont été observées dans plusieurs organismes bactériens comme *Streptococcus pyogenes* et *Mycobacterium tuberculosis*. Ces répétitions ont été retrouvées dans des archéobactéries. En 2000, on a mis en évidence que les répétitions étaient issues d'une même entité appelée SRSR pour *Short Regularly Spaced*

*Repeats* (Mojica et al, 2000). C'est finalement en 2002 que l'acronyme CRISPR est adopté par la communauté scientifique suite à un consensus (Jansen et al, 2002).

## II. 2. 1) Immunité bactérienne

En 2005, 3 groupes de recherche observent que les séquences d'ADN séparant les répétitions CRISPR sont des ADN invasifs qui appartiennent aux bactériophages ou aux plasmides. Par la suite, on a mis en évidence que le CRISPR était un système de défense chez les procaryotes (Barrangou et al, 2007). Lors de l'infection par un bactériophage par exemple, le virus expulse son information génétique dans la bactérie. La bactérie répond en coupant l'ADN du phage en petites séquences de 20 paires de bases (pb). L'une d'entre elles est capturée par des transferrines et intégrée dans le locus CRISPR comme nouvel "espaceur" (Figure 3).

Lors d'une seconde infection par le bactériophage la bactérie transcrit la région CRISPR. A cette étape la bactérie transcrit 3 nouvelles entités. La première est un ARN messager (mRNA) codant pour la nucléase Cas9. La seconde est l'ARN CRISPR (crRNA en anglais CRISPR RNA) immature pre-crRNA contenant l'information génétique étrangère et les répétitions. Le dernier, le tracrRNA (*trans-activating CRISPR RNA*), permet la maturation du crRNA (Figure 3). En effet, le tracrRNA et le pre-crRNA s'hybrident formant ainsi un ARN double brin. Celui-ci est clivé par la ribonucléase III donnant lieu à un ARN guide simple brin (sgRNA en anglais single guide RNA) (Karvelis et al, 2013) (Figure 3). Enfin, le sgRNA recrute la nucléase Cas9 puis la guide vers la région génomique homologue pour qu'elle clive l'ADN étranger et inactive ainsi la prolifération du bactériophage ou du plasmide (Figure 3). Ce système est un moyen de défense contre les invasions extérieures. C'est en quelque sorte une immunité de la bactérie. Celle-ci est adaptative car elle fonctionne comme chez l'homme en deux phases. La première infection permet de mettre en place une reconnaissance et une mémoire du bactériophage par la bactérie. La seconde infection permet à la bactérie de neutraliser le phage (Garnau et al, 2010) (Figure 3).



**Figure 3 : Immunité adaptative de la bactérie par le système CRISPR Cas9**

## II. 2. 2) Nucléase Cas 9

Il existe 2 classes de CRISPR/Cas, soit un complexe multi-protéine pour la classe 1, soit une protéine unique pour la classe 2. La Cas9 se situe dans la deuxième catégorie. Lors de l'acquisition, une courte séquence d'ADN viral est intégrée comme un nouvel "espaceur". Il s'agit d'un processus dépendant de Cas1-Cas2, ces deux Cas étant indispensables dans les deux classes de CRISPR/Cas. Ils sont présents parmi les gènes Cas représentés ci-dessus. Le CRISPR/Cas de classe 1 regroupe 3 Types (Type I, III et IV). Il en est de même pour la classe 2 (type II, V et VI). La Cas9 fait partie du type II car elle effectue une cassure double brin à bout net sur un ADN après avoir reconnu la région cible avec une petite séquence de l'ADN qui engage le processus de clivage. Cette séquence est appelée PAM en anglais (*Protospacer Adjacent Motif*).

### II. 2. 2. 1) Structure

Quatre domaines actifs caractérisent les nucléases Cas9. Deux domaines sont dits « domaines catalytiques » HNH et RuvC. Ces nucléases clivent l'ADN et forment une cassure double-brins. Les deux autres domaines sont dits « domaines de reconnaissance » : ils sont en charge d'amener la protéine à sa cible. Il s'agit du domaine REC (en anglais *alpha-helical recognition*) qui permet la fixation de la protéine aux sgRNA, et le domaine PI (en PAM interacting), région de la protéine de reconnaissance avec l'ADN (5'-NGG). Cette interaction a lieu juste avant le clivage (*Jinek, et al, 2014*).

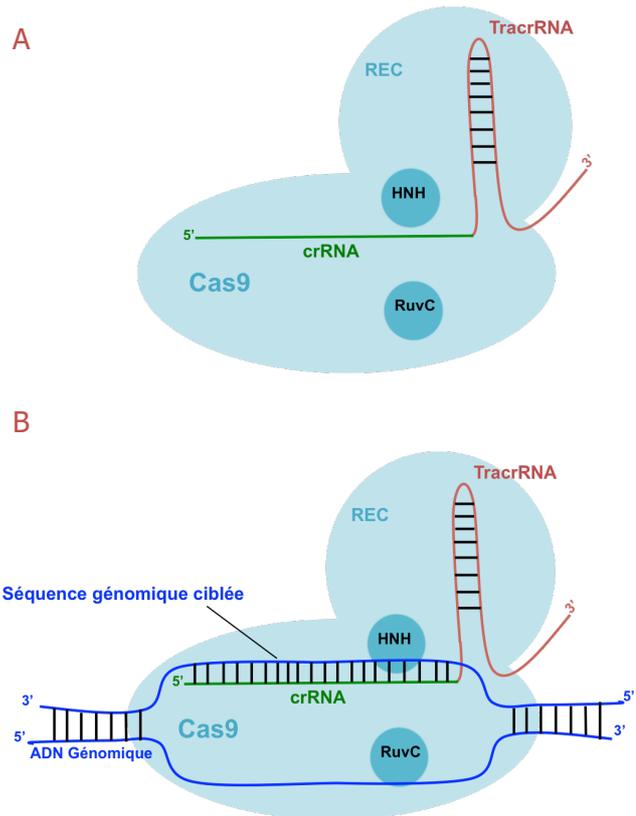
### II. 2. 2. 2) Interactions avec d'autres molécules

La Cas9 est une enzyme qui clive l'ADN. Pour que ce clivage se fasse de manière ciblée, il est nécessaire que la protéine se fasse guider.

### II. 2. 2. 2. 1) ARN guide

Le sgRNA est recruté par la Cas9 du le coté 3'.(Figure 4A).

Ensuite, ce complexe se fixe de manière spécifique sur une séquence d'ADN double-brin par la région 5' ou crRNA contenant une séquence homologue à l'ADN, qui s'hybride par complémentarité sur les 20 pb ainsi ciblé. La présence du site PAM est indispensable pour que la Cas9 clive l'ADN. Pour que la nucléase clive l'ADN (Figure 4B).



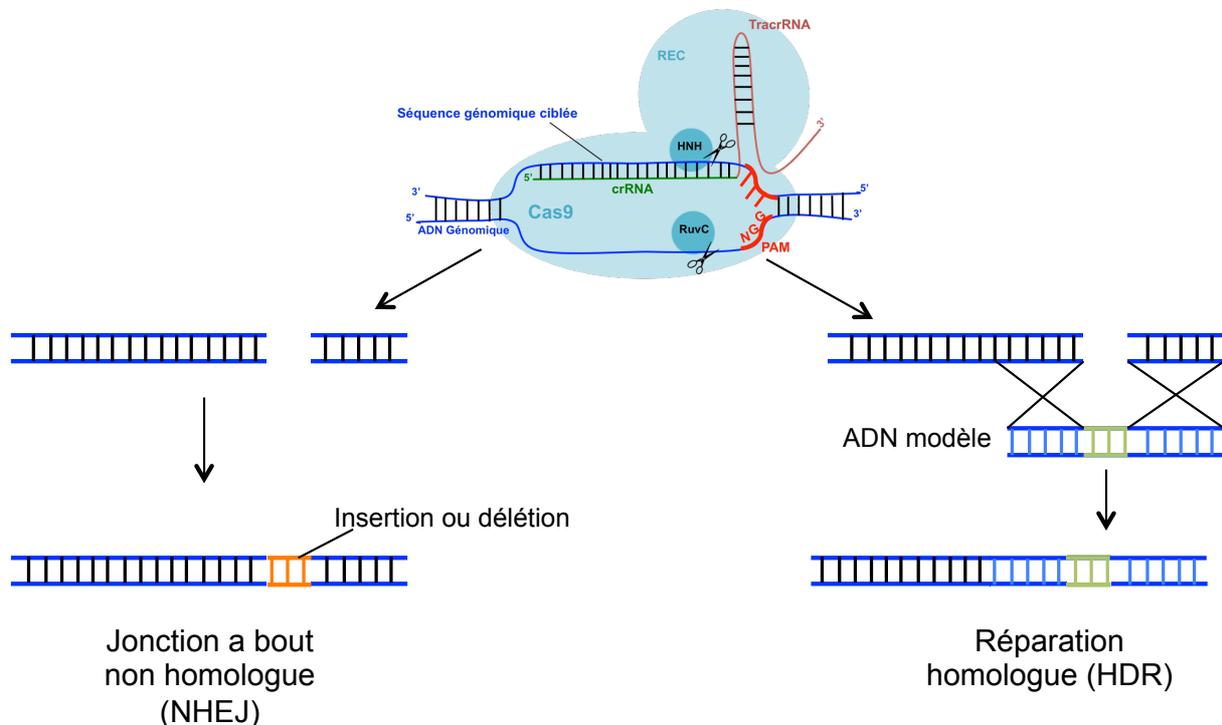
**Figure 4 : Interactions du sgRNA dans le systèmes CRISPR (A) Interaction sgRNA/Cas9 (TracrRNA). (B) Interaction sgRNA/ADN (crRNA).**

### II. 2. 2. 2. 2) Site PAM

Le site PAM est indispensable pour que la Cas9 clive l'ADN car, s'il est absent, la protéine ne clivera pas l'ADN. Un site PAM reconnu pas la Cas9 est constitué de 3 nucléotides. Il s'agit souvent de la séquence 5'-NGG mais, dans de récentes publications, il a été démontré que la séquence 5'-NAG permettrait dans une moindre mesure à la Cas9 de cliver l'ADN. Le site PAM fait partie de l'ADN. Il suit l'hybridation ADN/sgRNA de 20pb. Quand la Cas9 reconnaît le site PAM, elle active ses sites catalytiques HNH et RuvC et clive l'ADN en formant une cassure double brins. Ce processus inactive la prolifération du phage ou du plasmide entré dans la bactérie. C'est ce qu'on appelle l'immunité adaptative de la bactérie (*Hsu et al, 2014*) (Figure 5).



généétique. Dans le second cas, la modification génétique est comprise entre les deux bras d'homologie. RAD51 et BRAC 2 sont des facteurs qui contrôlent la HDR (*Jayathilaka et al 2008*). (Figure 6).



**Figure 6 : Stratégie de la cellule pour réparer son ADN après une cassure double brins causer par le système CRISPR-Cas9**

### **II. 2. 3. 2) Adaptation de CRISPR/Cas9 dans la modification génomique**

#### **des cellules eucaryotes**

En 2012 un article est publié, mettant en évidence les propriétés de modification de génome. Ce travail a été réalisé par une collaboration entre les équipes de Jennifer Doudna et Emmanuelle Charpentier (*Jinek, et al, 2012*). Début 2013, deux équipes se sont intéressées aux propriétés de ces éléments pour faire de la correction du génome dans des cellules eucaryotes. Pour cela elles ont adapté le système CRISPR/Cas9 d'une bactérie pathogène pour l'homme, le *Streptococcus pyogenes*, afin de modifier le génome de cellules murines et humaines (*Cong et al, 2013 ; Mali et al, 2013*). Ils ont montré que la Cas9 pouvait cliver l'ADN malgré des disparités dans l'homologie ADN cible et le crARN, ce qui engendre des recombinaisons aléatoires dans le génome (« off-targets »).

### **II. 2. 3. 3) « Off-targets »**

La séquence de 20pb qui permet le complexe ADN-ARN doit être spécifique d'une région du génome car, dans le cas contraire, des cassures en dehors de la région ciblée peuvent apparaître. Ce sont les « off-targets ». Des études ont montré que la spécificité de la Cas9 est liée aux 8 à 12 premières bases de séquence de guidage proximal PAM ([Fu et al, 2013](#) ; [Hsu et al, 2013](#)). Ces études démontrent que Cas9 tolère les disparités dans la séquence de guidage à partir du douzième nucléotide du site PAM. Elle réussit à cliver un ADN ayant jusqu'à cinq disparités avec la séquence du sgRNA si celles-ci sont situées entre le treizième et le vingtième nucléotide. Mises bout à bout, les fréquences de « off-targets » peuvent se rapprocher voire dépasser la fréquence de clivage de la région ciblée par la Cas9.

Il est important de noter que le PAM 5'-NAG peut également conduire à un clivage « off-target ».

D'autres facteurs rentrent en jeu comme la quantité de sgRNA et de Cas9 introduites dans les cellules ainsi que la durée de l'activation de la Cas9. En effet plus il y a de protéine dans la cellule plus elle clivera d'ADN. Par conséquent, il y aura plus de risque de voir apparaître des cassures d'ADN dans les « off-targets ». L'ouverture de la chromatine est aussi un facteur essentiel dans la fixation de la Cas9 sur l'ADN.

Le principal danger dans les « off-targets » est la modification involontaire de gènes ou de marques épigénétiques. Pour minimiser les risques, de nouvelles approches sont abordées. Il est probable qu'une combinaison parfaite de dix bases dans la région proximale de PAM de la séquence de guidage de l'ARNm soit suffisante pour mener la liaison de Cas9 à l'ADN ([Kuscu et al, 2014](#)).

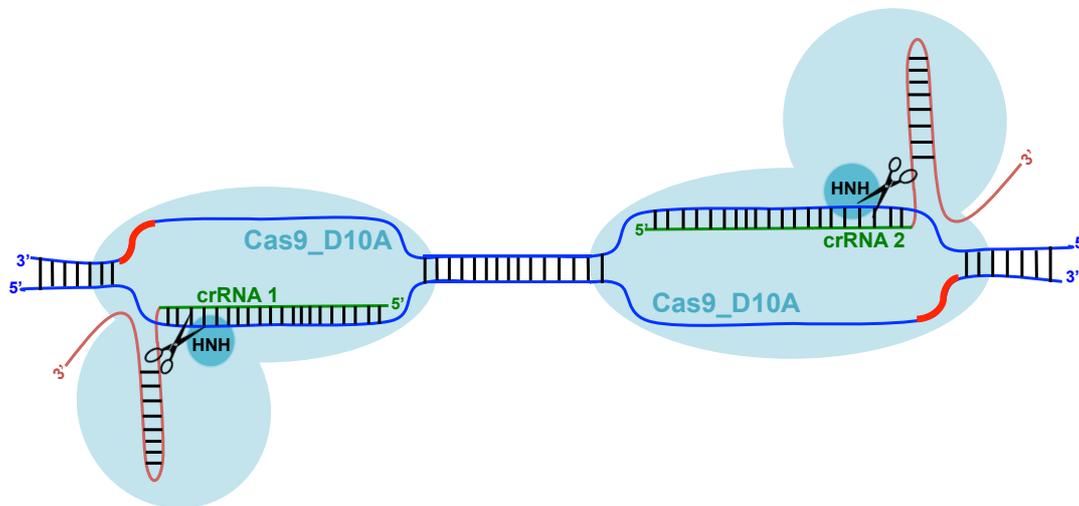
### **II. 2. 3. 4) Choix de la protéine Cas à utiliser**

Afin de limiter la présence de « off-targets » lors de la modification génétique ciblée, des équipes de recherches ont développé différentes méthodes.

#### **II. 2. 3. 4. 1) Cas9\_D10A**

En septembre 2013, une expérimentation a été faite en combinant une Cas9 ayant son site catalytique RuvC muté sur une paire de base (Cas9\_D10A) et deux sgRNA ciblant la région sur deux sites distincts. Cette combinaison induit deux cassures simple-brin (cassures réalisées par le domaine HNH de la Cas9\_D10A) sur les deux brins de l'ADN ([Ran et al, 2013](#)). Bien qu'il y ait deux sgRNA, donc potentiellement deux fois plus de risques d'avoir des « off-targets », la cellule réparant de façon fidèle les cassures simple-brin ([Dianov et Hübscher, 2013](#)), la probabilité que deux sites soient reconnus par la Cas9\_D10A dans la

même région génomique hors de la région ciblée est faible. Le risque de « off-targets » est donc 50 à 1500 fois inférieur que lors de l'utilisation de la nucléase (*Ran et al, 2013*) (Figure 7).



**Figure 7 : Système CRISPR-Cas9 pour effectuer une cassure double brins par la stratégie de la double nickases (Cas9\_D10A).**

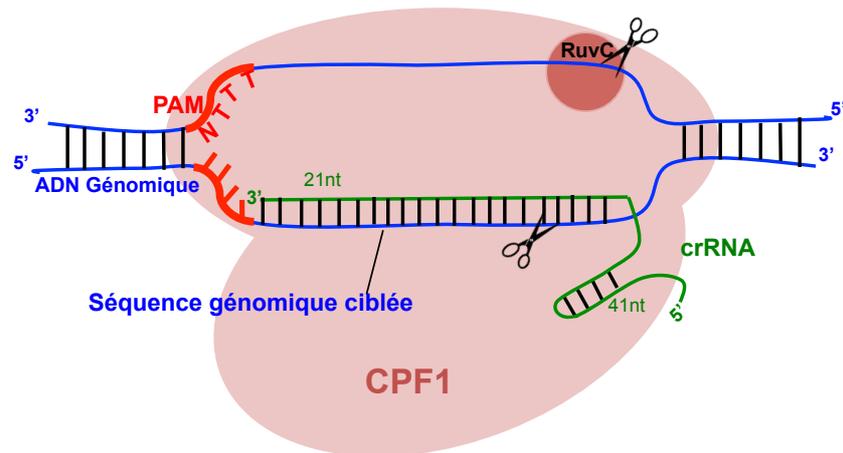
D'autres équipes ont expérimenté une nouvelle méthode pour effectuer de la correction du génome. Il s'agit d'allier la Cas9-D10A avec un sgRNA. La spécificité de la cible est améliorée et le nombre de « off-targets » est diminué. Cependant, du fait de la réparation fidèle de la cellule lors d'une coupure simple brin, l'efficacité de la modification génomique est diminuée (*Satomura et al, 2017 ; Gao et al, 2017*).

#### **II. 2. 3. 4. 2) CPF1 (CRISPR-associated endonuclease in *Prevotella* and *Francisella* 1)**

Le gène CPF1 a été découvert dans la région CRISPR de la bactérie *Francisella tularensis subsp.* Elle code pour des systèmes de défense fonctionnels. Contrairement aux systèmes Cas9, le système CRISPR contenant CPF1 possède trois caractéristiques. Premièrement, les répétitions CRISPR associées à la CPF1 sont transformées en crRNA matures sans l'exigence d'un tracrRNA (*Deltcheva et al, 2011*). Deuxièmement, le complexe CPF1-crRNA clive efficacement l'ADN ciblé grâce à une reconnaissance sur un site PAM de quatre nucléotides riches en T contrairement au site PAM du système Cas9 riche en G. Enfin, CPF1 introduit une coupure double brins de l'ADN avec 4 ou 5 nucléotides d'écart de manière cohésive loin du site PAM ce qui évite des délétions ou des insertions. La CPF1 possède le site catalytique

RuvC. Ce dernier clive l'ADN en formant une cassure double-brins. La mutation de ce site inactive son fonctionnement et donc le clivage des deux brins d'ADN (*Zetsche et al, 2015*).

Les régions riches en T sont plus rares que celles riches en G ce qui accroît la spécificité de la région génomique ciblée. De plus, le tracrRNA est absent du système CPF1 ce qui donne lieu à des sgRNA plus courts, plus rapidement matures et fonctionnels. La CPF1 peut donc être une bonne alternative au système CRISPR/Cas9 (*Figure 8*).



**Figure 8 : Système CRISPR pour effectuer une cassure double brins par la stratégie la nucléase CPF1**

#### **II. 2. 3. 4. 3) SpCas9-HF1**

La Cas9-HF1 est une Cas9 dont 4 acides aminés de la séquence (N497A, R661A, Q695, et Q926A) ont été modifiés. Ces modifications permettent à la Cas9 de reconnaître la région génomique ciblée et elles empêchent la Cas9 de cliver des « off-targets ». (*Kleinstiver et al, 2016*). Il n'y a pas assez de publications actuellement pour en dire plus, mais c'est très certainement une alternative de plus au système CRISPR standard.

#### **II. 2. 3. 4. 4) Utilisations des différentes méthodes.**

La Cas9 sauvage est de moins en moins utilisée, remplacée par les méthodes de clivage plus spécifiques afin de diminuer le risque de « off-targets ». La méthode à privilégier est le CRISPR/CPF1 car elle a une forte spécificité du fait de son site PAM 3'-NTTT. Cette méthode fonctionne aussi bien pour effectuer l'inactivation d'un gène (KO en anglais Knock-Out) que l'intégration ciblée d'un nouveau gène (KI en anglais Knock-In). Cependant, les séquences 3'-NTTT sont plus rares que les séquences 5'-NGG. Suivant l'organisme modèle, il est

possible de cibler une région dans le génome qui ne contient pas cette séquence. Dans ce cas là, les méthodes CRISPR/Cas9 seront privilégiées.

### **II. 2. 3. 5) Amélioration de la recombinaison homologue et diminution des « off-targets »**

En plus de développer des systèmes CRISPR/Cas de plus en plus performants, les chercheurs s'intéressent la recombinaison homologue après une cassure double brins par CRISPR lors d'une expérience CRISPR. Ils ont observé que le taux de recombinaison homologue après une cassure double-brins est dépendant de l'organisme modèle ainsi que des types cellulaires au sein d'une même espèce (*Kass et al, 2013*) (*Miyaoka et al, 2016*).

