



**HAL**  
open science

# Modélisation de l'épithélium bronchique humain par la technologie des cellules souches pluripotentes induites (iPS)

Caroline Sansac

► **To cite this version:**

Caroline Sansac. Modélisation de l'épithélium bronchique humain par la technologie des cellules souches pluripotentes induites (iPS). Médecine humaine et pathologie. Université Montpellier, 2016. Français. NNT : 2016MONTT014 . tel-01510250

**HAL Id: tel-01510250**

**<https://theses.hal.science/tel-01510250>**

Submitted on 19 Apr 2017

**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.

# THÈSE

Pour obtenir le grade de  
**Docteur**

Délivré par l'**UNIVERSITE DE MONTPELLIER**

Préparée au sein de l'école doctorale : **Sciences  
Chimiques et Biologiques pour la Santé**  
Et de l'unité de recherche : **IRMB – INSERM 1183**

Spécialité : **Biologie – Santé**

Présentée par **Caroline SANSAC**

**MODELISATION DE L'EPITHELIUM  
BRONCHIQUE HUMAIN PAR LA  
TECHNOLOGIE DES CELLULES SOUCHES  
PLURIPOTENTES INDUITES (iPS)**

Soutenue le 18 Octobre 2016 devant le jury composé de

Pr. John DE VOS, PUPH, Université de Montpellier  
Pr. Pascal CHANEZ, PUPH, Université d'Aix-Marseille  
Dr. Nathalie LEFORT, IR, iPS Plateforme IMAGINE  
Pr. Arnaud BOURDIN, PUPH, Université de Montpellier

Directeur de thèse  
Rapporteur  
Rapporteur  
Président du jury



IRMB, INSERM U 1183  
Hôpital St Eloi  
80 avenue Augustin Fliche  
34295 Montpellier, Cedex 5  
FRANCE

# THÈSE

Pour obtenir le grade de  
**Docteur**

Délivré par l'**UNIVERSITE DE MONTPELLIER**

Préparée au sein de l'école doctorale : **Sciences  
Chimiques et Biologiques pour la Santé**  
Et de l'unité de recherche : **IRMB – INSERM 1183**

Spécialité : **Biologie – Santé**

Présentée par **Caroline SANSAC**

**MODELISATION DE L'EPITHELIUM  
BRONCHIQUE HUMAIN PAR LA  
TECHNOLOGIE DES CELLULES SOUCHES  
PLURIPOTENTES INDUITES (iPS)**

Soutenue le 18 Octobre 2016 devant le jury composé de

Pr. John DE VOS, PUPH, Université de Montpellier  
Pr. Pascal CHANEZ, PUPH, Université d'Aix-Marseille  
Dr. Nathalie LEFORT, IR, iPS Plateforme IMAGINE  
Pr. Arnaud BOURDIN, PUPH, Université de Montpellier

Directeur de thèse  
Rapporteur  
Rapporteur  
Président du jury



“ Let it go, let it go. ”

*Elsa*

# Remerciements

**A mon jury de thèse :**

**Monsieur le Professeur Arnaud Bourdin**, Arnaud, merci de m'avoir suivie au cours de ces années, d'avoir participé aux comités, de m'avoir encouragée et même donné des fibroblastes pour m'entraîner ! Bref, merci.

**Monsieur le Professeur John De Vos**, Lavoisier disait : « Rien ne perd, rien ne se crée, tout se transforme », De Vos dira « Tout se reprogramme, tout se différencie », j'ajouterai « Enfin presque ».

**Monsieur le Professeur Pascal Chanez**, je tiens à vous assurer de ma profonde reconnaissance pour l'intérêt que vous portez à ce travail et pour votre participation à mon jury de thèse en qualité de rapporteur.

**Madame le Docteur Nathalie Lefort**, je vous remercie de me faire l'honneur de participer à ce jury et de juger mon travail. Veuillez trouver ici le témoignage de mon profond respect et ma sincère reconnaissance.

### **A mes trois co-internes :** parce que jamais deux sans trois

Olive (Loulou, impossible de me souvenir de toi avant STE, et comme tu l'as bien résumé : une très belle amitié, et quelle complicité !), mon Juju (il m'a d'abord fallu décoder le phrasé Maxi-BG avant de comprendre quel homme formidable tu es. Tic Tac...) et Engi (dernière arrivée et tu prends une place non négligeable, tu es quelqu'un de merveilleux petit chat, Mwahhhh)

### **A mes amis :**

Ma Caca, pour les fous rires, les larmes, les joies, les peines. Merci d'être la sœur que je n'ai pas eu. Spéciale dédicace à Golette et à Nounette (en police 72 comme promis).

En fait c'est vrai, les remerciements sont ce qu'il y a de plus difficile, alors comme d'habitude : bref, court et sans tourner autour du pot.

Rencontrés au travail, pendant l'internat ou à l'internat, d'ici ou d'ailleurs, de près ou de plus loin, d'hier, de maintenant et surtout de demain : Merci à tous de faire partie de ma vie.

Cosette (P2-power), Popo (aux quatre coins du monde ou presque), Delinger (STEPPPPPPP) et son Baba, Olga un poka de temps en temps, Anne-So.A (très jolie rencontre), Kamel et Perrine + Maëlle, Safia et Mika, Sonia, Anne-So, Virginie et Fabrice + Anoukette et Sibylle, , Syssy, Régine et Soso, Amadou, Patricia, Philippe, Pierre, Chris, SC, Carole et Cécile, Elenita et Alan + Dianita, Jean-Yves et Caroline + Maxime, Jean-François, Pascal, Boune et Elsa, Diane, Qiang + Ling, Fabienne.

To my foreigners friends met here in France, thanks for teaching me how to taste a good wine ☺. And for those that I've met overseas thanks for looking after me, I've felt like home there and it was good fun.

### **A mes collègues :**

De la team IRB : Said pourquoi n'es tu pas arrivé avant ? et tous les stagiaires de passage(s), Noémie, Nicolas, Cycy, David, ce fut un grand plaisir de vous encadrer

De la team ADV : Arnaud, Isabelle, Aurélie, Lulu (Ahhhh ma lulu, on s'est pas assez croisées pendant ces quatre ans mais, femelle, tu es exceptionnelle, ne change pas !), Sophichou.

### **Enfin à ma famille :**

Dieu sait qu'elle s'est réduite au cours de ces dernières années, et pourtant je reste chanceuse de vous avoir auprès de moi. Encouragée, aimée et supportée, après 35 000 ans d'études ça y est je suis (enfin) une adulte. Je ne le dis ni le montre jamais, parce que chez nous c'est comme ça on dit rien, au mieux on l'écrit, alors : j'vous aime très fort et Merci.





# Table des matières

<b>TABLE DES MATIERES.....</b>	<b>7</b>
<b>LISTE DES ABREVIATIONS.....</b>	<b>10</b>
<b>LISTE DES FIGURES .....</b>	<b>12</b>
<b>LISTE DES TABLES.....</b>	<b>15</b>
<b>INTRODUCTION .....</b>	<b>16</b>
<b>PARTIE 1 : LE POUMON.....</b>	<b>19</b>
<b>1. LE DEVELOPPEMENT DU SYSTEME RESPIRATOIRE.....</b>	<b>20</b>
<b>2. LE POUMON ADULTE.....</b>	<b>24</b>
2.1 ANATOMIE – PHYSIOLOGIE.....	24
2.2 HISTOLOGIE .....	27
<b>PARTIE 2 : LES CELLULES SOUCHES PLURIPOTENTES .....</b>	<b>30</b>
<b>1. LA PLURIPOTENCE .....</b>	<b>31</b>
1.1 GENERALITES.....	31
1.2 DECOUVERTE.....	31
1.3 PROPRIETES.....	32
1.3.1 <i>La morphologie des cellules.</i> .....	33
1.3.2 <i>Le comportement de ces cellules.</i> .....	33
1.3.3 <i>Les activités enzymatiques de ces cellules</i> .....	34
1.3.4 <i>Le réseau de facteurs de transcription de la pluripotence</i> .....	34
1.3.5 <i>Les facteurs extrinsèques de la pluripotence</i> .....	36
1.3.6 <i>L’expression de marqueurs membranaires de la pluripotence</i> .....	37
1.3.7 <i>La formation de corps embryoïdes, ou embryoid bodies (EB)</i> .....	38
1.3.8 <i>La formation de tératomes dans un animal immunodéprimé</i> .....	39
<b>2. CELLULES SOUCHES PLURIPOTENTES EMBRYONNAIRES.....</b>	<b>39</b>
<b>3. INDUCED PLURIPOTENT STEM CELLS : IPS.....</b>	<b>44</b>
3.1 LES PRECURSEURS DE LA REPROGRAMMATION .....	45
3.2 DECOUVERTE DES IPS.....	45
3.3 ROLE DES FACTEURS DE TRANSCRIPTION DANS LA REPROGRAMMATION .....	49
3.4 LE DEROULEMENT MOLECULAIRE DE LA REPROGRAMMATION .....	50