#### **II. 2. 3. 5. 1) SCR7 et RS-1**

Certaines équipes se sont penchées sur le l'inhibition de la réparation NHEJ qui est majoritaire lors d'une cassure double-brins. Le fonctionnement de NHEJ repose sur le recrutement de la ligase IV. Certaines molécules inhibent son activité. Comme le SCR7 pyrazine (SCR7) qui réduit le taux de NHEJ par 2 (*Chu et al, 2015*). Cependant l'action du SCR7 sur l'amélioration de taux de HDR est discutée. En effet, certains auteurs disent que le SCR7 améliore la HDR par un facteur 5. D'autres en revanche n'observent pas de différences significatives (*Song et al, 2016*). Ces derniers proposent la molécule RS-1 qui augmente la quantité de RAD51 dans la cellule. Celui-ci favorise l'entrée de la cellule dans le processus de HDR.

#### **II. 2. 3. 5. 2) Nocodazole et ABT-751**

En 2016, Yang et collaborateurs ont étudié l'importance de la phase du cycle cellulaire dans laquelle se trouvent les cellules souches pluripotentes induites humaines (hiPSC en anglais *human induced Pluripotent Cells*). Ils ont observé que les cellules non synchronisées ont un taux de HDR de 0,43% après une cassure double-brins induite par CRISPR/Cas9. Lorsque les cellules sont mises au contact du Nocodazole (inhibiteur de la prolifération des microtubules), elles sont synchronisées aux phases G2/M du cycle cellulaire et elles présentent alors un taux de HDR de 1,5%, soit 3,5 fois plus que les cellules non traitées. Un autre test est réalisé avec de l'ABT-751 qui a la même action que le Nocodazole. Les résultats sont similaires. Le taux de HDR dans les cellules est de 1,35% soit 3,1 fois plus que les cellules témoins (*Yang et al, 2016*).

#### **II. 2. 3. 5. 3) Anti-CRISPR**

Un autre facteur important dans la fréquence des « off-targets » est la durée de l'activité de la nucléase Cas9. Plus la Cas9 est fonctionnelle longtemps, plus le risque de voir la formation

de « off-target » est élevé. Pour palier cette difficulté, une équipe a développé la technique de la Cas9 inductible. Les cellules intègrent le gène de la Cas9 dont le promoteur est activé par la présence d'une molécule, par exemple la Doxycycline (Dox) (*González et al, 2014*). Lorsque les cellules sont cultivées en présence de Dox, la nucléase est synthétisée. Après 24h, la Dox est retirée et l'expression de la Cas9 est stoppée.

Un autre moyen de limiter la durée de l'activité de la Cas9 est d'injecter de l'ARN ou de la protéine purifiée dans les cellules (*Kim et al, 2014*). En effet la durée de vie d'un ARNm ou d'une protéine est très inférieure au plasmide. Mais la production de protéine Cas9 purifiée est coûteuse et sa conservation est compliquée. C'est pourquoi, cette méthode est peu utilisée.

Fin 2016, une publication est parue décrivant un système anti-CRISPR d'origine microbienne. Il s'agit d'un moyen pour les bactériophages de contrer la défense CRISPR des bactéries. Ces anti-CRISPR sont spécifiques des systèmes CRISPR-Cas de type I dont ne fait pas partie la Cas9 (Classe 2 type II). Pawluk et son équipe ont identifié trois familles anti-CRISPR qui inhibent spécifiquement le système CRISPR-Cas9 de *Neisseria meningitidis*. (*Pawluk et al, 2016*). Ils ont montré que lorsqu'ils intégraient sgRNA la NmeCas9 et le Acr (Anti-CRISPR), apparaissait une inhibition de la modification génomique. Des anti-CRISPR ont été récemment identifiés chez *Streptococcus pyogenes* (*Dong et al, 2017*).

## **II. 2. 4) Utilisation de la technologie CRISPR Cas9 dans les hESC**

Bien que la technologie CRISPR Cas9 soit relativement récente, de nombreuses données ont été publiées sur le sujet. Cette technologie est devenue en quelques années la référence dans le domaine de la modification génétique ciblée avec plus de 5000 publications ces trois dernières années. L'analyse, et surtout la comparaison du taux de HDR et donc de l'efficacité de la technologie, est impossible à faire sur l'ensemble des données publiées.

Je résume ici les résultats obtenus sur l'insertion d'un gène (KI) dans les hESC par la technique du CRISPR Cas9 et Cas9\_D10A.

Le résumé que je propose n'a pas pour objectif d'identifier une méthode idéale. Malgré le nombre raisonnable de publications réalisées sur les CRISPR dans les hESC, les données répertoriées ne sont pas ou peu comparables. Les conditions de cultures diffèrent entre deux laboratoires. De plus ce ne sont pas les mêmes régions génomiques qui sont ciblées.

Les résultats des articles montrant l'intégration d'une matrice de réparation après clivage de l'ADN par la technologie CRISPR/Cas9 dans des hESC sont résumés dans le tableau ci-dessous (*Zhu et al, 2015 ; Rong et al, 2014 ; Merkle et al, 2015*).

**Tableau 1 : Exemples de Knock-In dans des hESC à l'aide de la technologie CRISPR/Cas9.**

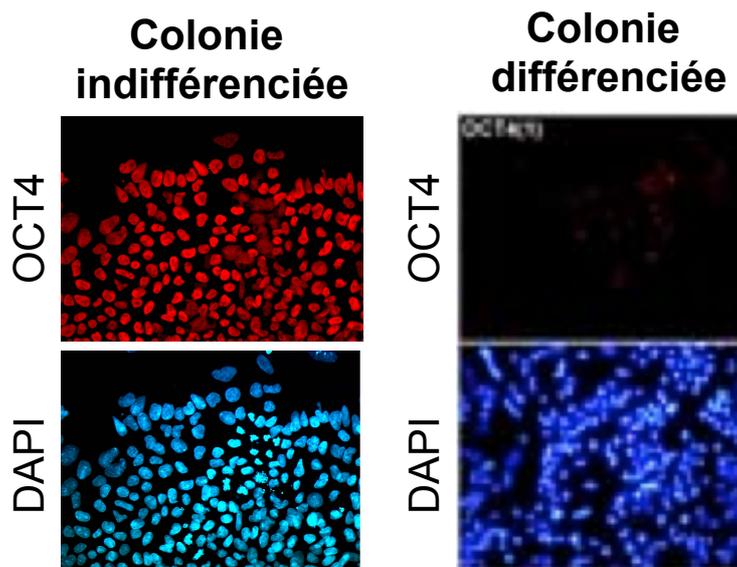
Gène ciblé	CRISPR	Gène	Outil	Clones positifs	Pourcentage de clones positifs	Efficacité Absolue
OCT4	Nucléase	eGFP-Puro	Gene-Pulser II	8/288	2,8%	$8.10^{-7}$
SFTPC	Nucléase	Puro	Gene-Pulser II	2/37	5%	$5.10^{-7}$
TRH-N	Double nickase	Néo	Neon	8/19	42%	$3.10^{-6}$
POMC-C	Double nickase	Néo	Neon	35/48	73%	$1.10^{-5}$
POMC-C	Simple nickase	Néo	Neon	0/18	0%	0%
CRH-C	Nucléase	Néo	Neon	28/30	93%	$1.10^{-5}$
CRH-C	Simple nickase	Néo	Neon	5/18	28%	$3.10^{-6}$

Nous observons que le taux de Knock-In dans les cellules hESC est très variable. Merkle et collaborateur en ont réalisé un nombre suffisant pour effectuer une analyse. Sur l'ensemble des gènes testés, 87% des clones résistants à l'antibiotique ont intégré le transgène dans le locus désiré après le clivage par la nucléase. Ce taux descend à 68% lorsqu'ils ont fait agir la double nickase. En ce qui concerne la simple nickase, le taux de KI est à 35%. (*Merkle et al, 2015*). Cependant, ils ne prennent pas en compte les « off-targets ».

# Projet EPHE

---

Le but de mon projet est de modifier génétiquement des cellules ES humaines afin de fabriquer des lignées marqueurs de l'état de pluripotence. Le gène *OCT4* est un gène clé de pluripotence des cellules ES. Il est fortement exprimé dans les cellules pluripotentes et son expression diminue lorsque les cellules se différencient (Figure 9). Il représente donc un bon candidat comme marqueur de l'état de pluripotence. J'ai entrepris de fabriquer une lignée de cellules hESC dans laquelle le gène codant pour la protéine fluorescente GFP (*Green Fluorescent Protein*) ou *Katushka 2* (*Kate2*) est intégré dans le génome de la cellule sous le contrôle transcriptionnel du gène *OCT4*. *OCT4* est exprimé spécifiquement dans les cellules pluripotentes (Figure 9).



**Figure 9 : Microphotographies de 2 colonies, l'une indifférenciée et l'autre différenciée, dans lesquelles sont observés l'expression du gène *OCT4* par immunofluorescence et le marquage des noyaux cellulaires au 4,6-diamidino-2-phenylindole (DAPI) (Images de la colonie différenciée tirée de l'article de Babaie et collaborateurs ; [Babaie et al, 2006](#)).**

# Matériels et Méthodes

---

## **I. Techniques de biologie cellulaire**

### **I. 1. Préparation des fibroblastes embryonnaires de souris (MEF)**

#### **I. 1. 1) Dérivation et congélation des MEF**

Les cellules souches embryonnaires humaines sont cultivées dans des boîtes gélatinées sur un tapis de cellules nourricières (MEFs). Afin d'utiliser le même lot de MEFs pendant plusieurs mois, nous réalisons un stock de cellules. Les MEFs sont dérivées à partir de fœtus de souris de la souche OF1 à 12,5 jours post-fécondation. Après dissection, les fœtus sont ensuite découpés en petits morceaux et dissociés par traitement enzymatique. Les cellules sont ensuite centrifugées et resuspendues dans du milieu MEF (Tableau 2), puis transférées dans des boîtes 100 mm (Corning 430167). Cette étape correspond au premier passage cellulaire (P1). Elle est réalisée par une personne ayant suivi la formation expérimentation animale. Le milieu est changé toutes les 24h. Trois jours plus tard, quand les cellules arrivent à confluence, toutes les cellules sont dissociées enzymatiquement (Trypsine/EDTA 1x (Gibco by Life 15400-054) pendant 5 minutes (min) à 37°C). Elles sont ensuite resuspendues dans 2 mL de milieu MEF. La boîte est rincée avec 3 mL de milieu MEF pour récupérer toutes les cellules, le contenu est transféré dans un tube de 15 mL. Celui-ci est centrifugé à 268 g pendant 5 min. Le culot de cellules est repris dans 4 mL de milieu MEF. Les cellules sont réensemencées au 1/4 dans de nouvelles boîtes. Cela correspond au passage P2. Lorsqu'elles sont de nouveau à confluence, elles sont trypsinées (P3) pour être congelées. Les MEFs sont numérotées, resuspendues dans du milieu MEF à une concentration finale de  $10^7$  cellules/mL. On ajoute ensuite au goutte à goutte un volume équivalent de milieu de congélation froid [80% FBS et 20% diméthylsulfoxyde (DMSO, sigma-aldrich D2650)]. La suspension cellulaire est ensuite aliquotée dans des cryotubes (Nunc) (1 mL par tube). Ces tubes sont congelés à -80°C dans une cryobox (Nalgène) puis transférés dans un container à azote liquide (-196°C) où on peut les conserver plusieurs mois.

#### **I. 1. 2) Décongélation des MEFs et traitement à la mycophénolate**

La préparation des cellules nourricières nécessite la décongélation d'un tube (ou plus) par semaine (selon les besoins). Après être sorti de l'azote liquide, le tube de MEFs est réchauffé

dans un bain Marie à 37°C. Nous ajoutons alors, goutte à goutte, 1 mL de milieu MEFs chaud. Nous récupérons les 2 mL obtenus dans un tube Falcon de 15 mL et nous diluons la suspension avec 10 mL. Après centrifugation, servant à éliminer le DMSO, les cellules sont resuspendues dans 10 à 30 mL de milieu MEFs et ensemencées dans 1 à 3 boîtes 100 mm (Corning). Les MEFs ainsi décongelées peuvent être amplifiées jusqu'au quatrième passage pour être utilisées comme cellules nourricières.

Les cellules hESC sont cultivées sur des MEFs inactivées pour qu'elles ne puissent pas se diviser et envahir les cellules d'intérêt. Pour cela, nous utilisons la mitomycine C (MMC, sigma-aldrich M0503) qui inhibe indirectement la synthèse de l'ADN en provoquant un pontage entre les deux brins complémentaires empêchant ainsi la réplication. Ce traitement empêche donc la division des MEFs sans les tuer, leur permettant ainsi de continuer à sécréter des facteurs de croissance. Pratiquement, le milieu des boîtes de 100 mm de MEFs à confluence est remplacé par 5 mL de MMC (5 µg/mL) pendant 2h à 37°C et 5% de CO<sub>2</sub>. Après incubation, les MEFs sont rincées 5 fois au PBS (Gibco – Life Technologie 14190). Elles sont ensuite trypsinées 5 min à 37°C et numérées à l'aide d'une cellule de Malassez. Les MEFs sont centrifugées 5 min à 268 g. Le culot cellulaire est repris dans du milieu MEF à 1.10<sup>6</sup> cellules/mL.

Avant d'ensemencer les MEFs mitomycinées, les boîtes sont traitées à la gélatine (sigma-aldrich G1393) 0,1% pendant 20 min à température ambiante. Nous ensemençons 1,8.10<sup>5</sup> MEF par boîte de 35 mm (Corning - 430165). Nous pouvons co-cultiver les hESC dans leur milieu adapté au minimum 12 h et jusqu'à 8 jours après l'ensemencement des MEF.

**Tableau 2 : Composition du milieu de culture des MEF**

<b>Milieu MEF</b>		<b>Composants (mL)</b>
<b>DMEM</b>		500
<b>FBS</b>	10%	50
<b>PSG 100X</b> (Pénicilline Streptomycine Glutamine)	1%	5,5
<b>NEAA</b> (Non Essential Amino Acid)	1%	5,5

## **I. 2. Culture des cellules ES humaines**

En raison des lois bioéthique, qui contrôlent la culture des cellules souches issues d'un embryon humain, les cellules hESC sont manipulées sous un poste de sécurité microbiologique de niveau II dans un laboratoire P2, au sein d'une enceinte dont l'accès est réglementé.

### **I. 2. 1) Préparation du bFGF**

Le bFGF est un facteur de croissance indispensable au maintien de l'autorenouveau des hESC. La préparation du bFGF se fait à partir d'une ampoule de 1mg de bFGF humain lyophilisé (Gibco - Life - PHG0263). Pour cela l'ampoule est centrifugée brièvement, le bFGF est dissous dans 100 mL de PBS plus BSA 0,1% (en anglais Bovin Serum Albumin (Gibco - Life- 15260037)) à une concentration de 10 mg/L. Des aliquots de 100 µL sont réalisés et conservés à -20°C. Lors de la culture des cellules hES, un aliquot de bFGF est décongelé, puis dilué au dixième dans 900 µL de PBS/BSA 0,1%. Le bFGF est utilisé à 5 µg/L, dans le milieu de culture.

### **I. 2. 2) Conditions de culture**

Les hESC sont cultivées à 37°C, dans une atmosphère à 5% de CO<sub>2</sub> et 5% d'O<sub>2</sub>, sur des MEF inactivées. Le milieu de base utilisé pour cultiver toutes les lignées étudiées est composé de knockout DMEM (Gibco – Life Technologies 10829-018) supplémenté avec 20% de Knock-Out Serum Replacement (KOSR) (Gibco – Life Technologies 10828-028), 1% d'une solution d'acides aminés non essentiels (NEAA, Gibco - Life Technologies 11140-035), 1% d'une solution de Pénicilline-Streptomycine-Glutamine à 29,2mg/mL (Gibco – Life Technologies 10378016) et 0.055 mM de β-mercaptoéthanol (Gibco – Life Technologies) Le milieu est supplémenté avec 5 µg/L de FGF. Les cellules hESC sont observées tous les jours pour voir l'évolution de la croissance des colonies ainsi que pour identifier une éventuelle contamination bactérienne ou fongique. Le milieu des cellules est changé tous les jours (2 mL de milieu par boîte). La densité cellulaire lors de l'ensemencement est importante pour maintenir l'état indifférencié des hESC. Une densité d'ensemencement trop faible augmente l'instabilité des cellules.

### **I. 2. 3) Passage cellulaire**

Les cellules hESC supportent mal la dissociation enzymatique. En effet, la répétition des passages cellulaires par cette méthode favorise l'apparition de mutations dans le génome. Mise à part des besoins ponctuels nécessitant une suspension unicellulaire (transfection, analyse par cytométrie en flux ou tri cellulaire), les cellules ES humaines sont repiquées

manuellement en amas cellulaires. Elles sont ensuiteensemencées dans des boites de 35 mm contenant  $1,8 \cdot 10^5$  de MEF inactivées. Le repiquage cellulaire consiste à gratter les colonies indifférenciées à l'aide d'un cône de 10  $\mu$ L pour décoller des amas cellulaires. Cette manipulation est réalisée 1 à 2 fois par semaine sous la PSMII à l'aide d'un microscope inversé au grossissement x40. Lors de l'amplification des cellules, une boîte-mère peut permettre d'ensemencer 2 à 4 boites-filles.

#### **I. 2. 4) Congélation et décongélation**

##### ***I. 2. 4. 1) Congélation***

Lors de la congélation, les cellules hES sont rincées avec 1 mL de PBS puis le PSB est retiré de la boîte. Il est alors ajouté 1 mL de cryostor (Bio Life Solution - CS10). Les cellules sont alors décollées de la boîte à l'aide d'un râteau (Biologix – 70-1180) et les cellules en suspension en amas dans le cryostor sont pipetées et placées dans un cryotube lequel est déposé dans une cryobox. Le tout est congelé à  $-80^{\circ}\text{C}$  pendant 24 à 48 heures. Le lendemain ou le surlendemain le cryotube est rangé dans un tank à azote liquide à  $-196^{\circ}\text{C}$ .

##### ***I. 2. 4. 2) Décongélation***

La décongélation se réalise comme celle des MEFs. Les cellules sontensemencées sur des MEFs fraîchement préparées dans le milieu ES. Environ 8 jours sont nécessaires pour que les cellules décongelées forment des colonies de taille suffisante pour pouvoir être repiquées pour la première fois.

#### **I. 2. 5) Préparation du milieu conditionné (MC)**

Le milieu conditionné (MC) est un milieu de culture enrichi en facteurs de croissance produits par les MEF. Pour produire le MC, les MEF sontensemencées dans une boîte 100 mm préalablement gélatiné à une densité de  $6,5 \cdot 10^4$  cellules/ $\text{cm}^2$ . Quand les MEF ont adhéré au fond de la boîte, le milieu MEF est remplacé par 38,5 mL de milieu ES supplémenté avec 150  $\mu$ L de bFGF. Le milieu est récupéré tous les jours et renouvelé toutes les 24 heures pendant une semaine. Il est filtré (Corning - 431097) et congelé à  $-20^{\circ}\text{C}$ . Le milieu MC est utilisé de manière ponctuelle lors de transfections par exemple lorsque les feeders MEF sont de mauvaise qualité ou absents.

### **I. 3. Transfection par électroporation**

La transfection permet d'intégrer un ou plusieurs plasmides dans une cellule par choc électrique. Deux appareils ont été utilisés, le Neon (thermo fisher scientific) et le Gene Pulser

II (Bio Rad). La préparation des cellules est la même pour les deux méthodes. Pour prévenir la différenciation, le milieu de culture des cellules hESC est remplacé par du MC, supplémenté avec 10 $\mu$ M de ROCKi (Rho-associated protein kinase inhibitor Y-27632, Calbiochem) et du bFGF 24h avant l'électroporation. Le jour de l'électroporation, le milieu des cellules est changé en présence de ROCKi et du bFGF. 5h après les boîtes sont rincées avec 1 mL de PBS 1X puis les cellules sont incubées 5 min avec 1 mL de trypsine EDTA 1X pour les dissocier en suspension unicellulaire. La suspension cellulaire est alors diluée dans du milieu ES pour inactiver la trypsine. Les cellules sont comptées, puis centrifugées pendant 5 min à 268 g. Le culot cellulaire est re-suspendu dans du PBS 1X. Le ou les plasmides sont ajoutés. La suspension cellules/plasmides est électroporée en suivant les protocoles des fabricants ([Tableau 3](#)). Les cellules transfectées sont ensuiteensemencées dans des boîtes 35 mm contenant chacune 180.000 MEF résistantes à la puromycine et 2 mL de MC supplémenté avec du ROCKi et du bFGF.

**Tableau 3 : Conditions d'électroporation de Gene Pulser II et de Neon**

Appareils	Bio Rad Gene Pulser II	Thermo Neon
Nombres de cellules	10.10 <sup>6</sup>	1.10 <sup>6</sup>
Quantité ADN	50 $\mu$ g	5 $\mu$ g
Cuves	Cuve de 0,4cm (800 $\mu$ L)	Cône Neon (100 $\mu$ L)
Tampons d'électroporations	PBS 1X	Buffer R
Paramétrage	500 $\mu$ F et 250V	2 pulses, 30 ms, 1050V

#### **I. 4. Sélection des clones cellulaires**

La sélection des clones permet d'éliminer les cellules n'ayant pas intégré la matrice de réparation lors de la transfection. Cette matrice contient un gène de résistance à la puromycine. La sélection commence 48h après la transfection. La puromycine est ajoutée dans le milieu ES à une concentration finale de 1  $\mu$ g/ $\mu$ L. Le traitement antibiotique dure une semaine. Les clones cellulaires résistants à la puromycine (puro) sont repiqués et placés individuellement dans un puits de plaque 24-puits, préalablementensemencé avec 40.000 MEF. Lorsque ces clones atteignent une taille suffisante, ils sont repiqués dans une plaque 12-puits préalablement préparés avec 86.000 MEF. Les clones sont repiqués et les cellules sont réensemencées dans une boîte 35 mm, et amplifiés pour une analyse PCR.

## I. 5. Culot sec

Les culots secs cellulaires sont utilisés pour des analyses ultérieures comme l'analyse génomique. Cette technique permet de conserver les cellules sans dégrader leur ADN. Les cellules hESC sont rincées deux fois avec 1 mL de PBS. Les cellules sont décollées à l'aide d'un râteau. Les cellules en suspension en amas sont centrifugées 5 min à 268 g. Enfin le surnageant est retiré et le culot de cellules est congelé à -80°C.

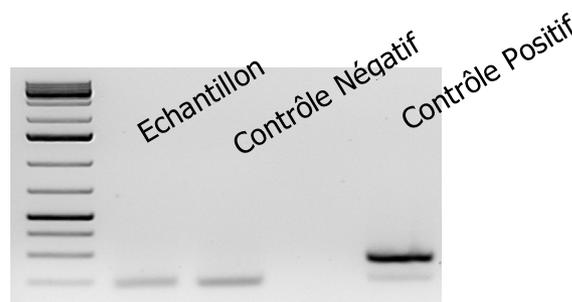
## II. Techniques de biologie moléculaire

### II. 1. Test PCR contamination Mycoplasme

Les mycoplasmes sont des bactéries sans paroi cellulaire. Ils sont présents en grande quantité dans l'environnement. Ce sont des contaminants des cultures cellulaires. Pour les identifier un test PCR est utilisé (Venor® GeM One Step Mycoplasma Detection kit - 11-8100). Pour effectuer ce test, 1 mL de surnageant cellulaire non renouvelé pendant 5 jours est prélevé dans un tube de 1,5 mL (Eppendorf). Ensuite l'échantillon est chauffé à 95°C pendant 5 min, et centrifugé 5 sec à vitesse maximale. Préalablement, 23 µL de OneStep Mix ont été aliquotés dans des tubes PCR. Le témoin négatif est réalisé avec 2 µL de PBS (Contrôle Négatif) ajouté au mix, 2 µL de « Positive Control DNA » (Contrôle Positif) complète le mix PCR du contrôle positif. Et 2 µL de surnageant cellulaire (Echantillon) sont ajoutés au dernier mix PCR. Les tubes PCR sont brièvement centrifugés.

Le programme d'amplification est le suivant : une dénaturation de 2 min à 94°C, puis 39 cycles comprenant une dénaturation de 30 sec à 94°C, une hybridation de 30 sec à 55°C et une élongation de 30 sec à 72°C et une élongation finale de 5 min à 72°C.

Ensuite 25 µL d'échantillon + 5 µL de Gel loading dye 6x sont analysés par électrophorèse sur un gel d'agarose à 1,5%, (20 à 40 min à 100V) (Figure 10).



**Figure 10 : Gel d'agarose du profil type d'un test Mycoplasme**

## **II. 2. Purification des plasmides par la méthode du PEG/LiCl**

Avant d'électroporer un plasmide dans des cellules eucaryotes, celui-ci est amplifié dans une bactérie compétente puis purifié.

### **II. 2. 1) Amplification plasmidique**

Les bactéries contenant le plasmide d'intérêt sont repiquées à partir d'une boîte de gélose LB (Lysogeny Broth) Agar (Sigma) + antibiotique dans 3 mL de LB avec de l'antibiotique. Elles sont incubées pendant 8 heures à 37°C sous agitation à 225 rpm. Les bactéries sont alors transvasées dans 250 mL de milieu LB avec de l'antibiotique. Elles sont cultivées à 37°C sous agitation pendant 12 à 14 heures.

### **II. 2. 2) Extraction plasmidique**

La culture bactérienne est centrifugée 15 min à 4.000 g à 4°C dans un pot à centrifugation de 250 mL. Le surnageant est éliminé et décontaminé à l'eau de javel. Le culot cellulaire est resuspendu dans 8 mL de GTE (Glucose Tris-HCl EDTA) contenant 200 µL de lysozyme (50 mg/mL), puis incubé 20 min afin de permettre la lyse de la paroi des bactéries. On ajoute alors 16 mL de NaOH/SDS (solution de lyse). Le pot est secoué à 5 reprises et laissé à incuber 5 min pour dégrader l'ADN bactérien. Douze mL de KAc/HCOOH (solution de neutralisation) sont rajoutés afin de précipiter l'ADN bactérien.

Le mélange est transvasé dans un tube de 50 mL (Falcon) et centrifugé 40 min à 4500 g. La solution située entre la pastille blanche supérieure et le culot est récupérée dans un autre tube 50 mL. Du PEG 40% (Sigma - Polyéthylène glycol 6000) est ajouté représentant un quart du volume final. L'ensemble est placé 1 h dans la glace afin de précipiter les acides nucléiques. Le tube est centrifugé 1h à 4.500 g à 4°C, puis le surnageant est éliminé. Le culot est solubilisé dans 900 µL de 5T.1E (5 mM Tris (pH7,5) / 0,1 mM EDTA) pendant 12-14 heures.

Le lendemain, les ARN sont précipités à l'aide de 500 µL de NH<sub>4</sub>Ac 7,5M pendant 15 minutes dans la glace. Les ARN ainsi précipités sont centrifugés 10 min à 17.000 g à 4°C. Le surnageant est transféré dans un tube de 2 mL auquel sont ajoutés 10 µL de RNase à 10 mg/mL, puis incubé 30 min à 37°C. On ajoute enfin 5 µL de Protéinase K à 20 mg/mL et on incube la solution à 56°C pendant 1h afin de dégrader les protéines. Les protéines résiduelles sont ensuite précipitées à l'aide de Phénol-CHCl<sub>3</sub> (1 volume de Phénol-CHCl<sub>3</sub> pour 1 volume de solution de plasmides). Après mélange au vortex pendant 1 min pour mélanger les 2 phases, le tube est centrifugé 5 min à 17.000g. La phase aqueuse (supérieure) est

récupérée et transvasée dans un autre tube. Si la phase aqueuse est trouble, il faut recommencer l'opération.

Pour précipiter l'ADN, 700 µL d'isopropanol sont ajoutés. Le tube est retourné à cinq reprises pour homogénéiser la solution. Puis il est incubé 1 h à -20°C, puis l'ADN plasmidique est centrifugé 20 min à 17.000 g à 4°C. Le surnageant est éliminé et le culot est solubilisé dans 500 µL d'H<sub>2</sub>O ultra-pure

### **II. 2. 3) Purification**

Pour obtenir un ADN super-enroulé pour les transfections, une précipitation au PEG/LiCl est réalisée. Pour ce faire, 180 µL de PEG/LiCl sont ajoutés, la solution est mélangée à l'aide du vortex et incubée 20 min à température ambiante. Le tube est centrifugé 15 min à 17.000g à 4°C. Le surnageant est éliminé et le culot est rincé 3 fois avec 1 mL d'éthanol à 70%. Il est séché brièvement puis resuspendu dans de l'H<sub>2</sub>O ultra pure pendant 12-14 heures à 4°C. Le plasmide est dosé avec le Nanodrop (Thermo Scientific).

## **II. 3. Fabrication des plasmides CRISPR**

La méthode utilisée pour réaliser une cassure double brin dans le génome est celle de la « double nickase ». Pour la concevoir, il est nécessaire d'avoir deux ARN pour guider la "nickase" sur les régions ciblées.

### **II. 3. 1) Design des sgRNA**

La conception des sgRNA est réalisée grâce à des algorithmes développés par des laboratoires, disponibles sur internet. Le premier a été conçu dans le laboratoire de Feng Zhang (<http://crispr.mit.edu>). D'autres ont depuis été développés, comme CHOPCHOP (<http://chopchop.rc.fas.harvard.edu>) ou CRISPOR (<http://crispor.tefor.net>). Ces algorithmes calculent les meilleures cibles pour un crRNA dans une région génomique. Ils prennent en compte les "off-targets" théoriques en fonction du nombre de disparités crRNA/ADN. L'algorithme qui a été utilisé pour concevoir les sgRNA est celui développé par Zhang.

### **II. 3. 2) Construction du plasmide CRISPR (Sakuma et al, 2014)**

Les amorces sont hybridées et phosphorylées dans du tampon de ligation [New England Biolabs (NEB)] (60 min à 37°C, 5 min à 95°C, puis refroidissement jusqu'à 25° à -0.1°C/sec). Les plasmides pX330A\_D10A (addgene - plasmid n° 48772) et pX330S (addgene Plasmide n° 58778) sont linéarisés avec l'enzyme de restriction BbsI. Pour éviter que le plasmide ne se re-circularise à la suite de la digestion, le plasmide est déphosphorylé par la phosphatase CIP (NEB). Les oligonucléotides hybridés et phosphorylés sont insérés dans les plasmides

préalablement purifiés avec la T4 DNA Ligase (NEB). La réaction est incubée 20 min à température ambiante.

Le produit de ligation est introduit dans des bactéries DH5α par transformation bactérienne, l'incubation dure 30 min dans de la glace. Puis un choc thermique est réalisé à 42°C pendant 45 sec. Les bactéries sont ensuite mises en culture sur du milieu solide LB agar+ Ampicilline (Sigma) à 100 mM pour les plasmides pX330A\_D10A et sur du milieu solide LB agar + Spectinomycine (Sigma) à 50 mM pour les plasmides pX330S. Les bactéries sont incubées à 37°C pendant la nuit.

Pour chaque transformation, cinq à six clones bactériens sont analysés par PCR avec une amorce sens correspondant au guide et une amorce CBA-REV (Annexe 1) présente dans le vecteur. Les clones sélectionnés sont repiqués sur une gélose LB agar + antibiotique pour effectuer une sauvegarde bactérienne, le reste de la colonie bactérienne est déposé dans un mix PCR pour amplification.

Pour chaque clone positif, les plasmides sont amplifiés par miniculture. Ils sont par la suite purifiés à l'aide du « GeneJET Plasmid Miniprep Kit » (Thermo). La séquence des plasmides est contrôlée par séquençage à l'aide des amorces précédemment citées et pBR322-For (Annexe 1) cela permet de contrôler l'intégration des guides dans le vecteur. (Eurofins genomics). Les résultats de séquençages sont analysés à l'aide du logiciel SerialCloner®.

Par la suite, chaque couple d'ARNg est cloné dans le même plasmide. Le sgRNA présent dans le plasmide pX330S intègre le plasmide pX330\_D10A. Ce clonage est réalisé grâce au système "golden gate assembly" (NEB) en suivant le protocole du fabricant. Le produit de ligation est ensuite transformé dans des bactéries DH5α par choc thermique à 42°C pendant 45 secondes. Les bactéries sont ensuite mises en culture sur du milieu solide LB + ampicilline. Pour chaque transformation, cinq à six clones sont analysés par PCR avec une amorce correspondant au guide, et CBA-Rev. Les clones sélectionnés sont repiqués sur une gélose LB + ampicilline pour effectuer une sauvegarde, et introduit dans le mix PCR comprenant les amorces du guide et de CBA-REV. Pour chaque clone positif, les plasmides sont extraits de maxiculture par la technique PEG/LiCl. L'insertion du couple de sgRNA dans le pX330\_D10A est enfin contrôlée par séquençage, et analysé avec SerialCloner®.

## **II. 4. Extraction ADN génomique**

L'extraction d'ADN génomique est réalisée à partir de culot cellulaire. Elle est réalisée avec le QIAamp DNA Mini Kit (Qiagen). Tout d'abord, le tube de cellules est décongelé dans la glace puis re-suspendu dans 200 µL de PBS. Ensuite 20 µL de « QIAGEN Protease » et 200 µL de Buffer AL sont ajoutés. Le tube est incubé à 56°C pendant 10 min, puis centrifugé brièvement pour faire tomber la condensation du tube. 200 µL d'éthanol 100% sont ajoutés. La solution est alors mélangée avec le vortex pendant 15 sec et centrifugée brièvement. Le lysat cellulaire est déposé sur la colonne « QIAamp Spin Column ». Lors de la centrifugation d'1 min à 8.000g, l'ADN est retenu dans la colonne tandis que les débris cellulaires sont élués. Pour rincer la colonne, 500 µL de « Buffer AW1 » sont ajoutés. Puis la colonne est centrifugée 1 min à 8.000g. 500µL de « Buffer AW2 » sont pipetés dans la colonne. Le tout est centrifugé 3 min à 17.000g. La centrifugation est répétée à vide 1 min à 17.000g pour éluer tout l'éthanol contenu dans les deux tampons de rinçages.

Enfin, la colonne est placée dans un tube de 1,5 mL.30 à 200 µL d'eau ultra pure sont ajoutés. Il faut alors attendre 1 à 5 min que toute la colonne soit mouillée. L'ADN est élué avec la centrifugation du tube 1min à 8.000g, puis dosé par spectrophotométrie à l'aide du nanodrop et dilué pour obtenir une solution d'ADN à 50 ng/µL.

# Résultats

---

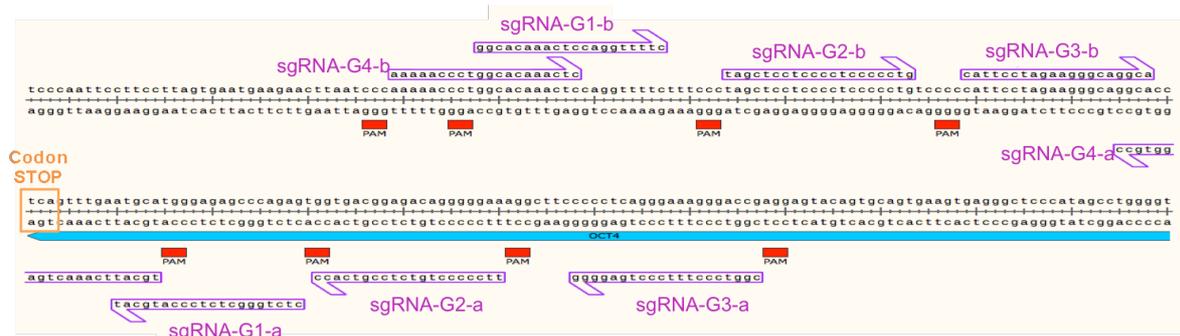
## **I. Outils nécessaires à la mise en place de la technologie CRISPR/Cas9**

Le système CRISPR/Cas9 n'est pas présent dans le génome des eucaryotes. Si on souhaite modifier le génome d'un eucaryote par ce système, il est donc nécessaire de l'y amener, c'est à dire d'apporter l'ARN guide (spécifique de la région ciblée), la protéine Cas 9 et la matrice de réparation (ADN homologue) selon le but recherché. Des plasmides permettent l'expression de ces différents éléments du système CRISPR/Cas9.

### **I. 1. Construction des plasmides d'expression des guides et de la Cas9**

#### **I. 1. 1) Design des guides (sgRNA)**

Mon projet a pour but d'insérer un gène codant pour une protéine fluorescente à la place du codon STOP du gène *OCT4* dans les cellules ES humaines. La première étape consiste donc à concevoir les guides permettant le recrutement de la Cas9 dans cette région. Dans ce but, j'ai d'abord analysé une séquence de 250 pb entourant le codon STOP du gène *OCT4*, grâce à l'algorithme CRISPR design (<http://crispr.mit.edu>). Etant donné que je vais utiliser la Cas9 D10A, j'ai spécifié le critère sgRNA permettant une double cassure simple brin. Parmi la liste de sgRNA proposée par l'algorithme CRISPR design, j'ai sélectionné 4 couples de guides (G1, G2, G3 et G4) ayant un score élevé. Ce score reflète l'efficacité de coupure par la Cas9 ainsi que la spécificité. Un score élevé signifie une faible probabilité de coupure non-spécifique ou "off-targets". Les couples de guides G1, G2, G3 et G4 encadrent le codon STOP du gène *OCT4* (Figure 11). L'étape suivante consiste à construire les plasmides d'expression des couples de sgRNA.



**Figure 11 : positionnement des 8 sgRNA conçus autour du codon STOP du gène *OCT4***

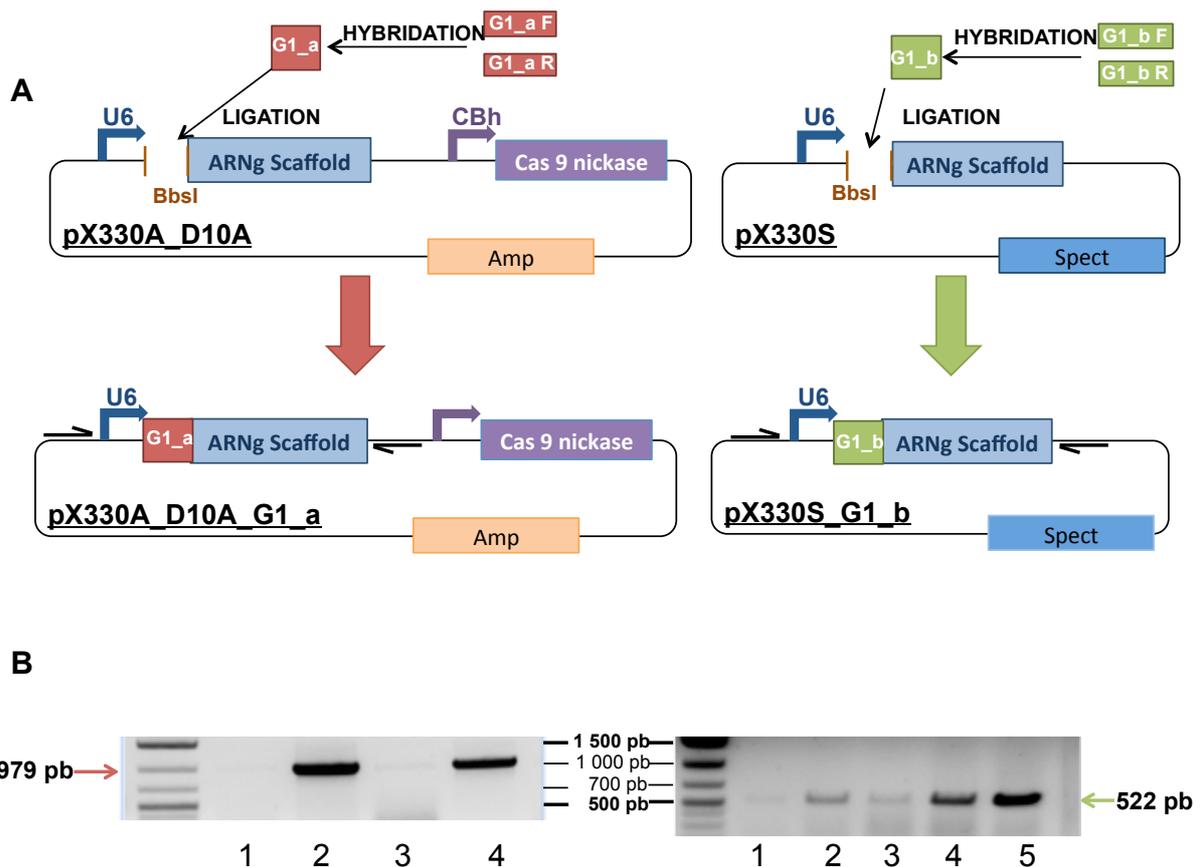
### **I. 1. 2) Construction des plasmides pX330A\_D10A\_sgRNA et pX330S\_sgRNA**

Chaque couple de guide (G1, G2, G3 et G4) est cloné dans un plasmide permettant également l'expression de la Cas9\_D10A (*Sakuma et al, 2014*). Cette construction se fait en deux étapes. Dans un premier temps, pour chaque couple, un des guides (ou sgRNA) est cloné dans le plasmide pX330A\_D10A et l'autre dans le plasmide pX330S2 (*Figure 12A*). Le plasmide pX330A\_D10A contient le promoteur U6 en amont d'un site pour l'enzyme de restriction BbsI et du tracrRNA permettant le recrutement de la Cas9\_D10\_A. Ce plasmide code également pour la Cas9\_D10A sous contrôle du promoteur CAG (CBh) et pour le gène de résistance à l'ampicilline. Le plasmide pX330S contient le promoteur U6 en amont d'un site pour l'enzyme de restriction BbsI et du tracrRNA permettant le recrutement de la Cas9\_D10\_A et le gène de résistance à la spectinomycine. Le promoteur U6 permet l'expression du guide composé des séquences sgRNA et tracrRNA. La séquence du sgRNAa est insérée dans le site BbsI du plasmide pX330A\_D10A et la séquence du sgRNAb dans le site BbsI du plasmide pX330S2, afin de générer les plasmides pX330A\_D10A\_sgRNAa et pX330S\_sgRNAb (*Figure 12A*). La séquence des sgRNA et la séquence des oligonucléotides correspondant sont représentées dans le *Tableau 4*.

**Tableau 4 :** (A) séquence des sgRNA. (B) Oligonucléotides sens et anti-sens contenant la séquence du sgRNA (minuscules) et la séquence nécessaire au clonage dans le site BsbI (majuscules). (C) Oligonucléotides double-brin obtenus après hybridation des oligonucléotides sens et anti-sens, insérés dans le site BsbI des plasmides pX330A\_D10A et pX330S2.