3.5	DIFFERENTES ESPECES REPROGRAMMEES.....	51
3.6	DIFFERENTS COCKTAILS DE REPROGRAMMATION.....	53
3.7	DIFFERENTES METHODES DE REPROGRAMMATION.....	57
3.8	DIFFERENTES CELLULES A REPROGRAMMER.....	59
3.9	LES APPLICATIONS DES IPS.....	62
<b>4.</b>	<b>TECHNIQUES DE CULTURE.....</b>	<b>64</b>
4.1	LES MILIEUX DE CULTURE.....	64
4.2	LES FEEDERS ET LES MATRICES EXTRACELLULAIRES.....	65
4.3	LES FACTEURS DE CROISSANCE.....	65
<b>5.</b>	<b>LIMITES DES IPS.....</b>	<b>66</b>
5.1	DIFFERENCES ENTRE IPS ET ES.....	66
5.2	DIFFICULTES DE LA MISE EN PLACE DES CONDITIONS DE PRODUCTION DE GRADE CLINIQUE.....	67
5.3	L'APPARITION D'ANOMALIES GENOMIQUES.....	69
	<b>PARTIE 3 : LA DIFFERENCIATION.....</b>	<b>71</b>
<b>1.</b>	<b>DIFFERENCIATION NON DIRIGEE.....</b>	<b>72</b>
1.1	LE CORPS EMBRYOÏDE.....	73
1.2	LE TERATOME.....	74
1.2.1	<i>Tératomes et tératocarcinomes en clinique.....</i>	<i>75</i>
1.2.2	<i>Tératomes dans l'animal immunosupprimé.....</i>	<i>76</i>
1.2.3	<i>Tissus obtenus dans les tératomes expérimentaux.....</i>	<i>78</i>
<b>2.</b>	<b>DIFFERENCIATION DIRIGEE.....</b>	<b>80</b>
2.1	MODELISER DES TISSUS <i>IN VITRO</i> .....	80
2.2	MODELISER L'EPITHELIUM BRONCHIQUE.....	83
	<b>PARTIE 4 : LES APPLICATIONS MEDICALES.....</b>	<b>87</b>
<b>1.</b>	<b>MEDECINE REGENERATRICE.....</b>	<b>88</b>
1.1	LES MEDICAMENTS DE THERAPIE INNOVANTE.....	89
1.2	DES MTI A BASE D'IPS.....	91
1.3	MEDECINE REGENERATRICE POUR LE POUMON.....	91
1.4	LE PREMIER ESSAI CLINIQUE UTILISANT DES IPS.....	93
<b>2.</b>	<b>MODELISATION DES TISSUS HUMAINS POUR L'INDUSTRIE PHARMACEUTIQUE.....</b>	<b>94</b>
2.1	MODELISER DES PATHOLOGIES.....	94
2.2	PRODUIRE DES ORGANES.....	95

2.3	DECOUVRIR DE NOUVEAUX MEDICAMENTS .....	97
2.4	DEVELOPPER LA PHARMACOGENOMIQUE .....	98
<b>RESULTATS .....</b>		<b>100</b>
<b>1.</b>	<b>LES TERATOMES .....</b>	<b>102</b>
1.1	RESUME .....	102
1.2	ARTICLE.....	103
1.3	DISCUSSION .....	104
<b>2.</b>	<b>LA DIFFERENCIATION DIRIGEE EN EPITHELIUM BRONCHIQUE .....</b>	<b>105</b>
2.1	PRODUCTION D'UNE LIGNEE D'IPS D'UN PATIENT ATTEINT DE DYSKINESIE PULMONAIRE .....	105
2.1.1	<i>La dyskinesie ciliaire primitive.....</i>	<i>105</i>
2.1.2	<i>Dérivation de deux lignées de fibroblastes cutanés et d'une lignée iPS de patients atteints de DCP .....</i>	<i>107</i>
2.2	PROTOCOLES DE DIFFERENCIATION.....	112
2.2.1	<i>Identification des marqueurs à suivre.....</i>	<i>112</i>
2.2.2	<i>Milieus .....</i>	<i>115</i>
2.2.3	<i>Matrices .....</i>	<i>118</i>
2.2.4	<i>Facteurs de croissance .....</i>	<i>118</i>
2.2.5	<i>Densité cellulaire et passage.....</i>	<i>125</i>
2.3	DISCUSSION .....	128
<b>DISCUSSION.....</b>		<b>133</b>
<b>1.</b>	<b>PERSPECTIVES .....</b>	<b>134</b>
<b>2.</b>	<b>LIMITES .....</b>	<b>135</b>
<b>3.</b>	<b>ALTERNATIVE .....</b>	<b>136</b>
<b>BIBLIOGRAPHIE.....</b>		<b>137</b>

# Liste des abréviations

ABM : agence de la biomédecine  
AFE : anterior foregut endoderm, intestin primitif antérieur  
ALI : air liquid interface, interface air-liquide  
ANSM : agence nationale de sécurité du médicament  
AP : phosphatase alcaline  
ATMP : advanced therapy medicinal product  
ATRA : acide tout trans rétinolique  
AZA : 5-azacytidine  
BMP : bone morphogenic protein  
BPCO : bronchopathie chronique obstructive  
BPF : bonnes pratiques de fabrication  
CCSP : club cell secretory protein  
CFTR : cystic fibrosis transmembrane conductance receptor  
CGP : cellule germinale primordiale  
CK : cytokératine  
CPM : carboxypeptidase M  
CSP : cellule souche pluripotente  
DCP : dyskinésie ciliaire primitive  
DE : definitive endoderm, endoderme définitif  
DPI : dépistage préimplantatoire  
EB : embryoid body, corps embryoïde  
EC : carcinome embryonnaire  
EMA : agence européenne du médicament  
EpCAM : epithelial calmoduline  
ES : cellule souche embryonnaire  
FGF : fibroblast growth factor  
FOX : forkhead box protein  
h : human, humain  
hFF : human foreskin fibroblast, fibroblaste de prépuce humain  
HLO : human lung organoid, organoïde humain de poumon  
HMG : high mobility group  
ICM : inner cell mass, masse cellulaire interne  
IL2RG : récepteur gamma de l'interleukine 2  
iPS : induced pluripotent stem cell, cellule souche pluripotente induite  
KLF : krüppel-like  
KO-SR : knockout serum replacement  
LIF : leukemia inhibitory factor  
lnc RNA : long non coding RNA, ARN long non codant  
m : mouse, souris  
MEC : matrice extracellulaire  
MEF : mouse embryonic fibroblast, fibroblaste embryonnaire murin  
MTI : médicament de thérapie innovante