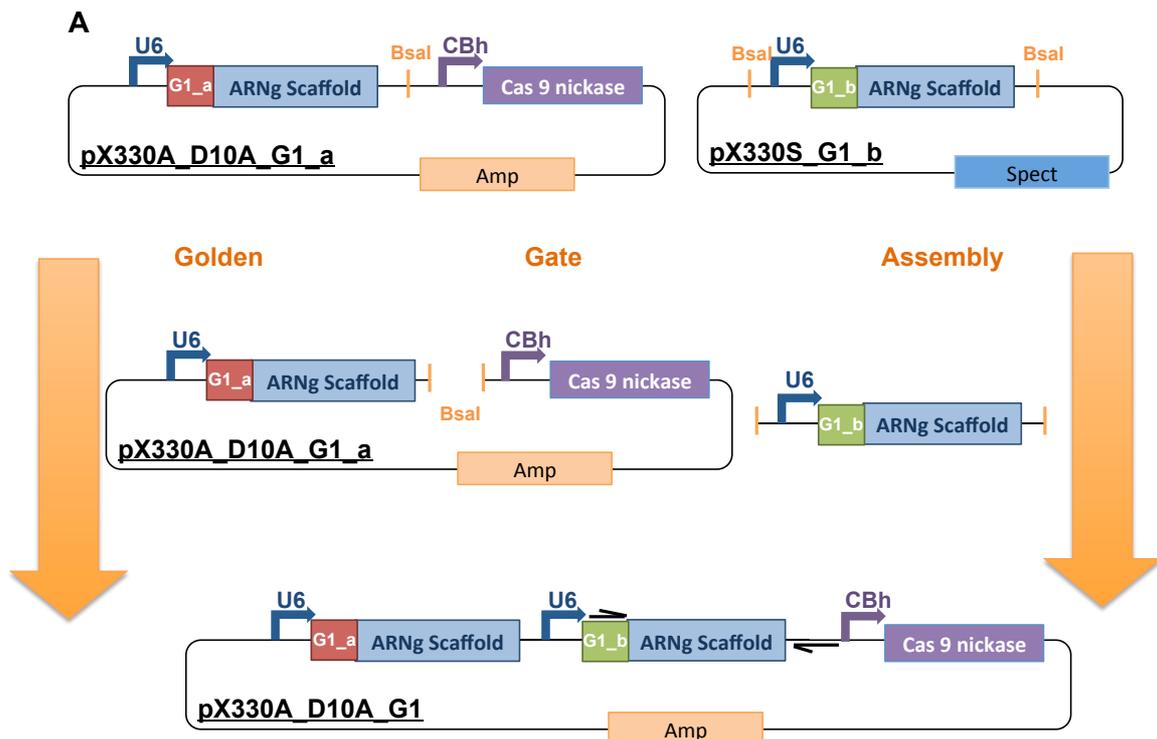
A : Séquence des sgRNA	B : Amorces synthétisées	C : Oligonucléotides double-brin
<b>G1a_</b> 5'-atgcatgggagagcccagag-3'	F: CACCGatgcatgggagagcccagag R: AAACctctgggctctccatgcatC	<pre>CACCGatgcatgggagagcccagag----           ----GatgcatgggagagcccagagGTTT</pre>
<b>G1b_</b> 5'-gaaaacctggagtttgcc-3'	F: CACCGgaaaacctggagtttgcc R: AAACggcacaaactccaggtttcC	<pre>CACCGgaaaacctggagtttgcc----           ----GgaaaacctggagtttgccGTTT</pre>
<b>G2a_</b> 5'-ggtgacggagacaggggaa-3'	F: CACCGggtgacggagacaggggaa R: AAACttccccctgtctcgtcaccC	<pre>CACCGggtgacggagacaggggaa----           ----GggtgacggagacaggggaaGTTT</pre>
<b>G2b_</b> 5'-cagggggagggaggagcta-3'	F: CACCGcagggggagggaggagcta R: AAACtagctctccccctccctgC	<pre>CACCGcagggggagggaggagcta----           ----GcagggggagggaggagctaGTTT</pre>
<b>G3a_</b> 5'-cccctcagggaaagggaccg-3'	F: CACCGcccctcagggaaagggaccg R: AAACcggctcctttcctgaggggC	<pre>CACCGcccctcagggaaagggaccg----           ----GcccctcagggaaagggaccgGTTT</pre>
<b>G3b_</b> 5'-tgctgccttctaggaatg-3'	F: CACCGtgctgccttctaggaatg R: AAACcattcctagaagggcaggcaC	<pre>CACCGtgctgccttctaggaatg----           ----GtgctgccttctaggaatgGTTT</pre>
<b>G4a_</b> 5'-ggcacctcagtttgaatgca-3'	F: CACCGggcacctcagtttgaatgca R: AAACtgcatcactgaggtgccC	<pre>CACCGggcacctcagtttgaatgca----           ----GggcacctcagtttgaatgcaGTTT</pre>
<b>G4b_</b> 5'-gagtttgccagggtttt-3'	F: CACCGgagtttgccagggtttt R: AAACaaaacctggcacaactcC	<pre>CACCGgagtttgccagggtttt----           ----GgagtttgccagggttttGTTT</pre>

L'insertion des oligonucléotides double-brin dans les plasmides pX330A\_D10A et pX330S2 est réalisée par ligation grâce à la T4 DNA Ligase. Le produit de ligation est ensuite utilisé pour transformer des bactéries DH5a. Après mise en culture de plusieurs clones, et extraction de l'ADN plasmidique, l'insertion des oligonucléotides est vérifiée par PCR dans 4 ou 5 clones. La présence d'un produit PCR d'une taille de 979 pb confirme l'insertion des oligonucléotides dans le plasmide pX330A\_D10A et une bande de 522 pb confirme l'insertion de l'oligonucléotide dans le plasmide pX330S (Figure 2B). Le résultat est confirmé par séquençage. Les plasmides ainsi obtenus sont nommés pX330A\_D10A-G1\_a et pX330S-G1\_b, dans le cas du couple de guide G1.

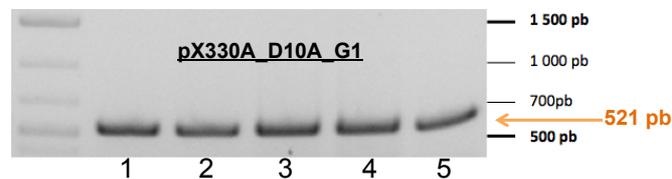


**Figure 12 : Construction des plasmides pX330A\_D10A-G1\_a et pX330S2-G1\_b. (A)** Hybridation des amorces codant pour l'ARNg, puis ligation des oligonucléotides double-brun dans les plasmides pX330A\_D10A et pX330S2. **(B)** Analyse par PCR des minipréparations des plasmides pX330A\_D10A\_G1a et pX330S2\_G1b puis électrophorèse sur gel d'agarose.

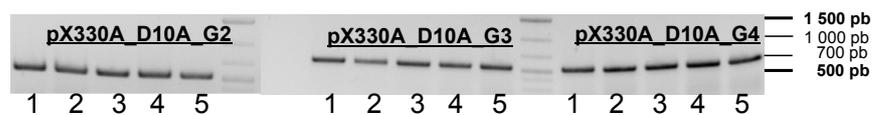
La deuxième étape consiste à insérer la cassette promoteur U6-Guide1b (G1\_b) du plasmide pX330S-G1\_b, dans le plasmide pX330A\_D10A-G1\_a, grâce à la technologie "Golden Gate Assembly" (NEB) (Figure 13A). Après digestion par l'enzyme de restriction BsaI, le fragment d'ADN contenant la cassette promoteur U6-Guide1b (G1\_b) est inséré par ligation grâce à la T4 DNA ligase, dans le plasmide pX330A\_D10A-G1\_a digéré par l'enzyme BsaI. Le produit de ligation est ensuite utilisé pour transformer des bactéries DH5α. Après mise en culture de plusieurs clones bactériens, et extraction de l'ADN plasmidique, l'insertion de la cassette promoteur U6-Guide1b (G1\_b) est vérifiée par PCR dans 4 ou 5 clones. La présence d'un produit PCR d'une taille de 521 pb confirme l'insertion de la cassette promoteur U6-Guide1b (G1\_b) dans le plasmide pX330A\_D10A-G1\_a (Figure 13B). Le résultat est confirmé par séquençage. Le plasmide ainsi obtenu est nommé pX330A\_D10A-G1. De la même manière, j'ai construit les plasmides pX330A\_D10A-G2, pX330A\_D10A-G3 et pX330A\_D10A-G4 (Figure 13C).



**B**



**C**



**Figure 13 : Construction des plasmide pX330A\_D10A-G1, plasmide pX330A\_D10A-G2, plasmide pX330A\_D10A-G3 et plasmide pX330A\_D10A-G4. (A)** Insertion du fragment BsaI contenant la cassette promoteur U6-Guide1b (G1\_b) dans le plasmide pX330A\_D10A-G1\_a grâce à la technologie Golden Gate assembly. **(B)** Analyse par PCR des minipréparations des plasmides pX330A\_D10A\_G1 puis électrophorèse sur gel d'agarose. **(C)** Analyse par PCR des minipréparations des plasmides pX330A\_D10A\_G2, pX330A\_D10A\_G3 et pX330A\_D10A\_G4 puis électrophorèse sur gel d'agarose.

## I. 2. Construction de la matrice de réparation

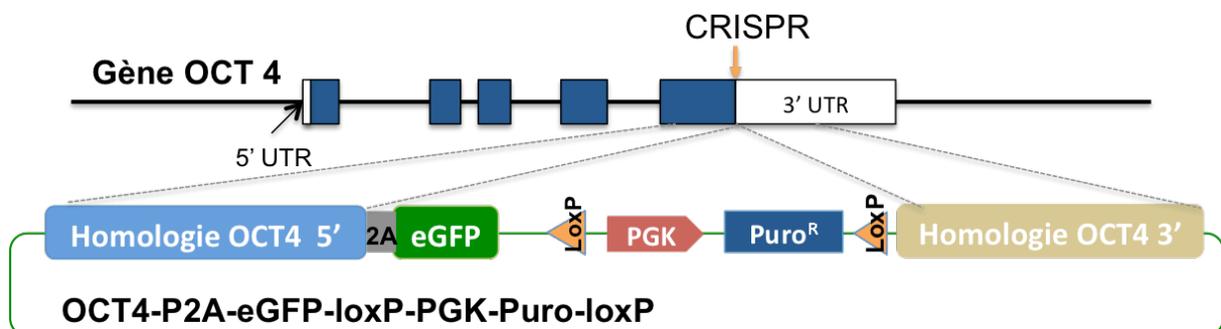
Le but de mon projet est de fabriquer une lignée de cellules ES humaines, rapportrice pour l'expression du gène *OCT4*. Pour cela, je vais utiliser la matrice de réparation conçue par Hockemeyer et Coll (*Hockemeyer et al, 2011*) (Figure 14). Cette matrice contient :

- deux bras d'homologies de 750pb situés de part et d'autre du codon STOP du gène *OCT4*.

- le gène de résistance à la Puromycine sous contrôle du promoteur pGK, le tout entouré de deux sites loxP. Il sera possible de sélectionner les cellules dans lesquelles il y aura eu recombinaison homologue, en présence de puromycine. La cassette pGK-Puromycine pourra ensuite être éliminée par transfection d'un plasmide d'expression de la recombinase Cre, qui induit une recombinaison entre les deux sites loxP.

- une séquence permettant l'expression de la protéine fluorescente GFP fusionnée à un peptide de clivage (P2A) dans sa partie N-Terminale. Après recombinaison homologue, les cellules exprimeront une protéine de fusion Oct4-P2A-GFP, qui sera ensuite clivée au niveau du peptide P2A.

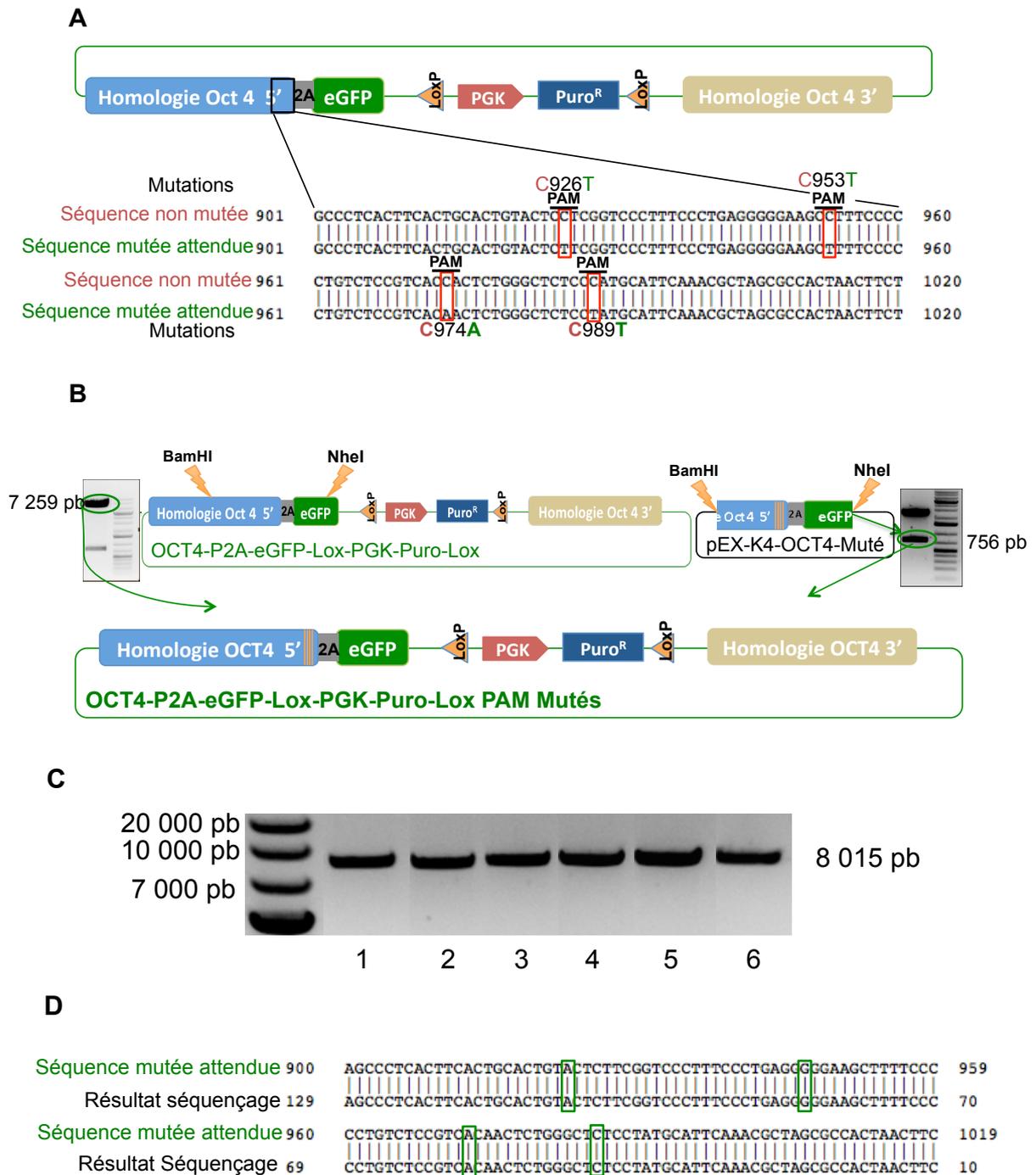
Cette matrice est contenue dans le plasmide pOCT4-eGFP-2A-Puro, disponible chez Addgene (plasmide n°31939).



**Figure 14 : Matrice de réparation permettant d'insérer le gène codant pour la GFP à la place du codon STOP du gène *OCT4*, dans les cellules hESC.**

### **I. 2. 1) Mutation des sites PAM dans la matrice de réparation**

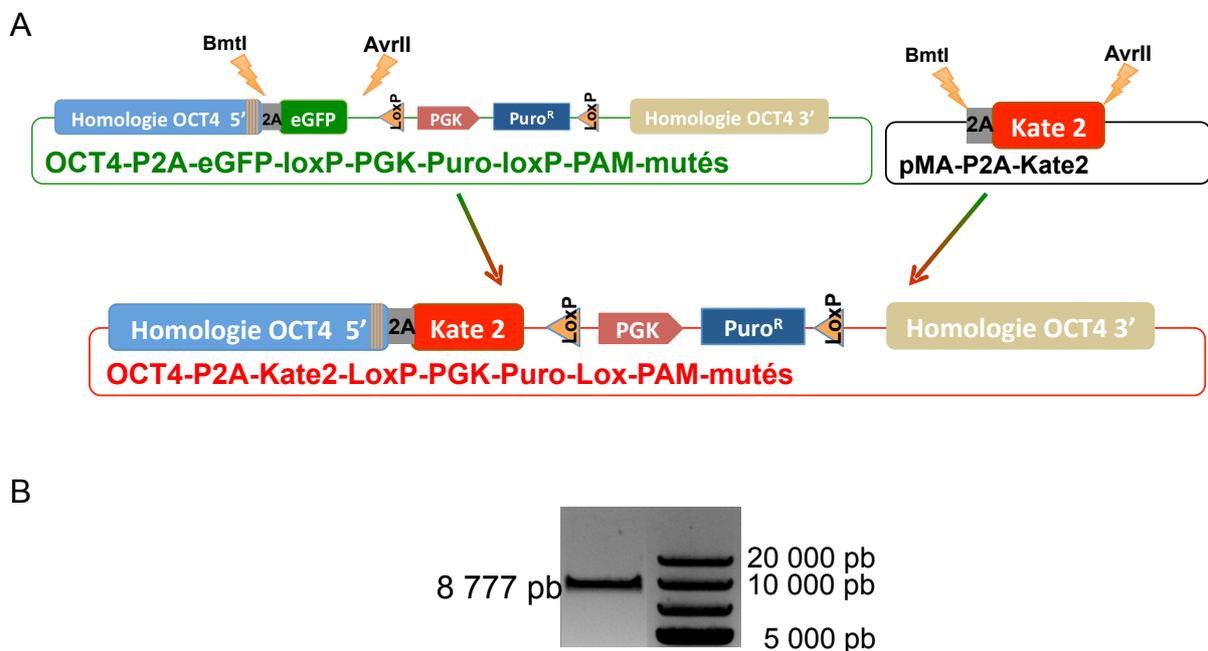
Les bras d'homologie de cette matrice contiennent les sites PAM correspondants aux guides que je j'ai conçu et cloné dans le plasmide d'expression pX330A\_D10A. Les sites PAM permettent la fixation de la Cas 9 et donc la coupure de l'ADN où se trouvent ces sites PAM. Avant de pouvoir utiliser cette matrice de réparation, je vais muter les sites PAM présents dans les bras d'homologie, de manière à éviter qu'elle soit coupée par la Cas9. Les mutations à réaliser sont les suivantes : C → T en position 926, C → T en position 953, C → A en position 974 et C → T en position 989 dans le plasmide pOCT4-eGFP-2A-Puro (Figure 15A). Ces quatre bases se situent dans un fragment de 756pb, encadré par un site de coupure pour l'enzyme de restriction BamHI et un site pour l'enzyme NheI, dans le plasmide pOCT4-eGFP-2A-Puro. J'ai fait synthétiser ce fragment avec les quatre mutations (Eurofins Genomics), que nous avons reçu sous la forme du plasmide pEX-K4-OCT4-Muté. Après digestion du plasmide pEX-K4-OCT4-Muté par les enzymes BamHI et NheI, la séquence contenant les sites PAM mutés a été insérée par ligation dans le plasmide pOCT4-eGFP-2A-Puro digéré par les enzymes BamHI et NheI (Figure 15B). Le produit de ligation est ensuite utilisé pour transformer des bactéries DH5α. Après mise en culture de plusieurs clones bactériens, et extraction de l'ADN plasmidique, l'insertion de la séquence mutée dans le plasmide pOCT4-eGFP-2A-Puro a été vérifiée par digestion avec les enzymes BamHI et NheI, puis séquençage (Figure 15C). Le plasmide ainsi obtenu est nommé pOCT4-eGFP-2A-Puro-PAM\_mutés.



**Figure 15 : Mutation des sites PAM dans la matrice de réparation Oct4-P2A-eGFP-PGK-Puro. (A)** Mutations à réaliser dans la matrice Oct4-P2A-eGFP-PGK-Puro **(B)** Insertion de la séquence contenant les sites PAM mutés dans le plasmide pOCT4-eGFP-2A-Puro **(C)** Gel d'agarose 0,8% des minipréparations pOCT4-eGFP-2A-Puro-PAM\_mutés, après linéarisation. **(D)** Séquençage du clone 4 montrant la présence des mutations C926T, C953T, C974A et C989T.

## I. 2. 2) Fabrication d'une matrice de réparation contenant le gène codant pour la protéine fluorescente Katushka (Kate2)

Les cellules dans lesquelles je vais introduire la matrice de réparation seront utilisées par la suite pour exprimer le système FUCCI, qui est un rapporteur du cycle cellulaire (*Bajar et al, 2016*). Ce système est basé sur l'utilisation d'une protéine fluorescente verte, la mAzami-Green (longueur d'émission 500 nm) et d'une protéine fluorescente orange, la Kusabira-Orange2 (longueur d'émission 560 nm). Afin d'éviter tout chevauchement des spectres d'émission des différentes protéines fluorescentes, j'ai remplacé le gène codant pour la GFP par le gène codant pour la protéine Katushka (Kate 2) (longueur d'émission 635 nm) dans la matrice de réparation. Après digestion du plasmide pMA-P2A-Kate2 par les enzymes BmtI et AvrII, un fragment de 1541 pb contenant le gène Kate2 a été inséré par ligation dans le plasmide pOCT4-eGFP-2A-Puro-PAM\_mutés digéré par les enzymes BmtI et AvrII (Figure 16A). Le produit de ligation est ensuite utilisé pour transformer des bactéries DH5α. Après mise en culture de plusieurs clones bactériens, et extraction de l'ADN plasmidique, l'insertion de la séquence Kate2 dans le plasmide pOCT4-eGFP-2A-Puro-PAM\_mutés a été vérifiée par digestion avec les enzymes BmtI et AvrII, puis séquençage (Figure 16B). Le plasmide ainsi obtenu est nommé pOCT4-Kate2-2A-Puro-PAM\_mutés.