MUC5AC : mucin 5AC  
NK : natural killer  
NKX2.1 : NK2 homebox 1  
NOD : souris diabétique non obèse  
NSG : NOD-SCID-Gamma  
OSKM : OCT4 SOX2 KLF4 c-MYC  
OSLN : OCT4 SOX2 LIN28 NANOG  
PL : lung progenitor, progéniteur bronchique  
RPE : épithélium pigmentaire de rétine  
SAG : agoniste de smoothed  
SAHA : suberoylanilide hydroxamic acid  
SCGB1A1 : secretoglobin 1A1  
SCID : immunodéficience sévère combinée  
SCNT : somatic cell nucleus transfert, transfert de noyau de cellules somatiques  
SeV : virus sendai  
SFTPC : surfactant protein C  
SHH : sonic hedgehog  
SOX : SRY-related HMG-box  
SRY : sex determining region on Y  
SSEA : stage-specific embryonic antigen  
SVF : serum veau foetal  
TGF- $\beta$  : transforming growth factor- $\beta$   
TSA : trichostatin A  
TTF1 : thyroid transcription factor 1  
TubIV : tubuline IV  
WKFBE : WNT3a KGF FGF10 BMP4 EGF  
ZO-1 : zona occludens 1

# Liste des figures

Figure 1 : Représentation du tube digestif primitif au premier mois de développement embryonnaire.....	20
Figure 2 : Le développement pulmonaire humain (d'après des documents de l'Université Paris V). .....	21
Figure 3 : Représentation de l'appareil respiratoire. ....	24
Figure 4 : Représentation des voies aériennes inférieures avec le nombre de divisions (Z). ....	26
Figure 5 : La double circulation pulmonaire (en bleu les veines, en rouge les artères) (adapté de (Aird, 2007)). ....	27
Figure 6 : Les propriétés des cellules souches pluripotentes. ....	32
Figure 7 : Images de CSP sur matrice (A) feeder et (B) Geltrex®, toutes deux à passage précoce P0 et P1 respectivement. ....	33
Figure 8 : Images de colonies CSP à passage (A) précoce – P0 – et (B) tardif – P58. ....	33
Figure 9 : Colonie d'ES après réaction à la phosphatase alcaline. ....	34
Figure 10 : Le noyau du circuit de régulation de la pluripotence (OCT4, NANOG et SOX2) (A) Boucle de rétrocontrôle positif du circuit et (B) Fonctionnement du circuit dans le maintien de la pluripotence et la répression de la différenciation (d'après (Jaenisch and Young, 2008) .....	35
Figure 11 : Schéma représentant les voies de signalisation extrinsèques pour le maintien de la pluripotence dans les CSP humaines (d'après 20535200). ....	37
Figure 12 : Image d'un corps embryoïde après 2 jours de différenciation, à partir des iPS iPCD02-30. ....	39
Figure 13 : Représentation de la première semaine du développement humain. ....	40
Figure 14 : Protocole de dérivation d'une lignée ES (adapté de (Odorico et al., 2001)). ....	43
Figure 15 : Image de colonies iPS (iPCD-02-30). ....	44
Figure 16 : Les phases de la reprogrammation (adapté de (David and Polo, 2014; Papp and Plath, 2013)). ..	51
Figure 17 : Evaluation des techniques de reprogrammation en fonction de leur efficacité et de leur sécurité (adapté de (Gonzalez et al., 2011)). ....	59
Figure 18 : Dérivation de fibroblastes, lignée FMO2. ....	60
Figure 19 : Immunofluorescence permettant la mise en évidence de l'expression de vimentine par les fibroblastes dérivés. ....	60
Figure 20 : Evaluation de la reprogrammation des cellules en fonction de leur accessibilité et de leur facilité à se reprogrammer (adapté de (Gonzalez et al., 2011)). ....	61
Figure 21 : Schéma résumant les différents choix possibles à chaque étape dans le but d'obtenir des iPS (adapté de (Maherali and Hochedlinger, 2008)) .....	62
Figure 22 : Schématisation des grandes étapes de la différenciation dirigée de l'arbre respiratoire (d'après (Kadzik and Morrisey, 2012)). ....	85
Figure 23 : Protocoles de différenciation en épithélium de l'appareil respiratoire (d'après (Gomperts, 2014)). Dans ce schéma, l'étape de ventralisation de l'AFE n'est pas individualisée .....	86

Figure 24 : Les deux grands types d'applications des iPS pour le développement de médicaments. ....	88
Figure 25 : Régénération connue d'un poumon adulte (A) dans les voies aériennes inférieures et (B) dans les voies respiratoires (d'après (Beers and Morrisey, 2011)).....	92
Figure 26 : Principe de la greffe de cellules progénitrices chez la souris (d'après (Snoeck, 2015)) .....	93
Figure 27 : illustration de la production d'organes dans un puits (d'après (Willyard, 2015)). .....	96
Figure 28 : Comparaison du processus conventionnel de découverte d'un médicament (A), avec le processus utilisant les cellules iPS de patient (B) (d'après (Grskovic et al., 2011)). .....	98
Figure 29 : Image de la dérivation de la lignée fibroblastique FMO3 9 jours après la biopsie. ....	107
Figure 30 : Immunofluorescence mettant en évidence l'expression par les fibroblastes de la lignée FMO2 de vimentine au cinquième passage. ....	108
Figure 31 : (A) culture de fibroblastes FMO2 (B) culture de fibroblastes FMO2 48h après la reprogrammation. ....	109
Figure 32 : (A) Premier clone apparu après reprogrammation : iPCD02-01 (P0), (B) iPCD02-30 adapté sur Geltrex dès son premier passage. ....	109
Figure 33 : Image d'une colonie iPS (iPCD02-26 (P1)) mal reprogrammée en aspect grappe de raisin. ....	110
Figure 34 : Caryotype de la lignée iPCD02-30 .....	110
Figure 35 : Immunohistochimie de tératome issu d'une lignée iPS (iPCD02-30) montrant du tissu d'origine (A) endodermique, (B) ectodermique et (C) mésodermique.....	111
Figure 36 : Immunofluorescence montrant l'expression des marqueurs de la pluripotence OCT4 et SOX2 par iPCD02-30.....	111
Figure 37 : Schéma de l'expression des facteurs de transcription le long du tube digestif d'un embryon de souris à E9.5. Abréviations: AFE, anterior foregut endoderm; DP, dorsal pancreatic bud; PFE, posterior foregut endoderm; VP, ventral pancreatic bud (d'après (Davenport et al., 2016b)).....	113
Figure 38 : Composition des milieux de différenciation du protocole de (Wong et al., 2015) (A) pour la différenciation des CSP de l'endoderme définitif jusqu'aux progéniteurs bronchiques (B) pour l'expansion des progéniteurs bronchiques (C) pour la polarisation et la maturation de l'épithélium bronchique .....	116
Figure 39 : Composition des milieux de différenciation du protocole de (Mou et al., 2012). ....	116
Figure 40 : Composition des milieux de différenciation du protocole de (Longmire et al., 2012).....	117
Figure 41 : Composition des milieux de différenciation du protocole de (Green et al., 2011).....	117
Figure 42 : Composition des milieux de différenciation du protocole de (Firth et al., 2014). ....	117
Figure 43 : Immunofluorescence mettant en évidence l'expression de FOXA2 dans des cellules (HS306) différenciées en DE par le protocole de S. Gerbal-Chaloin.....	119
Figure 44 : Mise en évidence de l'expression de FOXA2, par FACS, dans des cellules (HS306) différenciées en DE par le protocole de S. Gerbal-Chaloin.....	119

Figure 45 : Immunofluorescence mettant en évidence l'expression de FOXA2 dans les cellules (BJ-CTL-C1) différenciées en DE avec STEMdiff™ Definitive Endoderm Kit.....	120
Figure 46 : Expression relative des marqueurs de (A) pluripotence, (B) DE après différenciation de CSP (lignée HS306 et BJCTL1) en DE avec STEMdiff™ Definitive Endoderm Kit. ....	120
Figure 47 : Microscopie optique de cellules (iPCD02-30) en cours de différenciation, après 3 jours de traitement par le STEMdiff Definitive Endoderm kit.....	121
Figure 48 : Immunofluorescence mettant en évidence l'expression de FOXA2 dans les cellules (iPCD02-30) différenciées en DE par le protocole 25654758. ....	121
Figure 49 : Microscopie optique de cellules (iPCD02-30) différenciées en AFE (après 8 jours de différenciation selon le protocole (Wong et al., 2015)).....	123
Figure 50 : Microscopie optique des cellules (iPCD02-30) au stade PL (après (A) 24 et (B) 29 jours de différenciation selon le protocole (Wong et al., 2015)).....	124
Figure 51 : Microscopie optique des cellules (iPCD02-30) au stade PL (après 35 jours de différenciation selon le protocole (Wong et al., 2015)).....	125
Figure 52 : Expression relative des marqueurs (A) de pluripotence, (B) d'un épithélium bronchique polarisé et mature après différenciation de CSP (lignée iPCD02-30) selon le protocole (Wong et al., 2015). ...	125
Figure 53 : (A) Représentation d'une densité cellulaire (A) trop importante entraînant un gradient de concentration de cytokines (B) trop faible avec absence de communication paracrine. ....	126
Figure 54 : Microscopie cellulaire des cellules (iPCD02-30) après 9 jours de différenciation selon le protocole (Wong et al., 2015) et après plating au stade AFE à (A) 25 000 , (B) 35 000, (C) 50 000, (D) 100 000, (E) 150 000, (F) 175 000, (G) 200 000 cellules.....	127
Figure 55 : Microscopie optique de cellules (iPCD02-30) différenciées très hétérogènes à l'étape AFE après 9 jours de différenciation selon le protocole (Wong et al., 2015). ....	131



# Liste des tables

Tableau 1 : Représentation des différents marqueurs membranaires selon l'espèce et le temps (d'après (Henderson et al., 2002). .....	38
Tableau 2 : Vingt-quatre gènes candidats à la pluripotente, réduite à dix candidats puis quatre (adapté de (Takahashi and Yamanaka, 2006)).....	48
Tableau 3 : Liste non exhaustive des espèces dont les cellules ont pu être reprogrammées par la technologie des iPS.....	53
Tableau 4 : Liste des ARN et protéines capables de remplacer les facteurs de reprogrammation (d'après (Ho et al., 2011)) .....	55
Tableau 5 : Liste des facteurs et molécules chimiques capables d'améliorer la reprogrammation (adapté de (Feng et al., 2009b; Maherali and Hochedlinger, 2008)) .....	57
Tableau 6 : Récapitulatif de différents grands types de différenciation .....	72
Tableau 7 : Les protocoles de formation d'EB (adapté de (Kurosawa, 2007)) .....	74
Tableau 8 : Liste non exhaustive de protocoles de différenciation dirigée (adapté de (De Vos et al., 2016))	82
Tableau 9 : Liste des amorces utilisées que nous avons utilisées pour les PCR et qui sont validées par A. Wong (Wong et al., 2015) .....	115
Tableau 10 : Liste des anticorps utilisés pour les différentes immunofluorescences .....	115

# **INTRODUCTION**

Dans la Grèce antique, les Dieux vivaient sur le Mont Olympe. Hébé, leur échanton, leur versait l'ambrosie et le nectar. Ces mets divins leur permettaient de garder la jeunesse éternelle et l'immortalité. Hébé, plus connue sous son nom romain de Juventas, personnifie ces qualités. Prométhée, quant à lui, fut puni par Zeus pour avoir donné le feu sacré d'Olympe aux humains. Il fut condamné à se faire dévorer le foie par un aigle, un supplice sans fin puisque son foie se reconstituait toutes les nuits. Prométhée incarne le mythe de la régénérescence des organes. Les sciences médicales ne cessant de progresser, la réalité rejoint progressivement la mythologie : rajeunir un être humain, le régénérer.

Le poumon est un organe indispensable à la vie qui nous permet de respirer. L'appareil respiratoire comprend les voies aériennes supérieures - les voies nasales, les sinus, le pharynx et le larynx -, les voies aériennes inférieures - la trachée et les bronches -, et enfin les poumons.

Les maladies qui touchent l'appareil respiratoires peuvent se diviser en trois grands groupes : (i) les pathologies aiguës, d'origine bactérienne ou non, telles que la pneumonie et la bronchite ; (ii) les pathologies chroniques, tels que l'asthme et la broncho-pneumopathie chronique obstructive ; (iii) les pathologies génétiques dont la plus connue est la mucoviscidose, mais qui comportent également d'autres pathologies telles que la dyskinésie ciliaire primitive. Afin de pouvoir améliorer la prise en charge des patients et les traitements existants, il est essentiel de connaître le fonctionnement de l'appareil respiratoire à la fois au cours de son développement et à l'état adulte, et à la fois dans un contexte physiologie et pathologique.

La technologie des cellules souches pluripotentes, et notamment les cellules souches pluripotentes induites (iPS) récemment découvertes, offre une ouverture remarquable pour améliorer notre connaissance de l'appareil respiratoire normal et pathologique et développer de nouveaux traitements. En effet, les cellules souches pluripotentes ont la capacité de se différencier en tout type de tissu, y compris en tissu pulmonaire. La possibilité de créer *in vitro* de l'épithélium bronchique ou du tissu alvéolaire, sain ou malade, permet d'entrevoir de nouvelles applications médicales et pharmaceutiques. La

modélisation du tissu respiratoire permettra d'améliorer les traitements pharmaceutiques existants en les testant directement sur le tissu cible et permettra de la même manière le criblage de nouveaux candidat-médicaments. D'autre part, les cellules souches pluripotentes pourront constituer une source de cellules voire de tissus et d'organes pour une médecine régénératrice du poumon. Mon travail de doctorat s'est focalisé sur la mise au point de la modélisation des tissus respiratoires par l'utilisation de cellules souches pluripotentes.

# **PARTIE 1 : Le Poumon**

Le poumon est un organe vital. En effet, c'est l'organe de la respiration : il permet les échanges gazeux avec l'environnement. Il fait partie du système respiratoire qui comprend également les voies respiratoires.

## 1. Le Développement du système respiratoire

Après fécondation, le zygote devient rapidement un embryon didermique puis tridermique avec la formation des trois feuillettes au cours de la gastrulation pendant la troisième semaine du développement embryonnaire. Le soulèvement des feuillettes donne lieu dès le début de la quatrième semaine de développement au plan corporel de l'embryon. C'est le plissement crânio-caudal et le repliement des côtés gauche et droit du disque qui donnent à l'embryon des vertébrés l'aspect de tube, tractus gastro-intestinal d'origine endodermique, dans un tube, l'embryon. Le tube digestif primitif se divise en trois parties : le foregut, le midgut et l'hindgut (Figure 1).

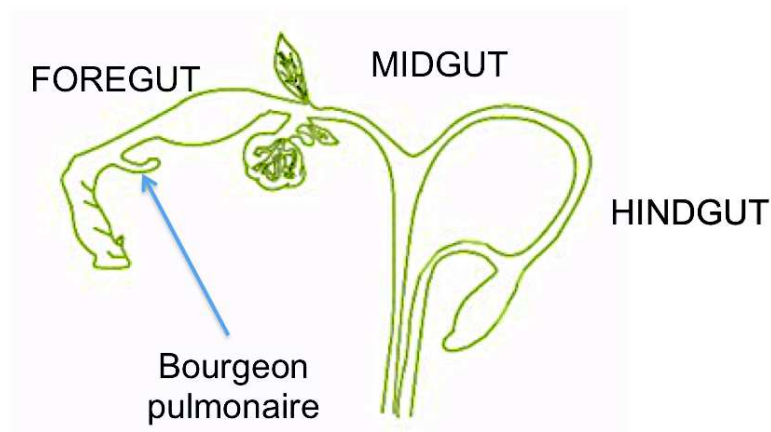


Figure 3 : Représentation du tube digestif primitif au premier mois de développement embryonnaire.