**Figure 16 : Remplacement de la GFP par Kate 2 dans la matrice de réparation. (A)** Insertion de la séquence contenant le gène Kate2 dans le plasmide pOCT4-2A-eGFP-Puro-PAM\_mutés **(B)** Gel d'agarose 0,8% d'une minipréparation pOCT4-2A-Kate2-Puro-PAM\_mutés, après linéarisation.

Cette nouvelle matrice permettra donc d'obtenir des cellules qui expriment une protéine de fusion OCT4-P2A-Kate2, après sélection en présence de puromycine.

## **II. Fabrication d'une lignée fluorescente révélant l'expression du gène OCT4**

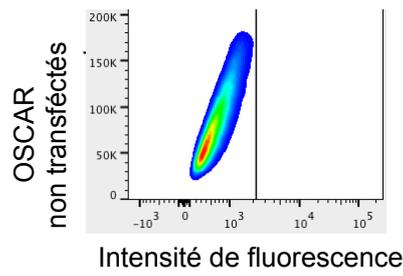
### **II. 1. Mises au point des conditions de transfections**

Les plasmides d'expression de la Cas9 et des guides (pX330A\_D10A\_G), ainsi que le plasmide codant pour la matrice de réparation (pOCT4-Kate2-2A-Puro-PAM\_mutés), sont transfectés dans les cellules ES humaines par électroporation. J'ai donc dans un premier temps mis au point les conditions d'électroporation. Pour cela, j'ai utilisé :

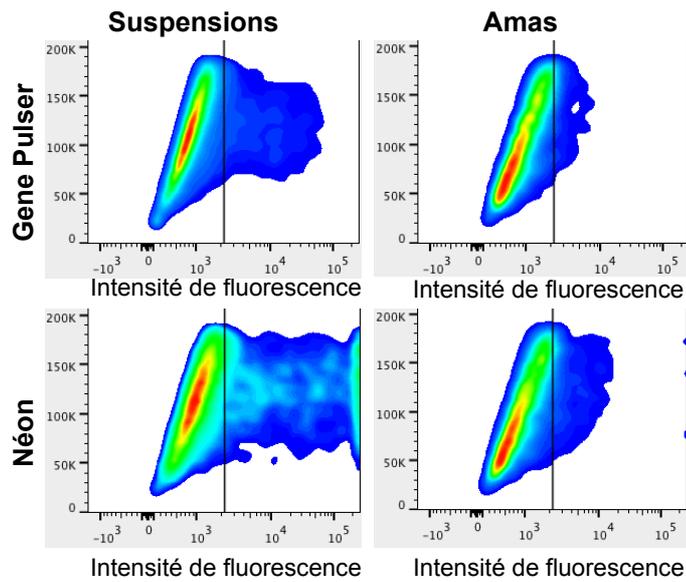
- le plasmide pGAE-GFP, qui permet l'expression de la GFP sous contrôle du promoteur ubiquitaire CAG ;
- deux types d'électroporateurs: le Gene Pulser II (BioRad) et le Système Neon (ThermoFisher) ;
- deux types de suspensions cellulaires: en amas (dissociation manuelle) ou unicellulaire (dissociation enzymatique et traitement au Rock inhibiteur).

J'ai réalisé ces tests dans les cellules ES humaines de la lignée OSCAR. 48h après la transfection, j'ai analysé l'expression de la GFP par cytométrie en flux dans ces différentes conditions (Figure 17).

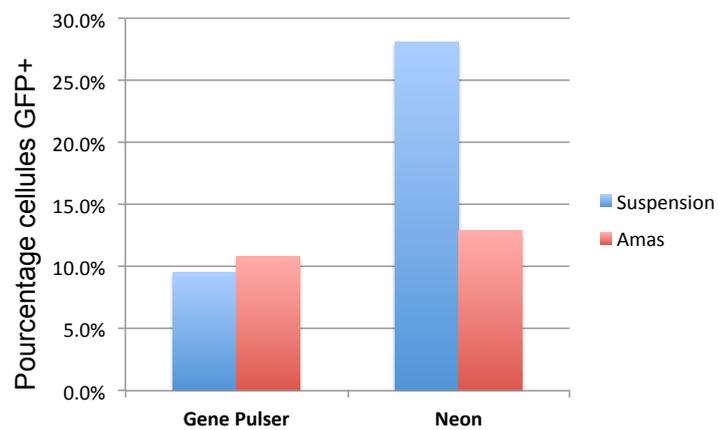
A



B



C



**Figure 17 : Analyse de l'expression de la GFP dans les cellules OSCAR transfectées avec le plasmide pGAE-GFP. (A) Cellules hESC OSCAR non transfectées (B) Dans des cellules hESC OSCAR transfectées avec l'appareil Gene Pulser II ou Neon, en amas et en suspension cellulaire. (C) Histogramme montrant le pourcentage d'hESC OSCAR exprimant la GFP en fonction de la condition de transfection.**

Dans les conditions de transfection avec l'appareil Gene Pulser II, le pourcentage de cellules GFP positives est de 10,8 % dans les cellules transfectées en amas et 9,56 % dans les cellules transfectées en suspension unicellulaire. Avec l'appareil Neon, le pourcentage de cellules GFP positives est de 12,9 % dans les cellules transfectées en amas et 28,1 % dans les cellules transfectées en suspension unicellulaire (Figure 17B et C). La condition d'électroporation en suspension unicellulaire avec l'appareil Neon semble donner les meilleurs résultats. Je vais donc utiliser ces conditions pour la transfection des plasmides pX330A\_D10A\_G2 et pOCT4-Kate2-2A-Puro-PAM\_mutés dans les cellules hESC OSCAR.

## **II. 2. Fabrication d'une lignée OSCAR OCT4-P2A-eGFP-Puro**

### **II. 2. 1) Obtention des clones**

Dans un premier temps, j'ai utilisé la matrice de réparation modifiée pOCT4-2A-eGFP-Puro-PAM\_mutés afin de me familiariser avec la technologie CRISPR/Cas9. Des cellules hESC OSCAR ont été co-électroporées avec les plasmides pX330A\_D10A\_G2 et modifiée pOCT4-2A-eGFP-Puro-PAM\_mutés. Les cellules sont ensuiteensemencées en condition clonale et les cellules ayant intégré la matrice de réparation sont sélectionnées en présence de puromycine. Après 7 jours de sélection, j'ai repiqué et amplifié les clones résistants. J'ai ainsi obtenu 12 clones dont seulement 10 ont pu être amplifiés et congelés.

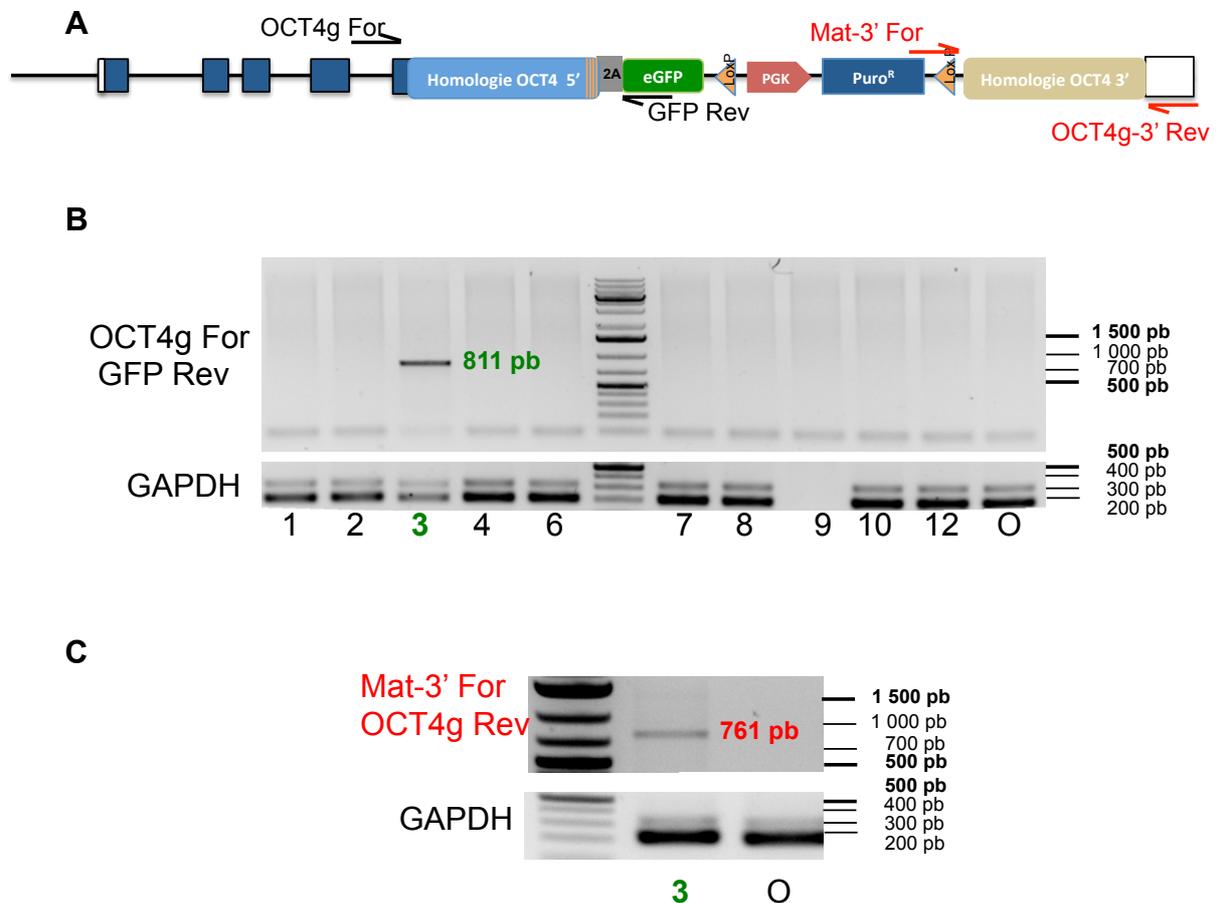
### **II. 2. 2) Caractérisation des clones**

J'ai tout d'abord vérifié l'intégration spécifique de la matrice de réparation dans le locus *OCT4*. En effet, la résistance à la puromycine peut être due à une intégration aléatoire de la matrice de réparation. La figure 8A représente la structure d'un allèle *OCT4* après intégration de la matrice de réparation. Pour cela, j'ai réalisé une PCR sur l'ADN génomique des 10 clones puromycine résistants en utilisant les amorces suivantes :

- OCT4g For, spécifique de la région génomique d'*OCT4* en amont du bras d'homologie 5' de la matrice (Figure 18A) (Annexe 2).

- GFP Rev, spécifique du gène codant pour la GFP, localisé dans la matrice de réparation (Figure 18A) (Annexe 2).

La présence d'un produit PCR d'une taille de 811 pb confirme l'intégration de la matrice de réparation à l'endroit désiré. Après PCR, seul le clone n°3 présente une bande d'une taille d'environ 800pb (Figure 18B).



**Figure 18 : Criblage des clones puromycine résistants obtenus après transfection de la matrice pOCT4-eGFP-2A-Puro. (A)** Schéma représentant la structure d'un allèle *OCT4* après recombinaison homologue avec la matrice pOCT4-2A-eGFP-Puro. Les flèches noires représentent les amorces OCT4g\_For et GFP\_Rev. Les flèches rouges représentent les amorces Mat-3'\_For et OCT4g-3'\_Rev. **(B)** Analyse par PCR de l'ADN génomique des clones résistants à la puromycine avec les amorces OCT4g\_For et GFP\_Rev. Le gène de la GAPDH est utilisé comme contrôle interne, **(C)** Analyse par PCR de l'ADN génomique du clone n°3 avec les amorces Mat-3'\_For et OCT4g-3'\_Rev. Le gène de la GAPDH est utilisé comme contrôle interne.

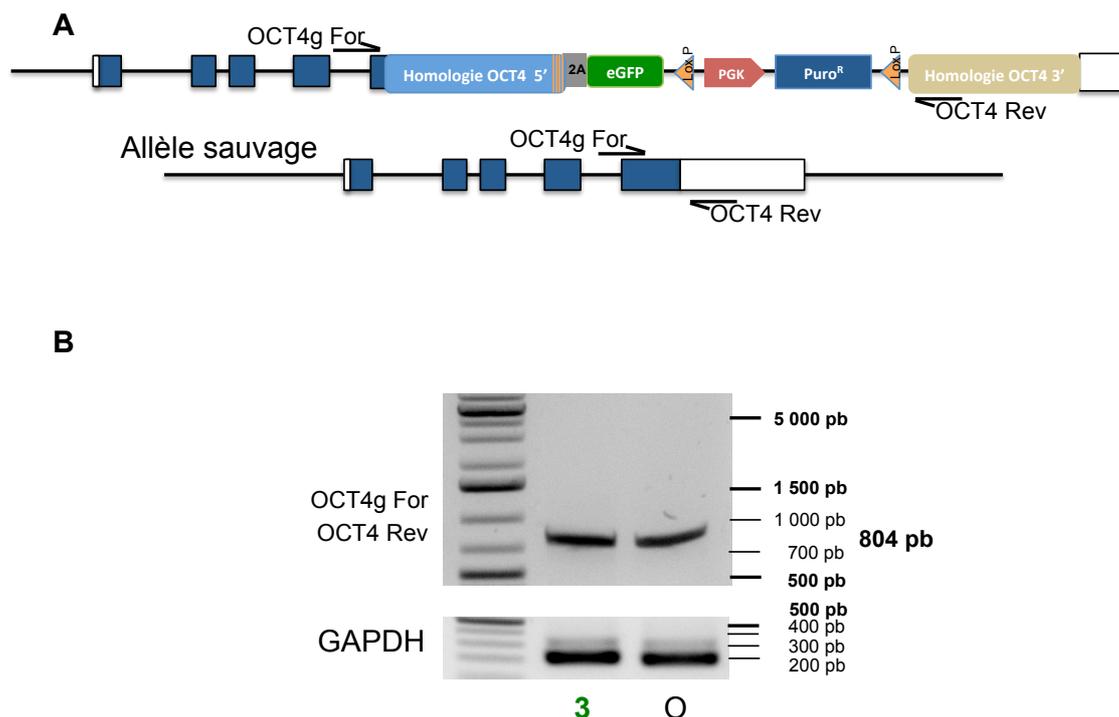
Sur le clone n°3, j'ai analysé l'intégration de la matrice de réparation du côté 3' du site d'intégration. Pour cela, j'ai réalisé une PCR sur l'ADN génomique avec les amorces suivantes :

- Mat-3'\_For, spécifique de la matrice en amont du bras d'homologie 3' (Figure 18A) (Annexe 2) ;

- OCT4g-3'\_Rev, spécifique de la région génomique du gène *OCT4* en aval du bras d'homologie 3' (Figure 18A) (Annexe 2).

La présence d'un produit PCR d'une taille d'environ 800 pb confirme l'intégration de la matrice de réparation à l'endroit désiré. Le clone OSCAR *OCT4-eGFP-Puro* n°3 a intégré la matrice de réparation, sans aucune anomalie aussi bien du côté 3' que du côté 5'. Ce résultat devra être confirmé par séquençage.

J'ai ensuite vérifié si un ou les deux allèles *OCT4* ont été mutés dans le clone OSCAR *OCT4-eGFP-Puro* n°3. Pour cela, j'ai réalisé une PCR sur l'ADN génomique, avec les amorces OCT4g\_For et OCT4\_Rev (Figure 19A) (Annexe 2). La présence d'un produit PCR d'une taille d'environ 800 pb indique la présence d'un allèle sauvage, où il n'y a pas eu d'intégration de la matrice de réparation. Une bande de 4127 pb devrait également être détectée, indiquant l'allèle où il y a eu recombinaison homologue (Figure 19B). Les conditions de PCR utilisées ne sont sans doute pas optimales pour l'amplification d'un fragment de 4127pb. Les cellules hESC OSCAR sont utilisées comme témoin positif.



**Figure 19 : Etude de l'hétérogénéité de l'insertion du pOCT4-eGFP-2A-Puro-PAM mutés dans le clone n°3. (A)** Schéma représentant la structure d'un allèle *OCT4* après recombinaison homologue avec la matrice pOCT4-eGFP-2A-Puro et un allèle sauvage du gène *OCT4*. Les flèches noires représentent les amorces OCT4g\_For et OCT4\_Rev. **(B)** Analyse par PCR de l'ADN génomique du clone n°3 et des cellules OSCAR. Le gène de la GAPDH est utilisé comme contrôle interne.

## II. 3. Fabrication d'une lignée OSCAR *OCT4-P2A-Kate2-Puro*

### II. 3. 1) Obtention des clones

Dans un second temps, j'ai utilisé la matrice de réparation pOCT4-2A-Kate2-Puro-PAM\_mutés. Des cellules hESC OSCAR ont été co-électroporées avec les plasmides pX330A\_D10A\_G2 et pOCT4-Kate2-2A-Puro-PAM\_mutés. Les cellules sont ensuite ensemencées en condition clonale, et les cellules ayant intégrées la matrice de réparation sont sélectionnées en présence de puromycine. Après 7 jours de sélection, j'ai repiqué et amplifié les clones résistants. J'ai ainsi obtenu 48 clones, dont 46 ont pu être caractérisés.

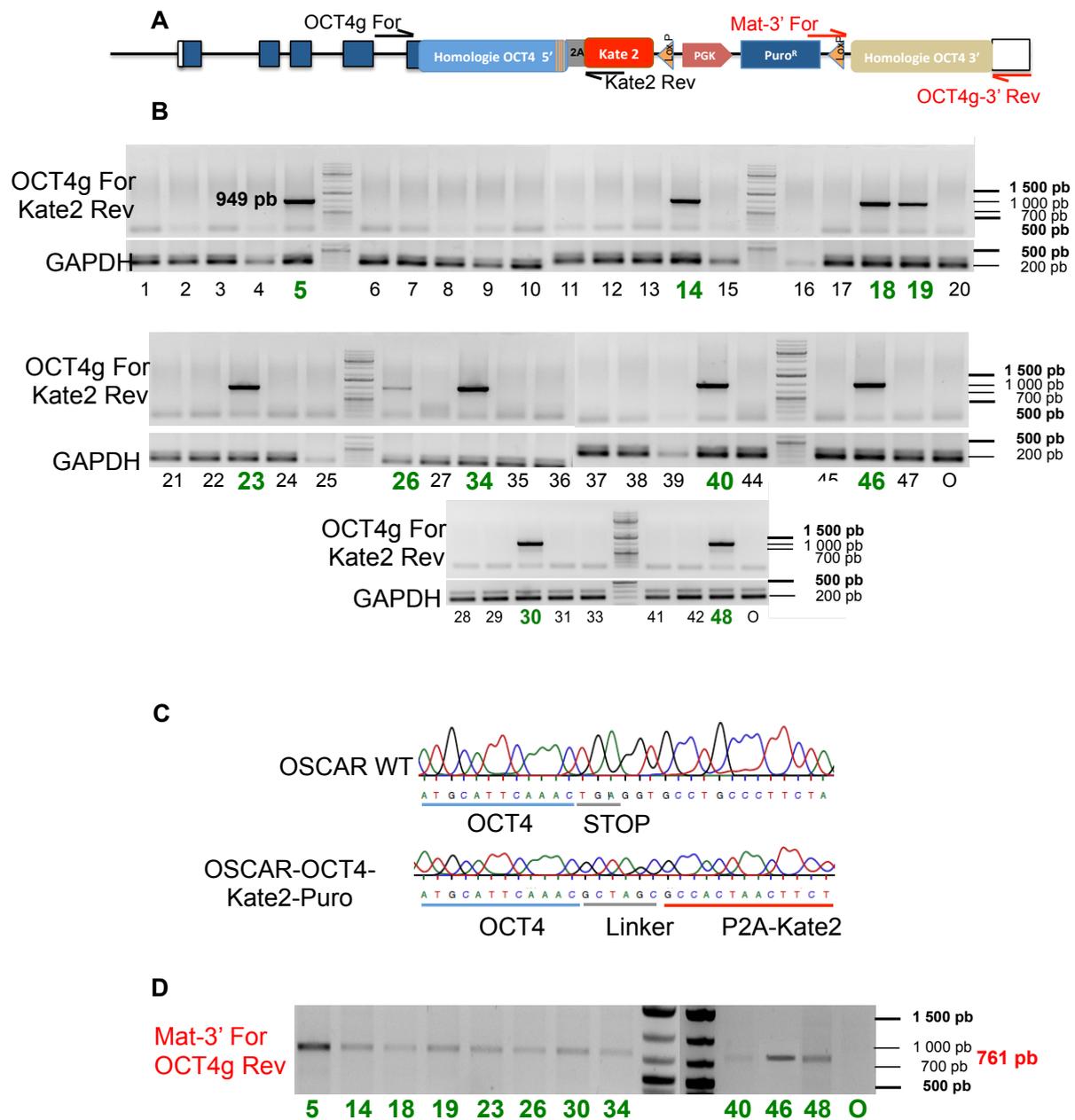
### II. 3. 2) . Caractérisation des clones

J'ai tout d'abord vérifié l'intégration de la matrice de réparation dans le locus *OCT4*. En effet, la résistance à la puromycine peut être due à une intégration aléatoire de la matrice de réparation. La Figure 20A représente la structure d'un allèle *OCT4* après intégration de la matrice de réparation. Pour cela, j'ai réalisé une PCR sur l'ADN génomique des 46 clones puromycine résistants en utilisant les amorces suivantes :

- OCT4g\_For, spécifique de la région génomique d'*OCT4* en amont du bras d'homologie 5' de la matrice (Figure 20A) ;
- Kate2\_Rev, spécifique du gène codant pour Kate2, localisé dans la matrice de réparation (Figure 20A) (Annexe 2).

La présence d'un produit PCR d'une taille de 949 pb confirme l'intégration de la matrice de réparation à l'endroit désiré. Après PCR, 11 clones présentent une bande d'une taille d'environ 1000 pb (Figure 20B).

Puis j'ai vérifié par séquençage l'insertion de la matrice de réparation dans les clones *OCT4-P2A-Kate2-Puro*. Les cellules hES OSCAR WT ont été utilisées comme contrôle négatif. Tous les clones ont intégré la matrice sans défaut du côté de l'homologie 5'. En effet le « linker » a remplacé le codon STOP du gène *OCT4* et est suivi par la séquence Kate2 (Figure 20C).



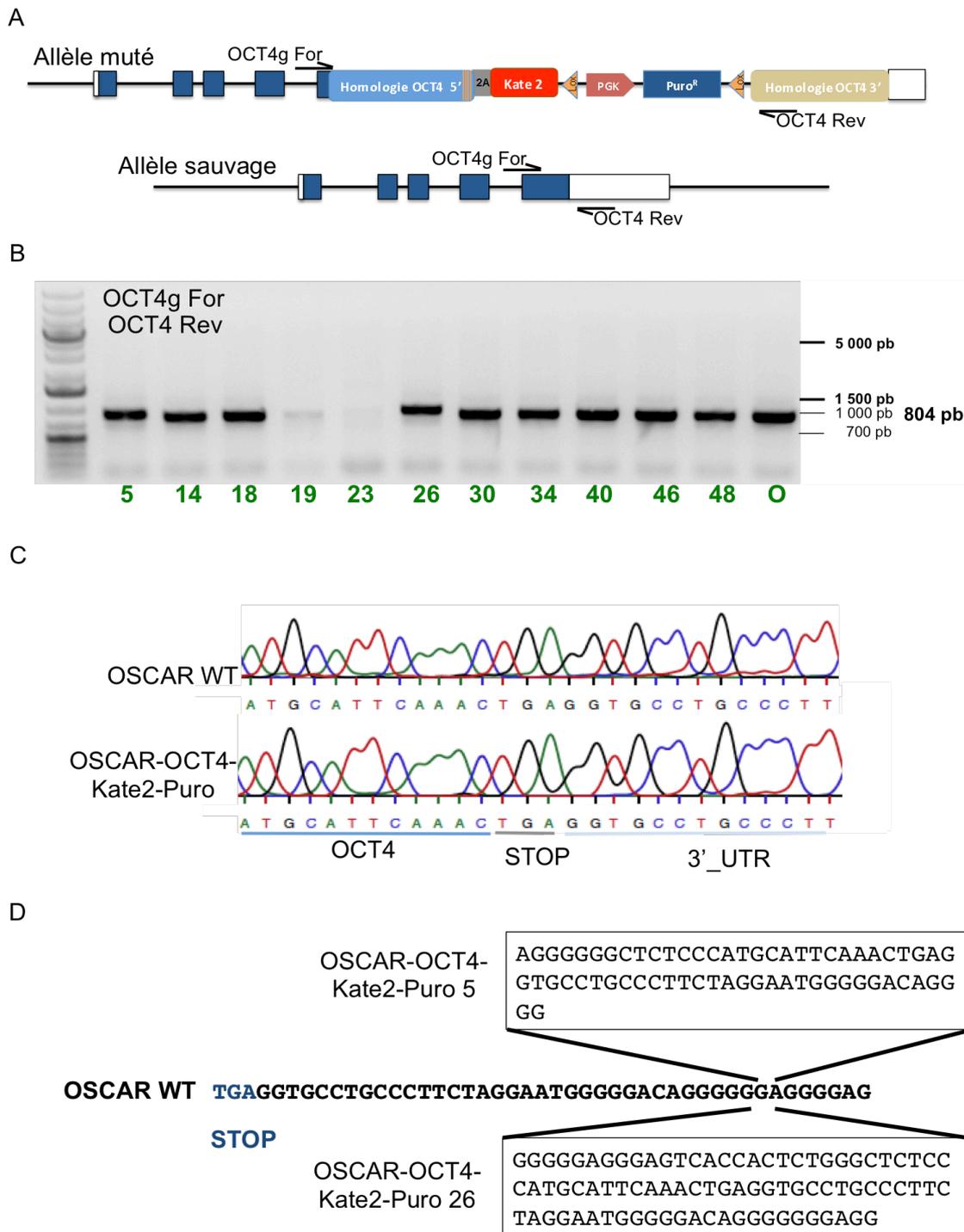
**Figure 20 : Criblage des clones puromycine résistants, obtenus après transfection de la matrice pOCT4-2A-Kate2-Puro-PAM\_mutés. (A)** Schéma représentant la structure d'un allèle *OCT4* après recombinaison homologue avec la matrice pOCT4-2A-Kate2-Puro-PAM\_mutés. Les flèches noires représentent les amorces OCT4g\_For et Kate2\_Rev. Les flèches rouges représentent les amorces Mat-3'\_For et OCT4g-3'\_Rev. **(B)** Analyse par PCR de l'ADN génomique des clones résistants à la puromycine, avec les amorces OCT4g\_For et Kate2\_Rev. Le gène de la GAPDH est utilisé comme contrôle interne. **(C)** Séquençage de la région du codon STOP du gène *OCT4* dans les cellules hESC OSCAR et les clones ayant intégré la matrice de réparation dans le locus du gène *OCT4*. **(D)** Analyse par PCR de l'ADN génomique des clones ayant intégré la matrice de réparation dans le locus du gène *OCT4*, avec les amorces Mat-3'\_For et OCT4g-3'\_Rev.

Dans les 11 clones résistants à la puromycine, j'ai analysé l'intégration de la matrice de réparation du côté 3' du site d'intégration. Pour cela, j'ai réalisé une PCR sur l'ADN génomique avec les amorces suivantes :

- Mat-3'\_For, spécifique de la matrice en amont du bras d'homologie 3' (Figure 20A) ;
- OCT4g-3'\_Rev, spécifique de la région génomique du gène *OCT4*, en aval du bras d'homologie 3' (Figure 20A).

La présence d'un produit PCR d'une taille de 761 pb confirme l'intégration de la matrice de réparation à l'endroit désiré (Figure 20D). Les clones OSCAR *OCT4*-Kate2-Puro ont intégré la matrice de réparation sans aucune anomalie aussi bien du côté 3' que du côté 5'.

J'ai ensuite vérifié si un seul ou les deux allèles *OCT4* sont mutés dans les clones OSCAR *OCT4*-Kate2-Puro. Pour cela, j'ai réalisé une PCR sur l'ADN génomique avec les amorces OCT4g\_For et OCT4g-3'\_Rev (Figure 21A). La présence d'un produit PCR d'une taille de 804 pb indique la présence d'un allèle sauvage où il n'y a pas eu d'intégration de la matrice de réparation. Une bande de 4889pb devrait également être détectée, indiquant l'allèle où il y a eu recombinaison homologue (Figure 21A). Les cellules hESC OSCAR sont utilisées comme témoin positif. Après PCR, une seule bande d'une taille d'environ 800 pb est détectée dans tous les clones (Figure 21B). Les résultats PCR montrent que la bande observée dans les clones 5 et 26 a une taille légèrement supérieure à celle observée dans les cellules OSCAR. Les résultats du séquençage ont montré qu'il y a une insertion de 62 pb dans la région du codon STOP du gène *OCT4* du clone 5 et de 85 pb pour le clone 26 (Figure 21D). Ces insertions ont été induites après le clivage par la Cas9\_D10A du second allèle du gène *OCT4* et de la réparation en NHEJ. Les 9 autres clones ont un allèle sauvage non muté (Figure 21C). Les conditions de PCR utilisées ne sont sans doute pas optimales pour l'amplification d'un fragment de 4 889pb. Ces résultats montrent donc que les clones OSCAR *OCT4*-Kate2-Puro sont hétérozygotes. Il y a eu recombinaison homologue sur un seul des deux allèles du gène *OCT4*.



**Figure 21 : Etude de l'hétérogénéité de l'insertion du pOCT4-eGFP-2A-Puro dans les clones positifs. (A)** Schéma représentant la structure d'un allèle *OCT4* après recombinaison homologue avec la matrice pOCT4-eGFP-2A-Puro et un allèle sauvage du gène *OCT4*. Les flèches noires représentent les amorces OCT4g\_For et OCT4\_Rev. **(B)** Analyse par PCR de l'ADN génomique des clones OSCAR *OCT4*-Kate2-Puro résistants à la puromycine et des cellules OSCAR. **(C)** Séquençage de la région du codon STOP du gène *OCT4* pour les cellules hESC OSCAR et du clone 46 OSCAR *OCT4*-Kate2-Puro. **(D)** Séquences insérées la région du codon STOP du gène *OCT4* pour les clones 5 et 26 OSCAR *OCT4*-Kate2-Puro.

Le tableau 5 montre le nombre de clones OSCAR *OCT4*-eGFP-Puro et OSCAR *OCT4*-Kate2-Puro obtenus. J'ai ainsi pu calculer l'efficacité relative, correspondant au ratio du nombre de clones positifs sur le nombre de clones résistants à la puromycine. J'ai déterminé l'efficacité absolue en calculant le ratio du nombre de clones positifs sur le nombre de cellules transfectées pour chaque expérience.

**Tableau 5 : Récapitulatif des résultats de la création de lignées OSCAR *OCT4*-P2A-Kate2-Puro**

Gène ciblé	CRISPR	Gène Fluo	Nb clones Puro <sup>R</sup>	Clones positifs	Hetero/homozygote	Efficacité Relative	Efficacité Absolue
OCT4	D10A_G2	eGFP	10	1	1 / 0	10% (1/10)	$1.10^{-7}$
		Kate2	46	11	9 / 0	19,6% (9/46)	$4,1.10^{-7}$ (9/22.10 <sup>6</sup> )

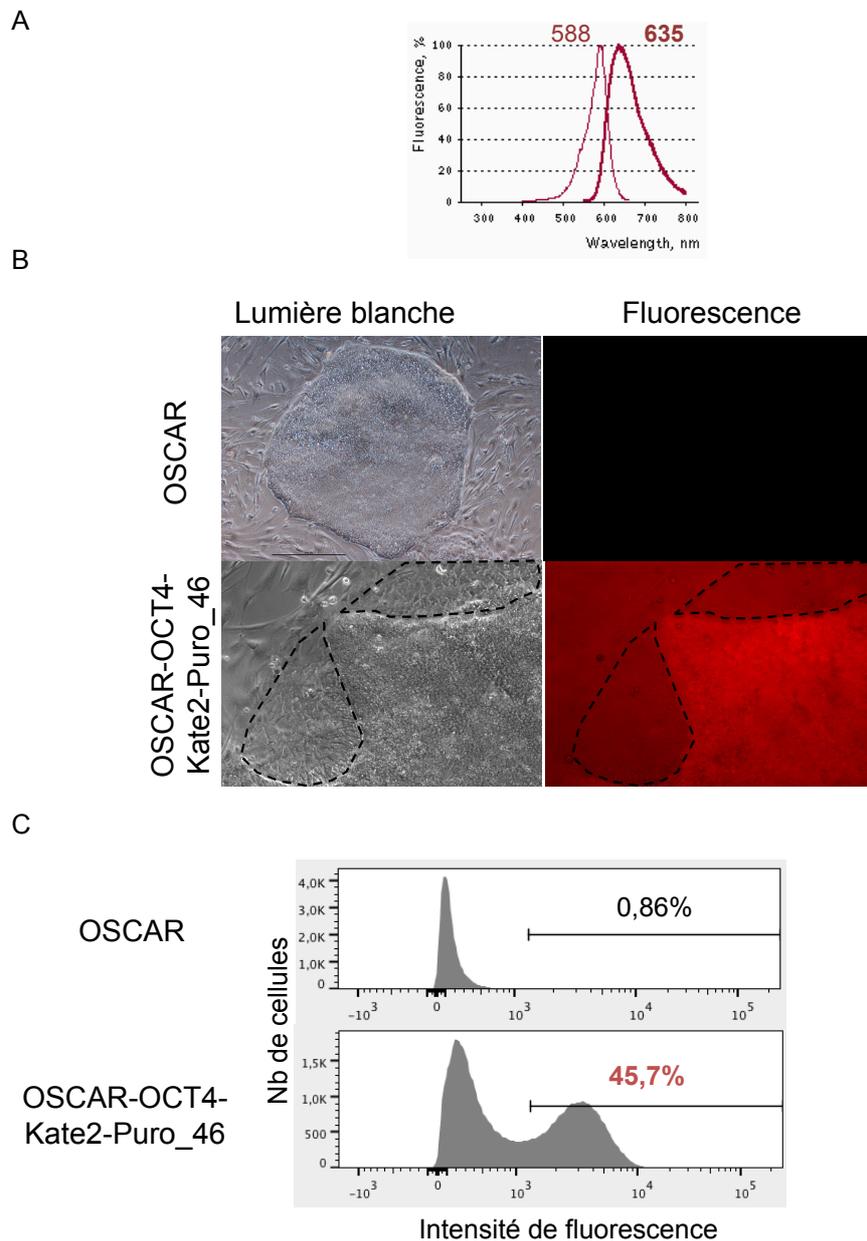
J'ai obtenu 1 seul clone OSCAR *OCT4*-eGFP-Puro et 9 clones OSCAR *OCT4*-Kate2-Puro. Cette augmentation pourrait être expliquée par la mutation des sites PAM dans la matrice de réparation pOCT4-Kate2-2A-Puro-PAM\_mutés.

### **II. 3. 3) Analyse de l'expression de la protéine Katushka dans la lignée OSCAR *OCT4*-Kate2-Puro.**

J'ai analysé l'expression de la protéine fluorescente Katushka dans la lignée OSCAR *OCT4*-P2A-Kate2-Puro. Etant sous contrôle du promoteur *OCT4*, la protéine Katushka doit présenter le même profil d'expression que la protéine *OCT4*. Son expression doit être forte et homogène dans les cellules indifférenciées. Après induction de la différenciation, l'expression de la protéine Katushka doit diminuer. J'ai donc analysé l'expression de la protéine Katushka dans les cellules OSCAR *OCT4*-Kate2 non différenciées et après induction de leur différenciation.

**II. 3. 3. 1) Analyse de l'expression de la protéine Kate2 dans les indifférenciées.**

L'analyse de l'expression de la protéine Kate2 a été réalisée dans le Clone OSCAR *OCT4*-Kate2-Puro n°46 (Figure 22).

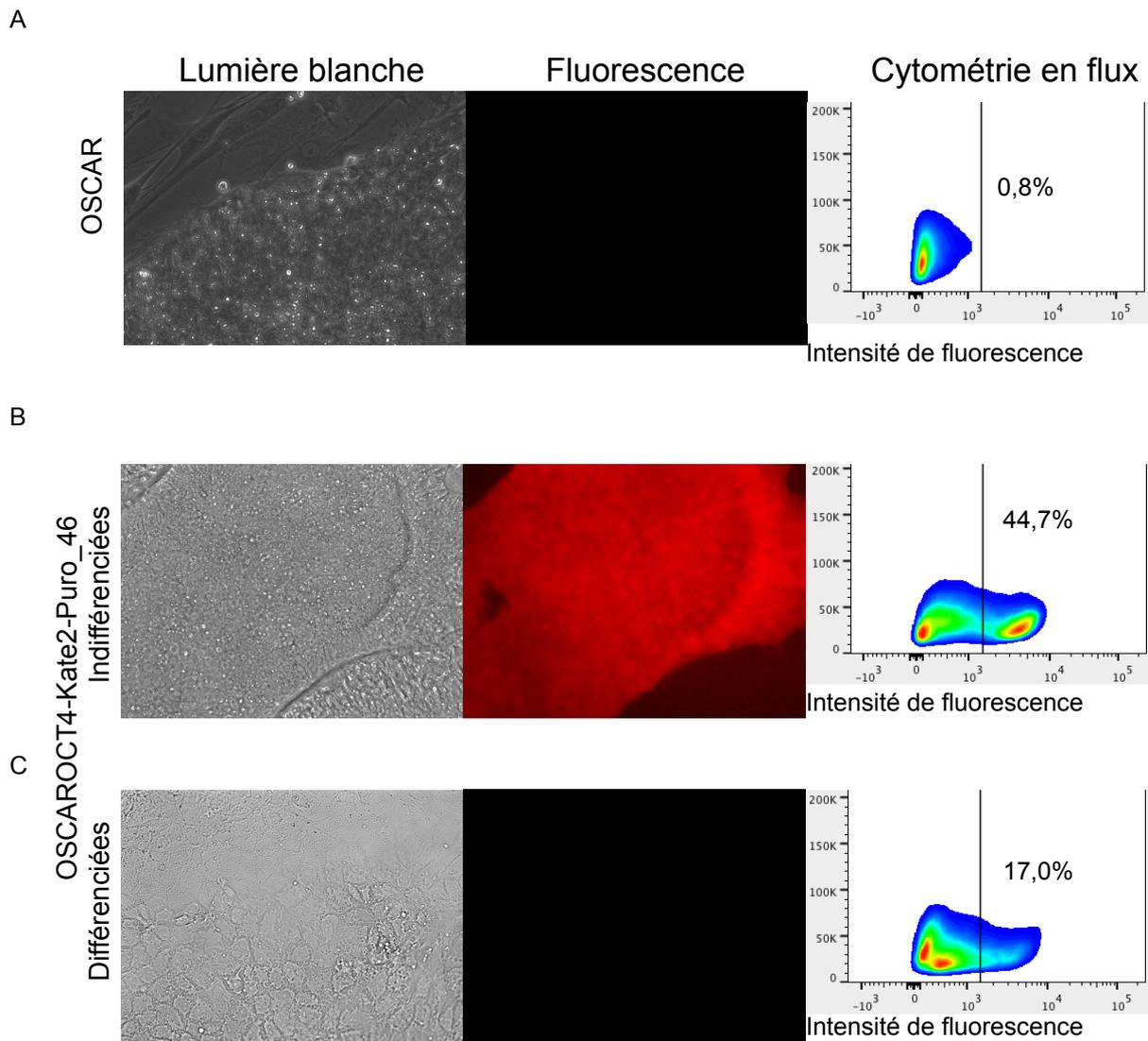


**Figure 22 : Analyse de l'expression de la protéine Kate2 dans les cellules OSCAR *OCT4*-Kate2-Puro\_46 (A) Spectre d'excitation et d'émission de la protéine Kate2. (B) Observation de l'expression de la protéine Kate2 dans des colonies de cellules OSCAR et de cellules OSCAR *OCT4*-Kate2-Puro#46, par microscopie à fluorescence. (C) Analyse en cytométrie en flux de l'expression de Kate2 dans les cellules OSCAR et de cellules OSCAR *OCT4*-Kate2-Puro#46.**

La protéine fluorescente Kate2 se caractérise par une longueur d'onde d'excitation de 588 nm et une longueur d'onde d'émission de 633 nm. Sa demi-vie est de 20 min (Figure 22A). L'expression de la protéine Kate2 peut être analysée par microscopie à fluorescence ou par cytométrie en flux. A l'état indifférencié, les cellules de la lignée OSCAR *OCT4*-Kate2 poussent sous forme de colonies sur un tapis de cellules nourricières (MEFs). Dans un premier temps, j'ai observé au microscope à fluorescence des colonies de cellules OSCAR *OCT4*-Kate2 (Figure 22B). L'expression de la protéine Kate2 peut être détectée dans les cellules OSCAR *OCT4*-Kate2. L'expression de Kate2 ne peut être détectée ni dans les cellules OSCAR sauvage, ni dans les MEFs. De plus, les cellules OSCAR *OCT4*-Kate2 différenciées observées à la périphérie de la colonie (Figure 22B, pointillés) n'expriment pas Kate2. Afin de déterminer le pourcentage de cellules positives, j'ai analysé l'expression de Kate2 par cytométrie en flux (Figure 22C). Les résultats montrent que 45,7% des cellules OSCAR *OCT4*-Kate2 expriment la protéine Kate2 (Figure 22C). L'absence d'expression de Kate2 dans plus de 50% des cellules pourrait être expliquée par la présence de la cassette pGK-Puro loxée en 3' du gène Kate2 (Zhu *et al*, 2015). En effet, cette séquence pourrait interférer avec des signaux de régulation situés dans la région 3' non traduite du gène *OCT4*.

### ***II. 3. 3. 2) Analyse de l'expression de la protéine Kate2 dans les cellules différenciées***

Afin d'induire la différenciation, les cellules OSCAR *OCT4*-Kate2-Puro\_46 sont cultivées en absence de bFGF pendant 7 jours. L'expression de la protéine Kate2 est ensuite analysée par microscopie à fluorescence et par cytométrie en flux (Figure 23). L'analyse a également été réalisée dans des cellules OSCAR WT et des cellules OSCAR *OCT4*-Kate2-Puro\_46 indifférenciées, cultivées en présence de bFGF.