Le foregut correspond à l'intestin antérieur et donnera l'œsophage, l'estomac, le duodénum et plusieurs diverticules qui bourgeonnent au niveau du mésogastre et donneront le foie, la vésicule biliaire et le conduit cystique ainsi que pancréas. Le midgut correspond à l'intestin moyen, soit le jéjunum, l'iléon, le caecum et le colon ascendant. Le hindgut correspond à l'intestin postérieur et donnera le colon transverse, le colon descendant, le colon sigmoïde et le rectum.

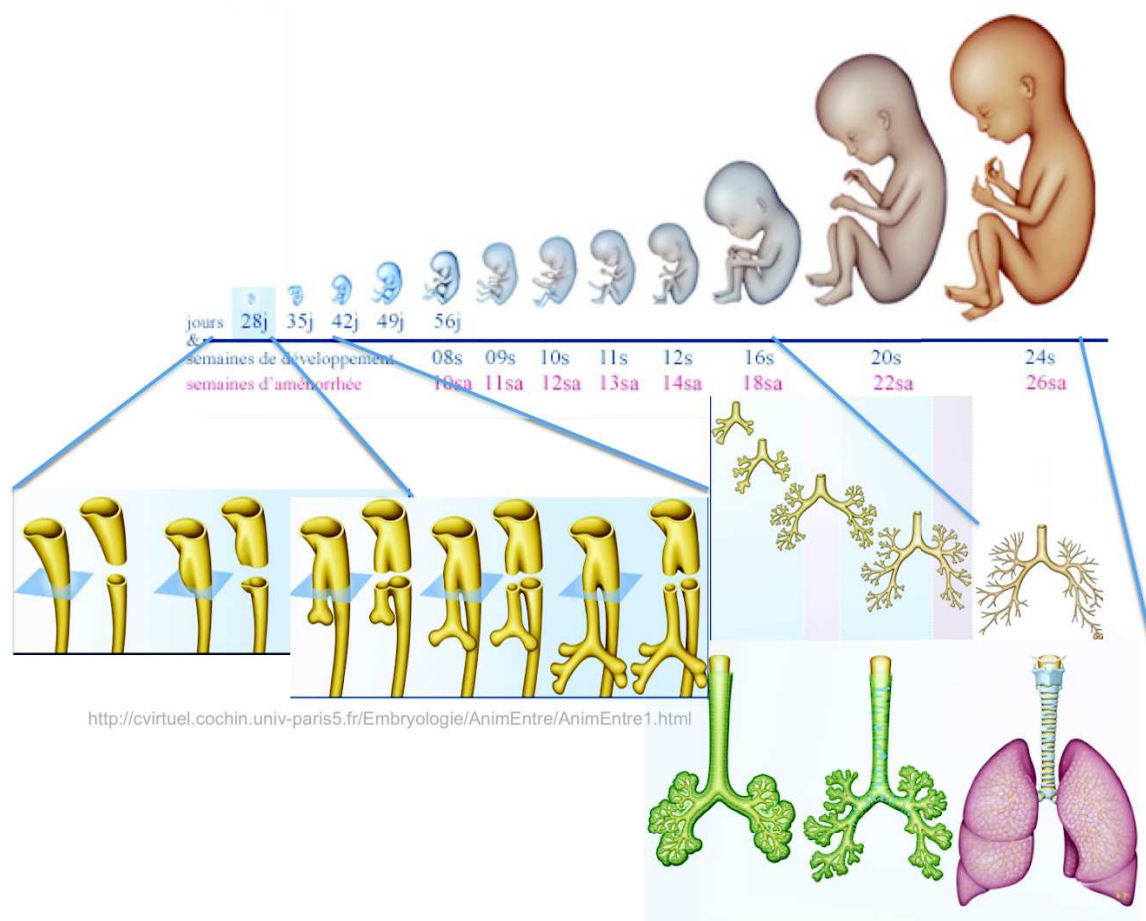


Figure 4 : Le développement pulmonaire humain (d'après des documents de l'Université Paris V).

Le développement pulmonaire se déroule tout au long de la grossesse et se termine après la naissance (Figure 2). Il se segmente en plusieurs étapes : embryonnaire, pseudo-glandulaire, canaliculaire, sacculaire et alvéolaire.

L'étape embryonnaire débute à partir du vingt-deuxième jour, par une évagination ventrale de l'intestin antérieur qui donne le diverticule respiratoire. Ce bourgeon pulmonaire est l'ébauche de l'appareil respiratoire, plus précisément de l'épithélium des voies aériennes et des alvéoles pulmonaires. Les autres constituants, comme le parenchyme pulmonaire et les anneaux bronchiques dérivent du mésenchyme et des vaisseaux environnants. Le diverticule progresse sous forme de tube droit, l'ébauche de trachée, parallèle au foregut et reste connecté à ce dernier par l'ébauche du larynx. Vers le trentième jour, la partie caudale du diverticule respiratoire se divise de façon dichotomique en deux bourgeons bronchiques primaires ou bronches souches. Le bourgeon bronchique gauche, presque horizontal, est plus petit que le bourgeon bronchique droit. La deuxième division a lieu pendant la cinquième semaine et est asymétrique. Le bourgeon gauche donne naissance à deux bronches secondaires ou lombaires alors que le bourgeon droit en donnera trois.

L'étape pseudo-glandulaire s'étend de la cinquième à la dix-septième semaine pour former les voies aériennes. Il s'agit d'une succession de division dichotomique donnant naissance aux bronches lombaires puis aux bronchioles terminales dont le diamètre diminue au fur et à mesure des divisions. Les voies de conduction de l'air sont totalement formées et bordées par un épithélium à cellules prismatiques précurseurs des cellules ciliées et des cellules sécrétrices.

La période canaliculaire s'étend de la dix-septième à la vingt-cinquième semaine et correspond à l'apparition des canalicules ou tubules, qui sont la partie respiratoire de l'arbre bronchique. L'unité respiratoire, ou acinus, dérive de la bronchiole terminale. Chaque acinus comprend trois à quatre divisions de bronchioles respiratoires prolongées par le canal alvéolaire. Cette étape du développement marque la modification du mésenchyme pulmonaire avec un envahissement d'un réseau capillaire qui sera le siège



des échanges gazeux. L'épithélium s'aplatit et la lumière des tubules s'élargit. Les pneumocytes de type II cubiques se différencient en pneumocytes de type I pavimenteux. A la fin de cette période, une partie du liquide amniotique est produite par l'épithélium pulmonaire. On peut ainsi mesurer la maturité des poumons du fœtus en mesurant dans le liquide amniotique la production de surfactant par les pneumocytes de type II.

La période sacculaire s'étend tout au long du troisième trimestre de la vingt-cinquième à la trente-sixième semaine. Il se forme les sacs terminaux, ou alvéoles primitives, marqués par la formation d'espaces aériens dans la partie respiratoire constitués d'alvéoles juxtaposés. Les parois sont tapissées des pneumocytes de types I et II. Les septa entre les sacs alvéolaires sont encore épais à ce stade. Les échanges gazeux sont possibles mais limités car les alvéoles sont encore peu nombreuses et immatures.

La dernière période, alvéolaire, commence au cours des dernières semaines de grossesse et se poursuit au delà de la naissance jusqu'à deux ans. Il se forme de nouveaux sacs alvéolaires définitifs et les premières alvéoles. A la naissance, on dénombre environ vingt à soixante-dix millions de sacs terminaux, pour un total de trois à quatre cent millions d'alvéoles. Cette dernière période est donc marquée par un processus d'alvéolisation mais aussi de septation. En effet, les septa s'amincissent et se multiplient afin de subdiviser davantage les alvéoles.

Il aura fallu près de huit mois in utero pour obtenir cet organe complexe permettant à l'organisme les échanges gazeux et la maturation se poursuit jusqu'à deux ans après la naissance.

## 2. Le poumon adulte

L'appareil respiratoire est composé des voies nasales, des sinus, du pharynx, du larynx, de la trachée, des bronches et de deux poumons (Figure 3). Ces derniers sont situés dans le thorax.