**Figure 23 : Analyse de l'expression de la protéine Kate2 dans les cellules OSCAR *OCT4-Kate2-Puro\_46* indifférenciées et différenciées.** Observation au microscope à fluorescence et analyse par cytométrie en flux de l'expression de Kate2 dans les cellules OSCAR **(A)** et dans les cellules OSCAR *OCT4-Kate2-Puro\_46* indifférenciées **(B)** et différenciées **(C)**.

Les cellules OSCAR *OCT4-Kate2-Puro\_46* analysées dans la figure 22 étaient à passage 34+16. Dans cette nouvelle analyse, j'ai utilisé des cellules à passage 34+21. L'analyse par cytométrie en flux montre que, à l'état indifférencié, 44,7% des cellules à P34+21 expriment la protéine Kate2, soit une diminution de 1% entre les deux expériences. L'expression de la protéine Kate2 est donc relativement stable au cours de la culture.

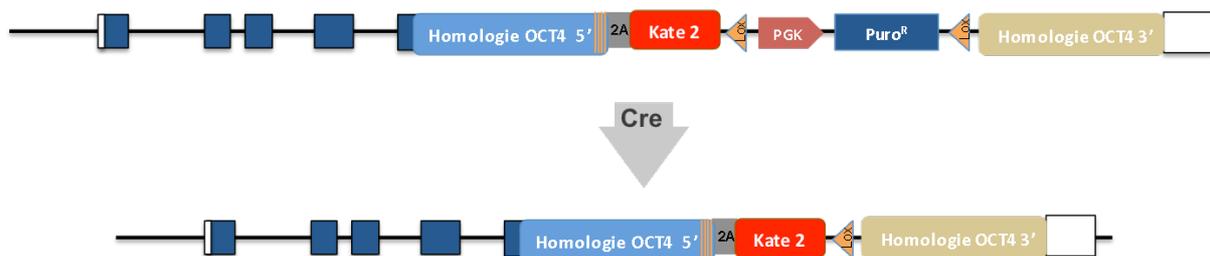
Après induction de la différenciation, l'expression de la protéine Kate2 n'est plus détectée par microscopie à fluorescence (Figure 23C). L'analyse en cytométrie de flux montre une diminution du niveau d'expression de Kate2 et du nombre de cellules Kate2 positives (17%).

Il faudrait analyser l'expression de Kate 2 après une période de différenciation plus longue afin d'observer une éventuelle absence d'expression de Kate2.

Ces résultats semblent donc montrer que l'expression de Kate2 reflète le patron d'expression de la protéine *OCT4* dans les cellules OSCAR *OCT4*-Kate2-Puro\_46. Cette expression, forte et homogène dans les cellules indifférenciées, diminue après induction de la différenciation.

## II. 4. Fabrication d'une lignée OSCAR *OCT4*-Kate2 avec excision de la cassette pGK-Puro loxée.

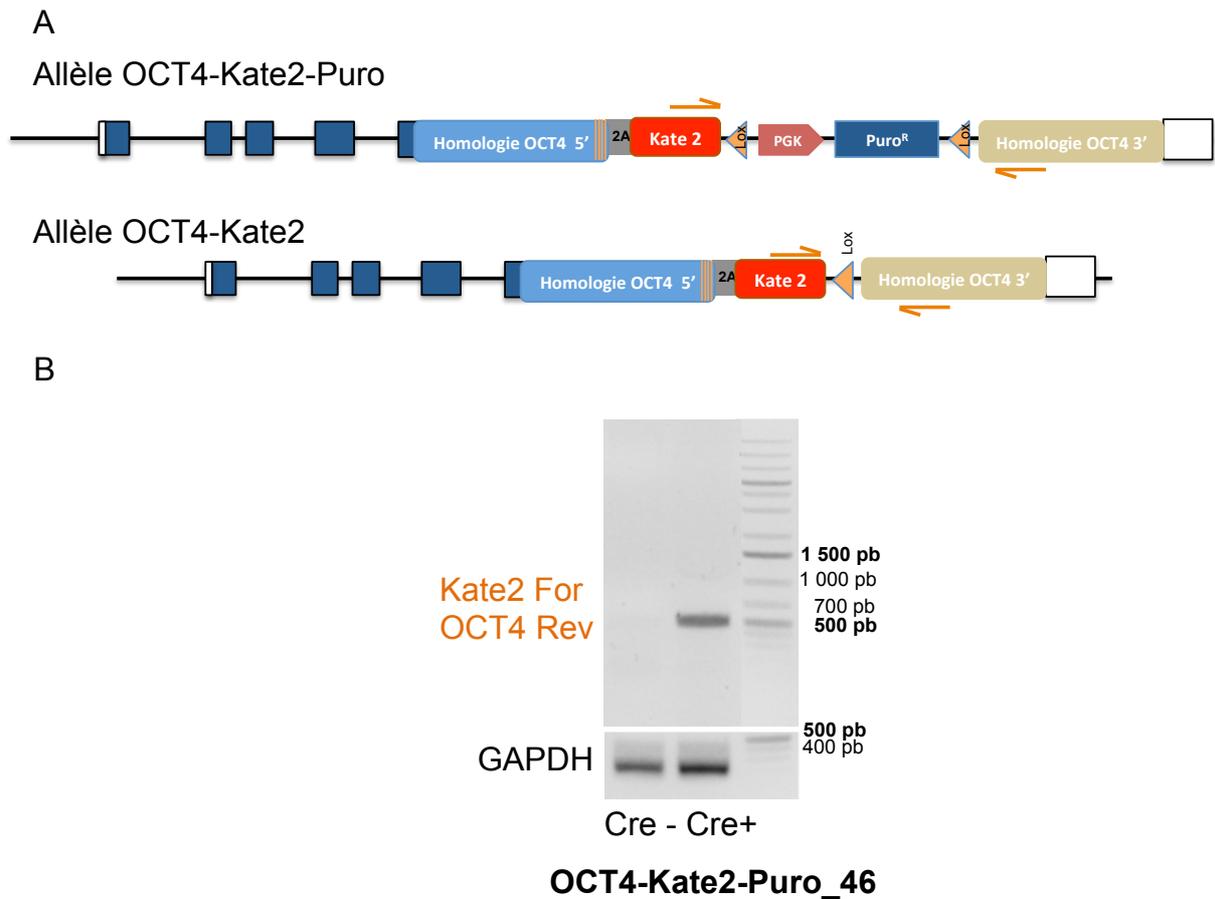
La matrice de réparation pOCT4-Kate2-2A-Puro-PAM\_mutés contient une cassette pGK-Puro encadrée par des sites Lox. Cette cassette peut être excisée grâce à l'action de la Recombinase Cre (Figure 24). Cela permettra d'analyser l'éventuelle interférence de cette cassette sur l'expression de la protéine Kate2. D'autre part, la résistance à la puromycine pourra être à nouveau utilisée pour modifier les cellules OSCAR-*OCT4*-Kate2.



**Figure 24 : Représentation schématique de l'excision de la cassette PGK-Puro par la recombinase Cre.**

### II. 4. 1) Analyse de l'excision de la cassette de sélection par PCR sur la population totale OSCAR *OCT4*-Kate2

Les cellules OSCAR *OCT4*-Kate2-Puro\_46 ont été transfectées avec un plasmide d'expression de la recombinase Cre. 48h après transfection, j'ai analysé par PCR l'excision de la cassette pGK-Puro avec les oligonucléotides Kate2 For et OCT4 Rev (Figure 25A) (Annexe 2). L'analyse a également été réalisée dans des cellules non transfectées.



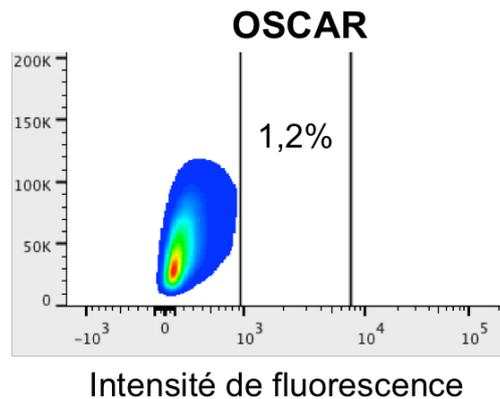
**Figure 25 : Excision de la cassette pGK-Puro dans les cellules OSCAR *OCT4-Kate2-Puro\_46*** (A) Schéma représentant la structure de l'allèle *OCT4-P2A-Kate2-Puro*, avant et après excision de la cassette pGK-Puro. Les flèches orange représentent les amorces Kate2\_For et OCT4\_Rev. (B) Analyse par PCR de l'ADN génomique des cellules OSCAR-*OCT4-Kate2-Puro\_46* avant et après transfection de la Cre Recombinase, à l'aide des oligonucléotides Kate2\_For et OCT4\_Rev. Le gène de la GAPDH est utilisé comme contrôle interne.

La présence d'un produit PCR d'une taille de 482 pb indique l'excision de la cassette de sélection. Une bande PCR de 2268 pb correspondant aux cellules où il n'y a pas eu excision devrait également être détectée (Figure 25A). Après PCR, une bande d'une taille d'environ 500 pb est détectée dans les cellules transfectées (Figure 25B). Ces résultats montrent que la transfection de la recombinaise Cre permet d'exciser la cassette pGK-Puro dans les cellules OSCAR *OCT4-Kate2-Puro\_46*. Il y a deux explications possibles pour expliquer l'absence d'un amplicon de 2268 pb. La première est que la recombinaise Cre a excisé la cassette de sélection dans toutes les cellules. La seconde est que l'amplification de plus de 2 000 pb à partir d'un ADN génomique est compliquée. Donc que ces amorces ne permettent pas de détecter une population cellulaire qui a conservé la cassette de sélection.

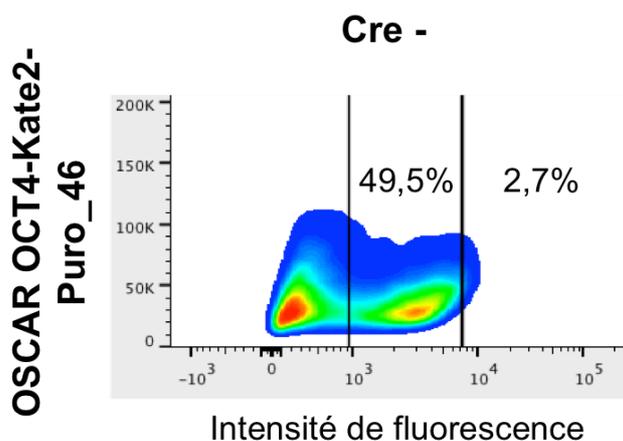
## II. 4. 2) Analyse de l'expression du gène *Kate2* après excision de la cassette pGK-Puro

Afin de vérifier l'effet de l'excision de la cassette de résistance sur l'expression de la protéine *Kate2*, j'ai analysé son expression dans les cellules OSCAR *OCT4-Kate2-Puro\_46* après transfection de la recombinaise Cre (Figure 26). Cette analyse a été réalisée par cytométrie en flux.

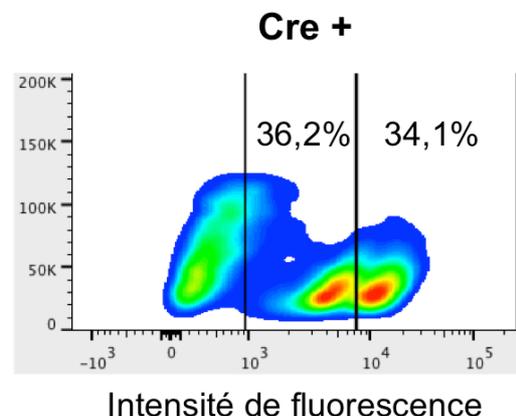
A



B



C



**Figure 26 : Analyse de l'expression de *Kate2* dans les cellules OSCAR *OCT4-Kate2-Puro* transfectées avec la recombinaise Cre par cytométrie en flux avant (B) et après (C) transfection. Les cellules OSCAR sont utilisées comme contrôle négatif (A).**

Dans les cellules OSCAR *OCT4-Kate2-Puro\_46* non transfectées, 52,2% des cellules expriment la protéine *Kate2*, à un niveau comparable à celui observé dans les figures 22 et 23. Dans les cellules OSCAR *OCT4-Kate2-Puro\_46* transfectées avec la recombinaise Cre,

deux populations exprimant la protéine Kate2 peuvent être observées : (i) une première population (36,2%) avec un niveau de fluorescence comparable à celui observé dans les cellules non transfectées (ii) une deuxième population (34,1%) avec un niveau de fluorescence plus élevé. Cette deuxième population pourrait correspondre aux cellules OSCAR *OCT4*-Kate2-Puro\_46 dans les lesquelles la cassette pGK-Puro a été excisée. Une vérification par PCR est en cours.

# Discussions et conclusion

---

L'équipe de Pierre Savatier étudie les mécanismes de pluripotence chez les espèces autre que la souris.

Chez la souris, les états de pluripotences, naïf et amorcé, sont bien définis. Les mESC s'auto-renouvellent à l'état de pluripotence naïf car elles sont capables de coloniser un embryon et de produire des chimères. De plus elles sont stables génétiquement. Les mEpiSC, un autre type de cellules pluripotentes chez la souris, s'auto-renouvellent à l'état amorcé de pluripotence. Elles sont instables génétiquement et ne colonisent pas l'embryon. Des analyses transcriptomiques ont permis de montrer que les cellules mESC se rapprochent de l'ICM de souris au stade préimplantatoire.

Chez l'homme, certains gènes sont exprimés dans l'embryon humain au stade morula et dans les cellules de l'ICM au stade blastocyste préimplantatoire. Ces gènes sont ensuite réprimés quand les cellules de l'ICM sont cultivées pour produire des cellules hESC dans les conditions de cultures décrites précédemment. Ces cellules s'auto-renouvellent à l'état de pluripotence dit « amorcées ».

Par mimétisme avec la souris, si l'on reprogramme des cellules hESC pour qu'elles aient un transcriptome qui se rapproche de l'ICM au stade préimplantatoire, alors les cellules seraient plus stables génétiquement et elles s'auto-renouvelleraient à l'état de pluripotence dit « naïves ».

Pour ces raisons l'équipe de Pierre Savatier cherche à établir des lignées de hESC dont le transcriptome se rapproche de celui de l'ICM humain. Pour cela l'équipe cherche à développer des lignées cellulaires rapportrices de l'états de la pluripotence avec l'introduction de marqueurs fluorescents sous contrôle du gène de pluripotence *OCT4* et du gène de pluripotence dit « naïf » *KLF17*.

# **I. Fabrication d'une lignée hESC rapportrice de l'état de pluripotence**

Lorsque j'ai commencé mon projet, une lignée de hESC appelée OSCAR avait été dérivée à partir d'un ICM d'embryon humain quelques années auparavant ([Chen et al, 2015](#)). Mon travail a consisté en la construction d'une lignée de hESC OSCAR marqueur de l'état de pluripotence, par l'introduction du gène de fluorescence Kate2 sous contrôle du promoteur du gène de pluripotence *OCT4*, cela afin de suivre l'évolution de la pluripotence des cellules. Pour créer ces cellules, j'ai utilisé la technologie CRISPR/Cas9 avec la méthode de la "double nickase".

Pour ce projet, j'ai obtenu 46 clones cellulaires résistants à l'antibiotique. La caractérisation par PCR m'a permis d'identifier 9 clones ayant intégrés la matrice de réparation et dont l'allèle sauvage n'a pas été muté. Cela me donne un taux relatif de 19,6% et un taux absolu de  $4,1 \cdot 10^{-7}$ . Zhu et collaborateurs ont publié des résultats sur l'introduction d'un gène fluorescent dont l'expression est contrôlée par le promoteur du gène de pluripotence *OCT4* dans des cellules souches embryonnaires humaines en utilisant la méthode de la nucléase. Ils ont obtenu 8 clones cellulaires fluorescents sur les 288 résistants à l'antibiotique. Ce résultats leurs donnent un taux de réussite relatif de 2,8% et un taux absolu de  $8 \cdot 10^{-7}$  ([Zhu et al, 2015](#)).

On remarque un nombre de clones cellulaires résistants à l'antibiotique plus important pour la méthode faisant agir la nucléase Cas9 que la méthode de la « double nickase ». Cela démontre l'efficacité de la nucléase Cas9. Cependant, la nucléase a induit 7 fois plus d'intégration de la cassette de sélection de façon aspécifique que la méthode de la "double nickase". Elle est donc moins spécifique que la méthode de la double "double nickase".

Dans un 2<sup>ème</sup> temps, j'ai excisé la cassette de sélection à la puromycine dont la séquence se situe entre le codon STOP du gène fusion *OCT4-Kate2* et la région 3'UTR du gène *OCT4* en utilisant le système Cre/Lox. Cela m'a permis d'observer une augmentation de l'intensité de fluorescence dans les cellules dont la cassette de sélection a été excisée. Ce phénomène d'inhibition de l'expression d'un gène causé par la présence de la cassette de sélection est retrouvé dans plusieurs publications ([Zhu et al, 2015](#) ; [Ema et al, 2006](#)). Il s'explique par le fait que le promoteur pGK est un promoteur très actif dans les cellules humaines. Il rentre en compétition dans le recrutement de l'ARN polymérase II avec les promoteurs endogène voisin. Ce qui à pour conséquence une diminution de la transcription des gènes voisins. Pour

les cellules OSCAR *OCT4-Kate2\_46* cela induit une diminution de l'expression du gène fusion *OCT4-Kate2* (*Pham et al, 1996*).

La principale source d'interrogation sur la technologie CRISPR/Cas9 est sa spécificité. En effet cette technologie est source de « off-targets ».

Pour éviter cela, lorsque j'ai conçu les sgRNAs qui ciblent la région autour du codon STOP du gène *OCT4*. J'ai observé leur score avec l'algorithme « <http://crispr.mit.edu/> ». Le score pour le couple de sgRNA était élevé (forte efficacité ; et faible « off-target » prédits). Récemment, j'ai refait le calcul pour vérifier les scores des sgRNA avec le même algorithme. Les scores ont diminué à cause des « off-targets ». J'ai vérifié les résultats en faisant un calcul avec un nouvel algorithme « CRISPOR ». Celui-ci confirme les résultats du premier. En effet, il trouve plusieurs centaines de « off-targets » prédits pour chacun des sgRNA. Ce changement brutal de score est dû à l'évolution des algorithmes. Ceux-ci sont de plus en plus puissants et donc précis.

Cependant, ce n'est pas parce que les guides que j'ai conçus ciblent potentiellement plusieurs centaines de régions génomiques de manière aspécifique, que la cassure simple brin créée par la Cas9\_D10A va induire une mutation. En effet la cellule répare de façon fidèle les cassures simple brin (*Dianov et Hübscher, 2013*).

Les cellules hESC OSCAR *OCT4-Kate2* étant créées, nous vérifierons l'absence de mutations aspécifiques en analysant le génome des cellules hESC OSCAR, qui me servira de génome de référence, et celui des cellules hESC OSCAR *OCT4-Kate2*, grâce à la technique « whole genome sequencing ».

Pour éviter l'effet des « off-targets », et améliorer le taux de recombinaison homologue de la cellule, j'aurais pu utiliser la Cpf1 qui est plus spécifique et efficace que la méthode de la « double nickase » grâce notamment aux sites PAM riches en T (*Zetsche et al, 2015*).

Par ailleurs, j'ai effectué une simulation de conception de sgRNA sur l'algorithme « CRISPOR » pour la nucléase Cpf1. Les résultats montrent des scores élevés. En effet, le nombre de "off-targets" prédits lorsque nous ciblons la région du codon STOP du gène *OCT4* varie entre 6 et 12.

Pour réduire le nombre de mutations non spécifiques, nous pourrions construire un vecteur codant pour la protéine Cpf1 sous le contrôle d'un promoteur inductible (*González et al,*

2014). Cela aura pour effet de contrôler la durée de l'expression de la nucléase Cpf1 afin de réduire sa durée d'action pendant laquelle l'ADN cellulaire peut être clivé.

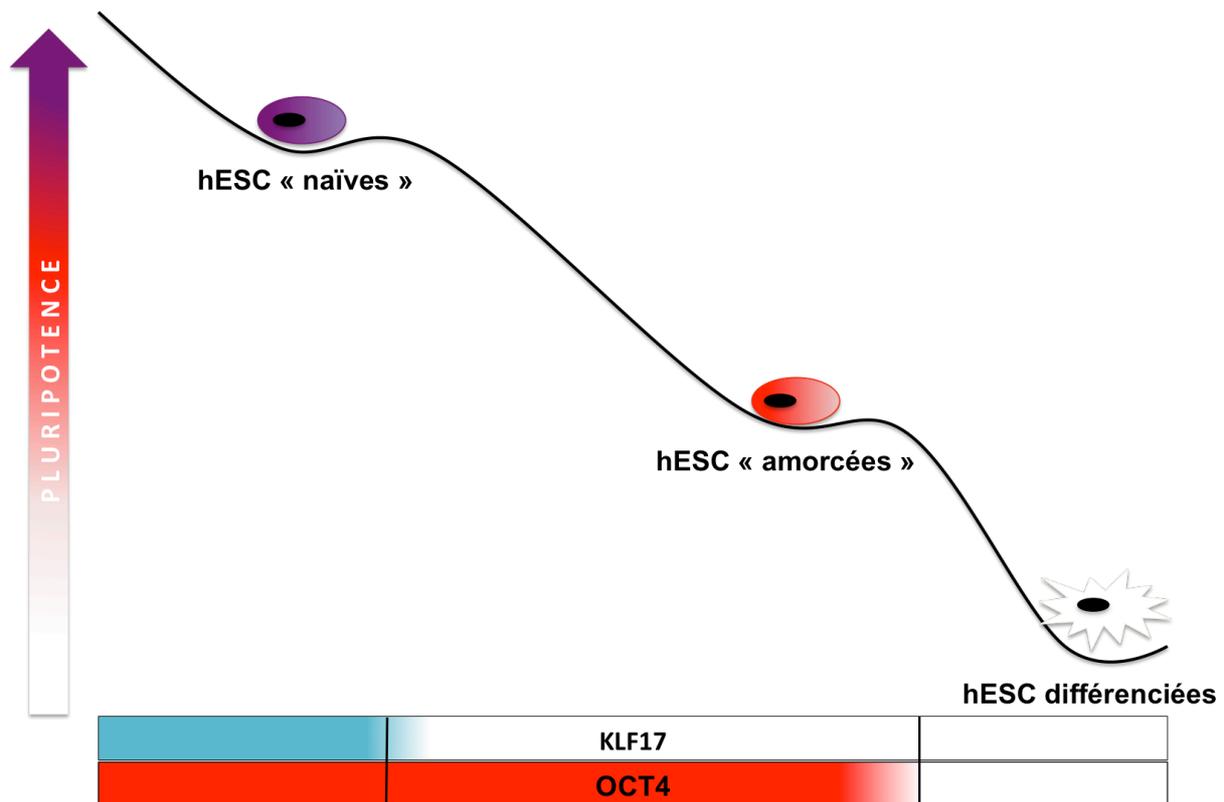
Enfin, des études ont montré que la cellule favorise le processus de recombinaison homologue quand elle est en phase G2/M du cycle cellulaire (Heyer et al, 2010). Pour favoriser l'intégration de la matrice de réparation dans le génome, nous pourrions enrichir la population cellulaire en phase G2/M en ajoutant du nocodazole 16h avant la transfection. (Yang et al, 2016, Becker et al, 2006).