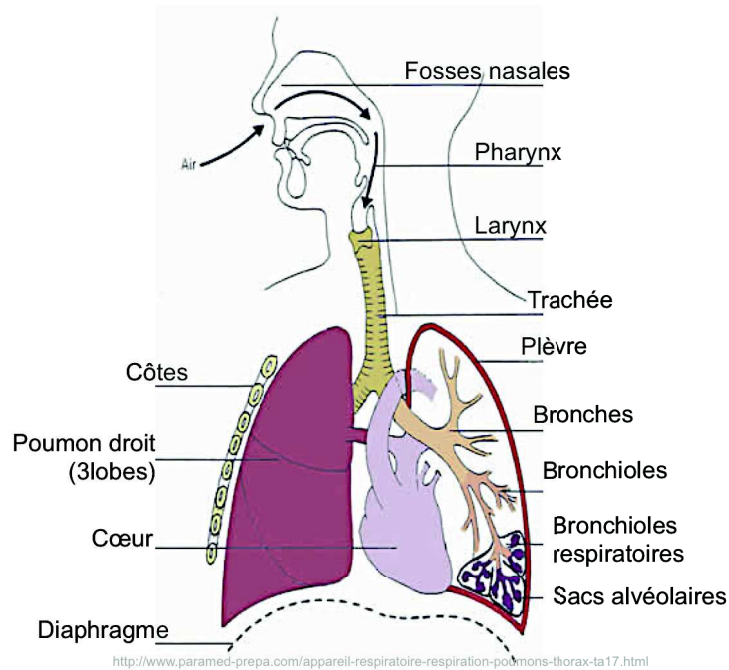


Figure 5 : Représentation de l'appareil respiratoire.

### 2.1 Anatomie – Physiologie

Les poumons sont posés sur le diaphragme, protégés par la cage thoracique et séparés par le médiastin. Les deux poumons ne sont pas symétriques. Le poumon droit a trois lobes alors que le gauche n'en a que deux afin de permettre un emplacement pour un autre organe vital, le cœur. Les lobes sont eux-mêmes divisés en plusieurs segments. La plèvre est une membrane à deux couches, viscérale et pariétale, et entoure les poumons.

Elle a un rôle protecteur mais aussi mécanique puisqu'elle facilite les mouvements des poumons pendant la respiration.

L'arbre bronchique se divise en deux parties : les voies conductrices de l'air et les voies respiratoires (Figure 4). Les voies aériennes permettent à l'inspiration de conduire l'air riche en oxygène jusqu'à la partie respiratoire et d'évacuer en sens inverse l'air riche en déchets carboniques à l'expiration. Il en existe deux types : les voies aériennes supérieures - les voies nasales, les sinus, le pharynx et le larynx, et les voies aériennes inférieures qui s'étendent de la trachée jusqu'aux bronchioles terminales. La trachée et les bronches sont tapissées par un épithélium cylindrique pseudostratifié (Jeffery, 1983). Il regroupe plusieurs types cellulaires mais toutes reposent sur la lame basale. Cet épithélium a également un rôle protecteur. En effet, le mucus qu'il produit associé aux battements de ses cils forment un système de transport muco-ciliaire. Ce qui permet d'éliminer les produits aéro-contaminants (les bactéries, les virus, les toxiques et autres poussières) qui ont été inhalés. Les voies respiratoires s'étendent des bronchioles respiratoires jusqu'aux sacs alvéolaires, comme leur nom l'indique c'est cette partie qui assurera les échanges gazeux entre l'air et le sang.

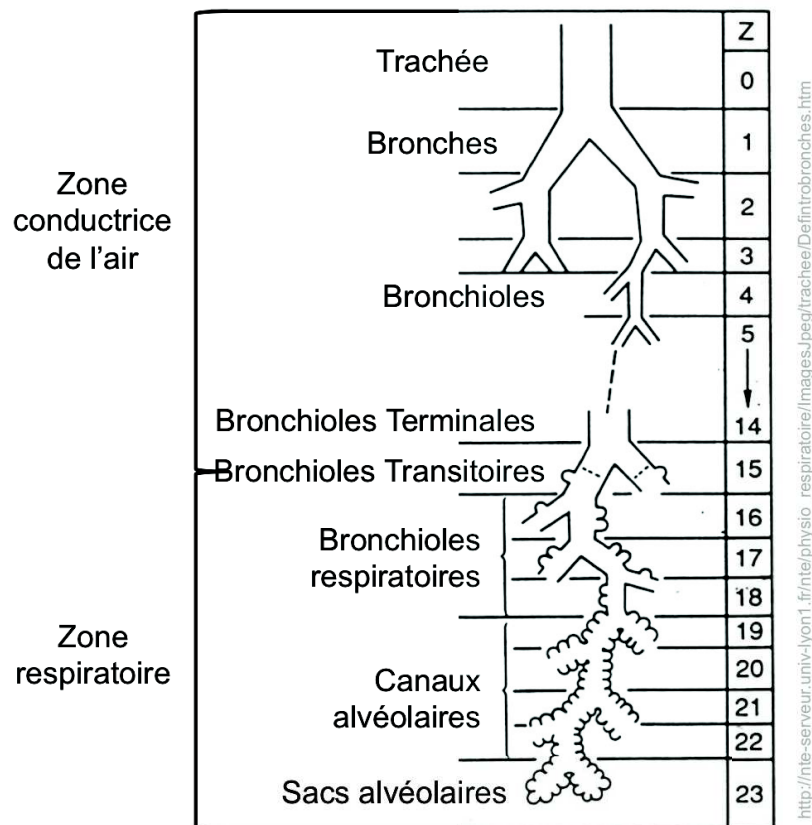


Figure 6 : Représentation des voies aériennes inférieures avec le nombre de divisions (Z).

Les poumons bénéficient d'une double vascularisation (Figure 5). Il existe d'une part une circulation bronchique assurée par les vaisseaux bronchiques, qui font partie de la circulation systémique, et qui approvisionnent le poumon en nutriments et oxygène, et les débarrassent de leurs déchets métaboliques. Il existe d'autre part une circulation pulmonaire placée entre le cœur droit et le cœur gauche : c'est la circulation fonctionnelle, assurant les échanges gazeux au niveau alvéolo-capillaire.

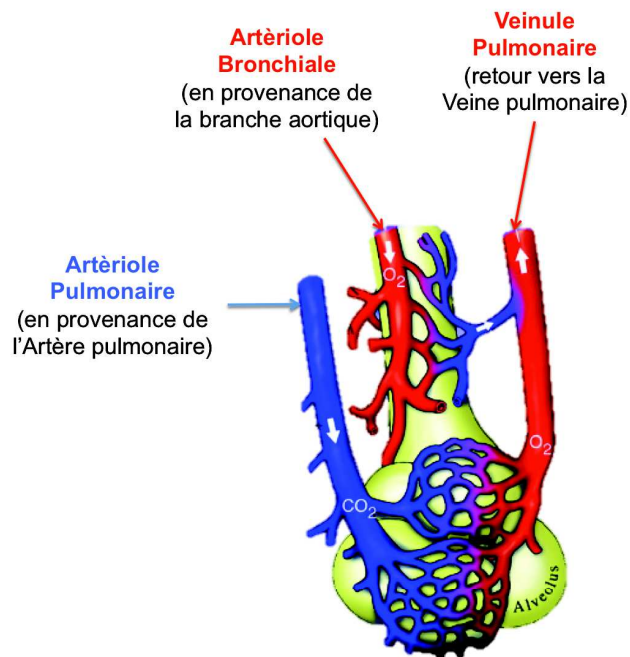


Figure 7 : La double circulation pulmonaire (en bleu les veines, en rouge les artères) (adapté de (Aird, 2007)).

## 2.2 Histologie

Les voies aériennes inférieures ont un rôle de conduction de l'air, de réchauffement de l'air inspiré et de protection. L'épithélium qui le compose lui permet de jouer ces fonctions et évolue au fur et à mesure des divisions bronchiques et bronchiolaires. De la trachée jusqu'aux bronchioles terminales, l'épithélium évolue d'un épithélium pseudostratifié vers un épithélium monocouche cuboïde. Il est composé principalement par des cellules basales, des cellules ciliées, des cellules caliciformes, des cellules neuroendocrines et des cellules Club (aussi appelées cellules en dôme ou cellules de Clara). Le pourcentage de répartition de ces cellules varie au fur et à mesure des divisions de l'arbre respiratoire.

Les cellules basales sont les seules cellules qui ne sont pas en contact avec la lumière bronchique, plus petites et ovoïdes. Elles donnent à l'épithélium sa cohésion. Elles sont

également considérées comme des cellules progénitrices permettant une régénération limitée de l'épithélium.

Les cellules ciliées sont prismatiques. Comme leur nom l'indique, elles ont leur pôle apical recouvert de cils (200 à 300 par cellules, de 5 à 7  $\mu\text{m}$ ). Les cils des cellules ont un mouvement rotatoire qui se synchronise. C'est le moteur de l'appareil muco-ciliaire. Toutes pathologies ciliaires rendant caduque ce moteur entraînent un encombrement bronchique par une non-évacuation du mucus, et des infections à répétitions par la non-élimination des agents pathogènes.

Les cellules caliciformes ou cellules à mucus sont prismatiques. Elles sécrètent des mucines. Ce sont des glycoprotéines à haut poids moléculaire qui sont le composant principal du mucus. Le mucus est une barrière physico-chimique qui protège l'épithélium des agressions en emprisonnant les agents pathogènes et autres poussières inhalées. Le mucus est évacué grâce à l'action ciliaire, il remonte jusqu'au carrefour aéro-digestif où il sera déglutit ou expectoré.

Les cellules neuroendocrines, de forme pyramidale, sont minoritaires. Elles appartiennent au système APUD ou neuroendocrinien diffus. Elles peuvent être isolées ou regroupées en îlots innervés, les corpuscules neuro-épithéliaux, préférentiellement retrouvés au niveau des bifurcations bronchiques. Elles ont à la fois un rôle de sécrétion (bombésine, calcitonine, sérotonine), de chémorécepteurs sensibles à l'hypoxie et de niche de cellules souches dans la régénération mais aussi la néoplasie.

Les cellules Club sont cuboïdales. Elles sont plus localisées vers les bronchioles terminales. Elles sécrètent CCSP (Club Cell Secretory Protein), ou SCGB1A1 (Secretoglobin 1A1), et des protéines du surfactant. Elles participent également au renouvellement de l'épithélium comme cellules progénitrices.

Les bronchioles respiratoires forment une zone de transition située entre la zone de conduction et la zone des échanges gazeux. Il y a trois générations de bronchioles respiratoires. Leur structure histologique est globalement identique à celle des bronchioles

terminales, mais avec la présence d'alvéoles, de plus en plus nombreuses dans la partie la plus distales.

La zone respiratoire à proprement parler est celle des échanges gazeux. L'épithélium alvéolaire, de type pavimenteux simple, comporte deux types cellulaires : les pneumocytes de type I et les pneumocytes de type II. Les pneumocytes de type I sont perméables aux gaz mais imperméables aux liquides, ils composent 97% de la paroi alvéolaire. Les pneumocytes de type II sécrètent une protéine tensio-active : le surfactant. Les échanges gazeux se font à travers la barrière alvéolo-capillaire qui englobe la cellule endothéliale, la lame basale et le pneumocyte de type I.

# **PARTIE 2 : Les Cellules**

## **Souches Pluripotentes**



Les cellules souches pluripotentes (CSP) sont des cellules souches définies par deux propriétés fondamentales : l'auto-renouvellement et la capacité à se différencier en tous les types cellulaires. Elles comprennent les cellules souches embryonnaires (ES) et les cellules souches pluripotentes induites (iPS).

# 1. La pluripotence

## 1.1 Généralités

La pluripotence est un état furtif dans le développement embryonnaire qui permet aux cellules de se différencier et donner des tissus issus des trois feuillets embryonnaires : endoderme, mésoderme et ectoderme. Avoir pu trouver les conditions pour préserver cette propriété *in vitro* à long terme, à partir d'embryons ou d'iPS, a été une véritable prouesse technologique et cela a créé une percée dans le domaine des biotechnologies.

## 1.2 Découverte

Historiquement, la propriété de pluripotence a été découverte dans les tératocarcinomes dans les années 50, puis étudiée dans les tératocarcinomes (Stevens and Little, 1954). Des cellules pluripotentes apparaissent dans le développement embryonnaire au stade du blastocyste, mais cet état est éphémère et disparaît dès le stade de gastrulation de l'embryon. Le challenge scientifique a été de capter cette propriété pour produire des lignées artificielles de cellules pluripotentes embryonnaires, c'est à dire des cellules souches embryonnaires. Martin Evans et Matthiew Kaufman ont été les premiers à dériver ces cellules ES chez la souris en 1981 (Evans and Kaufman, 1981), puis James Thomson chez l'humain en 1998 (Thomson et al., 1998). L'utilisation d'embryons pour

obtenir ces lignées soulève deux problèmes significatifs. Le premier, éthique, est la destruction d'un embryon ; le second, scientifique, réside dans l'impossibilité de choisir le patrimoine génétique de l'embryon de départ.

2016 marque le trente cinquième anniversaire des ES murines, le dix-huitième des ES humaines et le dixième des iPS.

### 1.3 Propriétés

La pluripotence et le caractère souche d'une lignée de cellules souches pluripotentes sont caractérisés par plusieurs techniques afin de qualifier ces propriétés (Figure 6) (Asprer and Lakshmiathy, 2015; Loring and Rao, 2006).

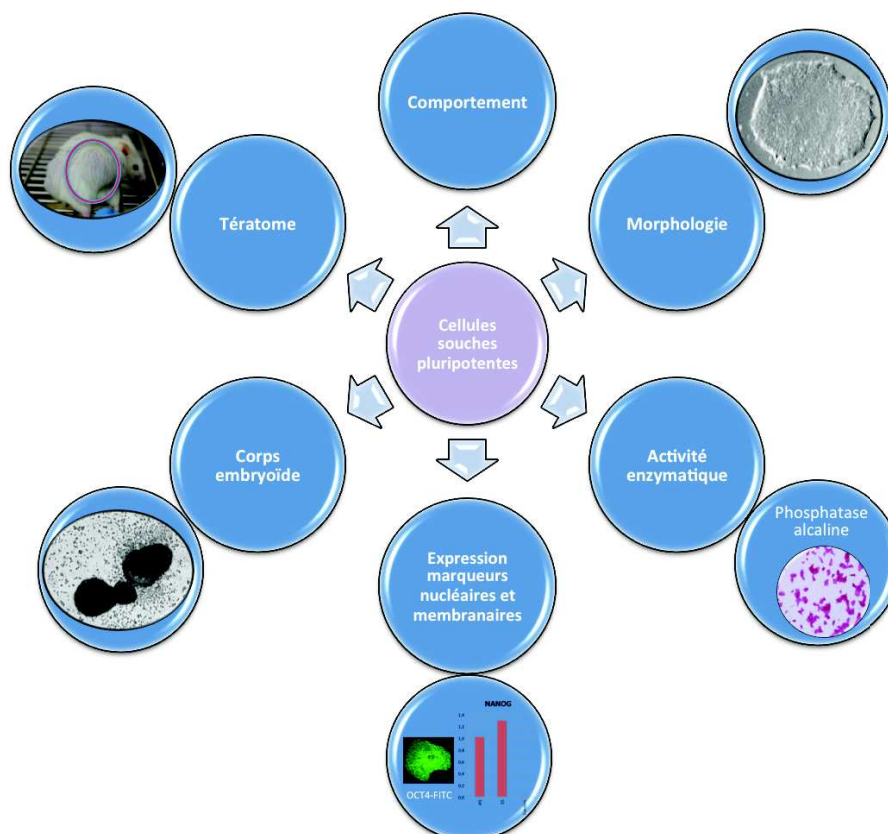


Figure 8 : Les propriétés des cellules souches pluripotentes.