Toutes ces modifications dans le protocole seront appliquées pour effectuer une nouvelle intégration de gène dans les cellules OSCAR *OCT4*-Kate2-Cre+.

## **II. Perspectives**

Afin de poursuivre la caractérisation des cellules OSCAR *OCT4*-Kate2-Cre+, il est essentiel de trier la population dont l'expression de Kate2 est la plus intense. Ensuite, nous contrôlerons que les cellules ne possèdent plus la cassette de sélection par PCR et par cytométrie en flux. Enfin, l'absence de modification chromosomique sera vérifiée en effectuant un caryotype. Les caractères de pluripotence des cellules OSCAR *OCT4*-Kate2 seront contrôlés par des analyses transcriptomiques, des immuno-marquages sur des protéines de pluripotence et sur la formation de tératomes.

Une fois les cellules hESC OSCAR *OCT4*-Kat caractérisées, nous les utiliserons pour effectuer une insertion d'un second gène de fluorescence (Turquoise 2) à la place du codon STOP d'un gène de pluripotence « naïf ». Le gène choisi est *KLF17*. Il est exprimé dans des cellules de l'embryon humain au stade morula et dans des cellules en culture qui sont dites « naïves ». L'intérêt de fabriquer une lignée cellulaire possédant deux marqueurs de pluripotence réside dans le fait que nous pourrions suivre l'évolution de la pluripotence de cellules en culture. Dans cette voie, nous pourrions tester des dizaines de conditions de cultures et connaître l'état de pluripotence des cellules en les observant par fluorescence. Ainsi, les cellules hESC différenciées n'exprimeront aucun marqueur fluorescent. Les cellules hESC qui s'auto-renouvelleront à l'état « amorcé » exprimeront seulement le marqueur Kate2. Les cellules hESC reprogrammées qui s'auto-renouvellent à l'état « naïf » exprimeront les marqueurs Turquoise 2 et Kate2. Nous pourrions alors sélectionner une condition de culture pour laquelle les cellules expriment les gènes de fluorescences Turquoise2 et Kate2 (Figure 27).



**Figure 27 : Profils de pluripotence des cellules hESC OSCAR *OCT4*-Kate2-KLF17-Turquoise2 en fonction de leur fluorescence.** Les cellules hESC différenciées n'expriment ni le gène *OCT4* ni le gène *KLF17*. Les cellules hESC qui s'auto-renouvellent à l'état « amorcé » expriment seulement le gène *OCT4*. Les cellules hESC reprogrammées qui s'auto-renouvellent à l'état « naïf » expriment les le gènes *OCT4* et *KLF17*.

### **III. Conclusion personnelle**

Le projet EPHE que j'ai réalisé dans l'équipe de Pierre Savatier m'a permis, d'un point de vue technique, de développer mes compétences en culture cellulaire et en biologie moléculaire. J'ai par ailleurs beaucoup appris sur la technologie CRISPR/Cas9. D'un point de vue plus personnel, j'ai intégré une équipe qui, en plus d'être compétente et reconnue dans son domaine, est agréable et travaille dans une ambiance joviale. Le travail d'équipe est de rigueur, ce qui favorise les échanges d'idées et permet de résoudre des problèmes rapidement. J'ai beaucoup appris au contact de chaque personne de l'équipe. Je leur souhaite le meilleur pour la suite, avec l'envie et l'espoir de revenir, un jour, travailler avec eux.

# Bibliographie

---

- Amit, M., Carpenter, M.K., Inokuma, M.S., Chiu, C.-P., Harris, C.P., Waknitz, M.A., Itskovitz-Eldor, J., and Thomson, J.A. (2000). Clonally Derived Human Embryonic Stem Cell Lines Maintain Pluripotency and Proliferative Potential for Prolonged Periods of Culture. *Dev. Biol.* *227*, 271–278.
- Armstrong, L., Lako, M., Lincoln, J., Cairns, P.M., and Hole, N. (2000). mTert expression correlates with telomerase activity during the differentiation of murine embryonic stem cells. *Mech. Dev.* *97*, 109–116.
- Babaie, Y., Herwig, R., Greber, B., Brink, T.C., Wruck, W., Groth, D., Lehrach, H., Burdon, T., and Adjaye, J. (2007). Analysis of Oct4-Dependent Transcriptional Networks Regulating Self-Renewal and Pluripotency in Human Embryonic Stem Cells. *Stem Cells* *25*, 500–510.
- Bajar, B.T., Lam, A.J., Badiie, R.K., Oh, Y.-H., Chu, J., Zhou, X.X., Kim, N., Kim, B.B., Chung, M., Yablonovitch, A.L., et al. (2016). Fluorescent indicators for simultaneous reporting of all four cell cycle phases. *Nat. Methods* *13*, 993–996.
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science* (80-. ). *315*, 1709–1712.
- Becker, K.A., Ghule, P.N., Therrien, J.A., Lian, J.B., Stein, J.L., Van Wijnen, A.J., and Stein, G.S. (2006). Self-renewal of human embryonic stem cells is supported by a shortened G1 cell cycle phase. *J. Cell. Physiol.* *209*, 883–893.
- Brons, I.G.M., Smithers, L.E., Trotter, M.W.B., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A., et al. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* *448*, 191–195.
- Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* *113*, 643–655.
- Chen, H., Aksoy, I., Gonnot, F., Osteil, P., Aubry, M., Hamela, C., Rognard, C., Hochard, A., Voisin, S., Fontaine, E., et al. (2015). Reinforcement of STAT3 activity reprogrammes human embryonic stem cells to naive-like pluripotency. *Nat. Commun.* *6*, 7095.
- Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J., and Voytas, D.F. (2010). Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* *186*, 756–761.
- Chu, V.T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., and Kühn, R. (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mamm. *Nat. Biotechnol.* *33*, 543–548.
- Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex genome engineering using CRISPR/Cas systems. *Science* *339*, 819–823.
- Coronado, D., Godet, M., Bourillot, P.Y., Tapponnier, Y., Bernat, A., Petit, M., Afanassieff, M., Markossian, S., Malashicheva, A., Iacone, R., et al. (2013). A short G1 phase is an intrinsic determinant of naïve embryonic stem cell pluripotency. *Stem Cell Res.* *10*, 118–131.

Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E. (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 471, 602–607.

Dianov, G.L., and Hübscher, U. (2013). Mammalian base excision repair: The forgotten archangel. *Nucleic Acids Res.* 41, 3483–3490.

Dong, D., Guo, M., Wang, S., Zhu, Y., Wang, S., Xiong, Z., Yang, J., Xu, Z., and Huang, Z. (2017). Structural basis of CRISPR–SpyCas9 inhibition by an anti-CRISPR protein. *Nature* 546, 436–439.

Ema, M., Takahashi, S., and Rossant, J. (2006). Deletion of the selection cassette, but not cis-acting elements, in targeted Flk1-lacZ allele reveals Flk1 expression in multipotent mesodermal progenitors. *Blood* 107, 111–117.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292, 154–156.

Fong, H., Wong, R.C.-B., and Donovan, P.J. (2012). Transcriptional regulation of TRKC by SOX2 in human embryonic stem cells. *Stem Cell Res.* 8, 206–214.

Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., and Sander, J.D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* 31, 822–826.

Gao, Y., Wu, H., Wang, Y., Liu, X., Chen, L., Li, Q., Cui, C., Liu, X., Zhang, J., and Zhang, Y. (2017). Single Cas9 nickase induced generation of NRAMP1 knockin cattle with reduced off-target effects. *Genome Biol.* 18, 13.

Garneau, J.E., Dupuis, M.-È., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71.

González, F., Zhu, Z., Shi, Z.-D., Lelli, K., Verma, N., Li, Q. V, and Huangfu, D. (2014). An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. *Cell Stem Cell* 15, 215–226.

Hansis, C; Grifo, J A; Krey, L.C., Hansis, C., Grifo, J. a, and Krey, L.C. (2000). Oct-4 expression in inner cell mass and trophectoderm of human blastocysts. *Mol. Hum. Reprod.* 6, 999–1004.

Hockemeyer, D., Wang, H., Kiani, S., Lai, C.S., Gao, Q., Cassady, J.P., Cost, G.J., Zhang, L., Santiago, Y., Miller, J.C., et al. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. *Nat. Biotechnol.* 29, 731–734. Heyer, W.-D., Ehmsen, K.T., and Liu, J. (2010). Regulation of Homologous Recombination in Eukaryotes. *Annu. Rev. Genet.* 44, 113–139.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31, 827–832.

Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. *Cell* 157, 1262–1278.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in

- Escherichia coli*, and identification of the gene product. *J. Bacteriol.* *169*, 5429–5433.
- Jansen, R., Embden, J.D., Gaastra, W., and Schouls, L.M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* *43*, 1565–1575.
- Jayathilaka, K., Sheridan, S.D., Bold, T.D., Bochenska, K., Logan, H.L., Weichselbaum, R.R., Bishop, D.K., and Connell, P.P. (2008). A chemical compound that stimulates the human homologous recombination protein RAD51. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 15848–15853.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA – Guided. *337*, 816–822.
- Jinek, M., Jiang, F., Taylor, D.W., Sternberg, S.H., Kaya, E., Ma, E., Anders, C., Hauer, M., Zhou, K., Lin, S., et al. (2014). Structures of Cas9 Endonucleases Reveal RNA-Mediated Conformational Activation. *Science* (80- ). *343*, 1247997–1247997.
- Karvelis, T., Gasiunas, G., and Siksnys, V. (2013). Programmable DNA cleavage *in vitro* by Cas9. *Biochem. Soc. Trans.* *41*, 1401–1406.
- Kass, E.M., Helgadottir, H.R., Chen, C.-C., Barbera, M., Wang, R., Westermark, U.K., Ludwig, T., Moynahan, M.E., and Jasin, M. (2013). Double-strand break repair by homologous recombination in primary mouse somatic cells requires BRCA1 but not the ATM kinase. *Proc. Natl. Acad. Sci.* *110*, 5564–5569.
- Keller, G.M. (1995). *In vitro* differentiation of embryonic stem cells. *Curr. Opin. Cell Biol.* *7*, 862–869.
- Kim, S. (2014). Supplemental Information\_Highly Efficient RNA-guide genome editing... *Genome Res.* *128*, 1–32.
- Kleinstiver, B.P., Tsai, S.Q., Prew, M.S., Nhu, T., Welch, M.M., Lopez, J.M., Mccaw, Z.R., Martin, J., Joung, J.K., Unit, P., et al. (2016). HHS Public Access. *34*, 869–874.
- Kuscu, C., Arslan, S., Singh, R., Thorpe, J., and Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat. Biotechnol.* *32*, 677–683.
- Lengner, C.J., Gimelbrant, A.A., Erwin, J.A., Cheng, A.W., Guenther, M.G., Welstead, G.G., Alagappan, R., Frampton, G.M., Xu, P., Muffat, J., et al. (2010). Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. *Cell* *141*, 872–883.
- Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. *Science* *339*, 823–826.
- Martin, G.R., and Evans, M.J. (1974). The morphology and growth of a pluripotent teratocarcinoma cell line and its derivatives in tissue culture. *Cell* *2*, 163–172.
- Merkle, F.T., Neuhausser, W.M., Santos, D., Valen, E., Gagnon, J.A., Maas, K., Sandoe, J., Schier, A.F., and Eggan, K. (2015). Efficient CRISPR-Cas9-Mediated Generation of Knockin Human Pluripotent Stem Cells Lacking Undesired Mutations at the Targeted Locus. *Cell Rep.* *11*, 875–883.
- Miyaoka, Y., Berman, J.R., Cooper, S.B., Mayerl, S.J., Chan, A.H., Zhang, B., Karlin-Neumann, G.A., and Conklin, B.R. (2016). Systematic quantification of HDR and NHEJ reveals effects of locus, nuclease, and cell type on genome-editing. *Sci. Rep.* *6*, 23549.

- Mojica, F.J.M., Díez-Villaseñor, C., Soria, E., and Juez, G. (2000). Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol. Microbiol.* *36*, 244–246.
- Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* *123*, 917–929.
- Pawluk, A., Amrani, N., Zhang, Y., Garcia, B., Hidalgo-Reyes, Y., Lee, J., Edraki, A., Shah, M., Sontheimer, E.J., Maxwell, K.L., et al. (2016). Naturally Occurring Off-Switches for CRISPR-Cas9. *Cell* *167*, 1829–1838.e9.
- Pesce, M., Gross, M.K., and Scholer, H.R. (1998). In line with our ancestors: Oct-4 and the mammalian germ. *Bioessays* *20*, 722–732.
- Pham, C.T., MacIvor, D.M., Hug, B.A., Heusel, J.W., and Ley, T.J. (1996). Long-range disruption of gene expression by a selectable marker cassette. *Proc. Natl. Acad. Sci. U. S. A.* *93*, 13090–13095.
- Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., et al. (2013). Double nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity. *Cell* *154*, 1380–1389.
- Rong, Z., Zhu, S., Xu, Y., and Fu, X. (2014). Homologous recombination in human embryonic stem cells using CRISPR/Cas9 nickase and a long DNA donor template. *Protein Cell* *5*, 258–260.
- Sakuma, T., Nishikawa, A., Kume, S., Chayama, K., and Yamamoto, T. (2015). Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. *Sci. Rep.* *4*, 5400.
- Satomura, A., Nishioka, R., Mori, H., Sato, K., Kuroda, K., and Ueda, M. (2017). Precise genome-wide base editing by the CRISPR Nickase system in yeast. *Sci. Rep.* *7*, 2095.
- Savatier, P., Huang, S., Szekely, L., Wiman, K.G., and Samarut, J. (1994). Contrasting patterns of retinoblastoma protein expression in mouse embryonic stem cells and embryonic fibroblasts. *Oncogene* *9*, 809–818.
- Savatier, P., Lapillonne, H., van Grunsven, L.A., Rudkin, B.B., and Samarut, J. (1996). Withdrawal of differentiation inhibitory activity/leukemia inhibitory factor up-regulates D-type cyclins and cyclin-dependent kinase inhibitors in mouse embryonic stem cells. *Oncogene* *12*, 309–322.
- Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* *336*, 688–690.
- Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* *336*, 688–690.
- Smith, J., Bibikova, M., Whitby, F.G., Reddy, A.R., Chandrasegaran, S., and Carroll, D. (2000). Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res.* *28*, 3361–3369.
- Song, J., Yang, D., Xu, J., Zhu, T., Chen, Y.E., and Zhang, J. (2016). RS-1 enhances

CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nat. Commun.* *7*, 10548.

Takeda, J., Seino, S., and Bell, G.I. (1992). Human Oct3 gene family: CDNA sequences, alternative splicing, gene organization, chromosomal location, and expression at low levels in adult tissues. *Nucleic Acids Res.* *20*, 4613–4620.

Thomson, J.A. (1998). Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science* (80-. ). *282*, 1145–1147.

Williams, R.L., Hilton, D.J., Pease, S., Willson, T. a, Stewart, C.L., Gearing, D.P., Wagner, E.F., Metcalf, D., Nicola, N. a, and Gough, N.M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* *336*, 684–687.

Wilson, T.E., Grawunder, U., and Lieber, M.R. (1997). Yeast DNA ligase IV mediates non-homologous DNA end joining. *Nature* *388*, 495–498.

Xu, C., Inokuma, M.S., Denham, J., Golds, K., Kundu, P., Gold, J.D., and Carpenter, M.K. (2001). Feeder-free growth of undifferentiated human embryonic stem cells. *Nat. Biotechnol.* *19*, 971–974.

Yang, D., Scavuzzo, M.A., Chmielowiec, J., Sharp, R., Bajic, A., and Borowiak, M. (2016). Enrichment of G2/M cell cycle phase in human pluripotent stem cells enhances HDR-mediated gene repair with customizable endonucleases. *Sci. Rep.* *6*, 21264.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., Van Der Oost, J., Regev, A., et al. (2015). Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell*.

Zhu, Z., Verma, N., González, F., Shi, Z.-D., and Huangfu, D. (2015). A CRISPR/Cas-Mediated Selection-free Knockin Strategy in Human Embryonic Stem Cells. *Stem Cell Reports* *4*, 1103–1111.

# Annexes

Annexe 1 : Amorces PCR pour la caractérisation des plasmides pX330A\_D10A et pX330S2

Noms	Séquences
pBR322_For	5'-GGGAAACGCCTGGTATCTTT-3'
CBA_Rev	5'-GAGATGGGGAGAGTGAAGCA-3'

Annexe 2 : Amorces PCR pour la caractérisation des cellules OSCAR OCT4-Kate2-Puro

Gènes	Identification des clones	Séquences	Tm (°C)	Taille Amplicon (pb)
GFP	Positifs en 5'	OCT4g For : 5'-AGTCCAAAGCTTGCCCTTGTCACC-3' GFP Rev: 5'-AGCTCCTCGCCCTTGCTCACC-3'	72	811
	Positifs en 3'	Mat-3' For : 5'-CCCTCGAAGAGGTTCACTAGG-3' OCT4g-3' Rev : 5'-TGAAGTGTGGCAGAGCCAGG-3'	68	756
Kate2	Positifs en 5'	OCT4g For : 5'-AGTCCAAAGCTTGCCCTTGTCACC-3' Kate2 Rev : 5'-GGCCCACCTCCACCACTTTG-3'	72	949
	Positifs en 3'	Mat-3' For : 5'-CCCTCGAAGAGGTTCACTAGG-3' OCT4g-3' Rev : 5'-TGAAGTGTGGCAGAGCCAGG-3'	68	756
	Hétérozygotes	OCT4g For : 5'-AGTCCAAAGCTTGCCCTTGTCACC-3' OCT4 Rev : 5'-CCTGGCACAACTCCAGG-3'	67	804
	Sans cassette de sélection	Kate2 For : 5'-GCCCGTCATGCAAAAAGAAGAC-3' OCT4 Rev : 5'-CCTGGCACAACTCCAGG-3'	68	470

ECOLE PRATIQUE DES HAUTES ETUDES

SCIENCES DE LA VIE ET DE LA TERRE

Fabrication de cellules souches embryonnaires humaines exprimant un marqueur fluorescent de la pluripotence grâce à la technologie CRISPR/Cas9

Antoine Cadiou

15 décembre 2017

RESUME

Le développement de nouvelles technologies comme CRISPR/Cas9 facilite la création de mutations ciblées dans le génome des cellules souches embryonnaires humaines. Cette technologie, découverte et caractérisée en 2012, ne cesse de se développer. Au départ, il existait uniquement la Cas9 sauvage qui crée une cassure double-brin à bouts francs dans la cellule. Très vite, la méthode de la Cas9\_D10A (double nickase) s'est développée pour palier au manque de spécificité de la précédente méthode. Enfin la nucléase Cpf1 a été ensuite découverte et est utilisée pour effectuer des modifications génétiques. La nucléase Cpf1 est plus spécifique et efficace que les précédentes méthodes. Je rapporte dans ce mémoire la création d'une lignée de cellules souches embryonnaires humaines dont le gène de pluripotence *OCT4* est fusionné au gène *Katushka* par la méthode de la « double nickase » de la technologie CRISPR/Cas9. Pour cela j'ai construit le vecteur permettant de créer la cassure double brin dans le locus du gène *OCT4* et j'ai modifié la matrice de réparation contenant le gène *Katushka* ainsi que la cassette de sélection à un antibiotique. L'analyse des clones cellulaires se fait en deux étapes. Cela implique, premièrement la sélection des clones résistants à l'antibiotique et l'identification des clones correctement ciblés, deuxièmement l'analyse de l'expression de la protéine fluorescente par microscopie et cytométrie en flux, avant et après l'excision de la cassette de sélection. Enfin, je présente en perspective les expériences à réaliser pour arriver à obtenir des cellules souches embryonnaires humaines exprimant un marqueur fluorescent de la pluripotence dite « naïve ». Ces expériences demanderont de réaliser une nouvelle intégration d'un gène de fluorescence dans le locus d'un gène spécifique de l'état naïf. Cette modification génomique fera également intervenir la technologie CRISPR/Cas9.

MOTS-CLES : Cellules souches embryonnaires, CRISPR/Cas9, Knock-In, *Katushka*, *OCT4*, humain