### 1.3.1 La morphologie des cellules.

Elles possèdent un rapport noyau/cytoplasme élevé. Elles s'organisent en amas compact avec des bords bien définis appelés colonies (Figure 7).

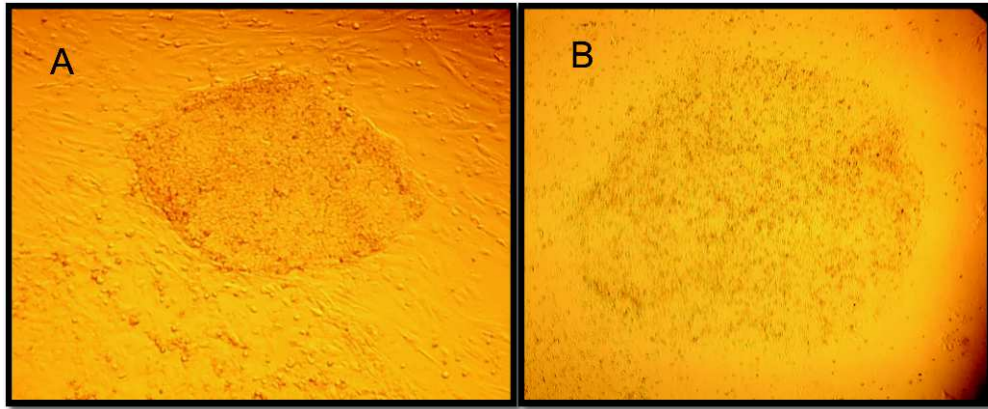


Figure 9 : Images de CSP sur matrice (A) feeder et (B) Geltrex, toutes deux à passage précoce P0 et P1 respectivement.

### 1.3.2 Le comportement de ces cellules.

Elles sont capables de garder leurs propriétés souche et pluripotente, après de nombreux passages et théoriquement indéfiniment (Figure 8).

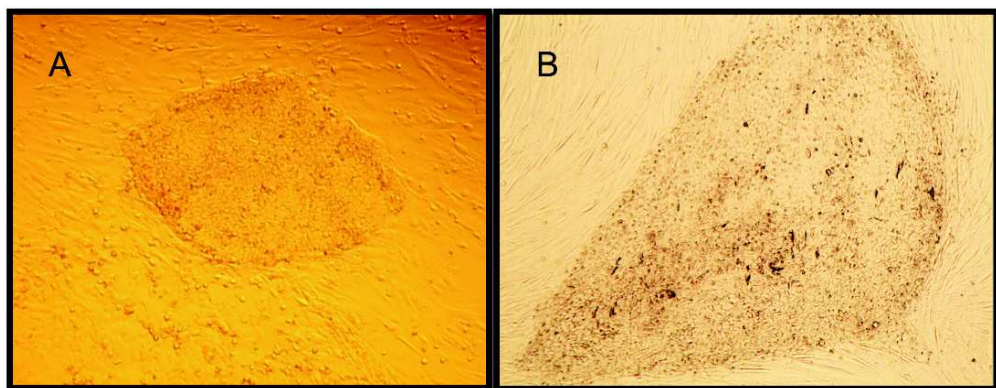


Figure 10 : Images de colonies CSP à passage (A) précoce – P0 – et (B) tardif – P58.

### 1.3.3 Les activités enzymatiques de ces cellules

L'activité enzymatique de la phosphatase alcaline (AP) est retrouvée chez l'homme sous forme de différentes isoenzymes, dont une est tissu non –spécifique exprimée à la fois dans le foie, le rein et les os, alors qu'une autre de ces isoenzymes est exprimée de façon élevée dans les cellules souches pluripotentes (Figure 9) (Singh et al., 2012). Il en va de même pour l'activité enzymatique de la télomerase ; elle est associée au caractère immortel des cellules souches embryonnaires, cancéreuses et germinales (Hiyama and Hiyama, 2007).

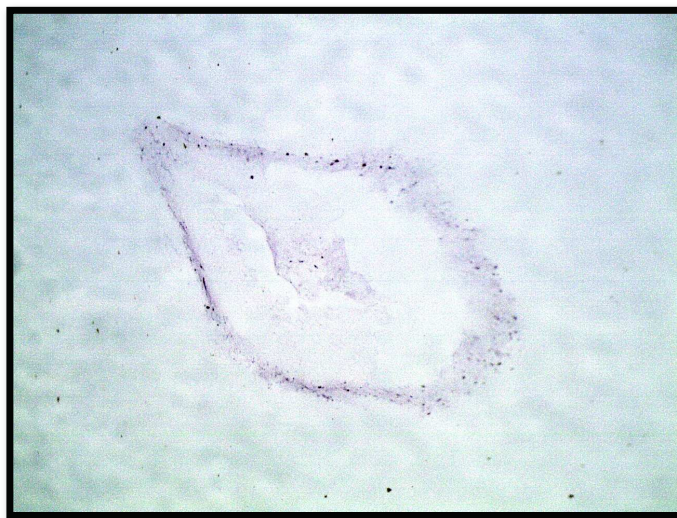
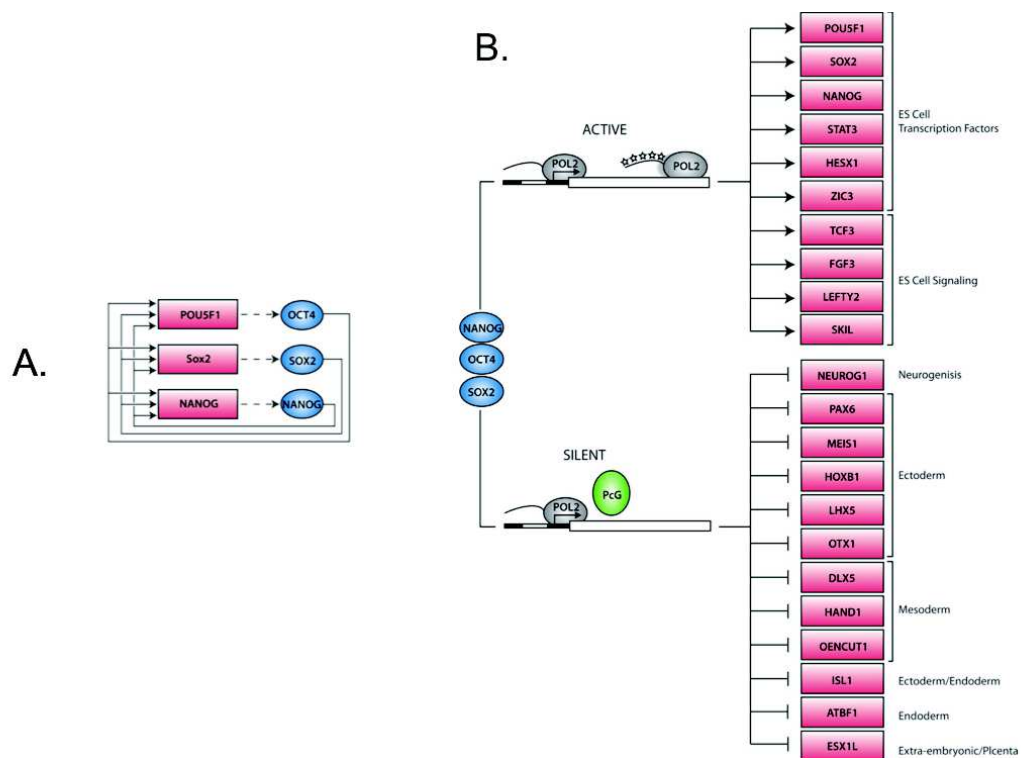


Figure 11 : Colonie d'ES après réaction à la phosphatase alcaline.

### 1.3.4 Le réseau de facteurs de transcription de la pluripotence

Les facteurs de transcription caractéristiques de la pluripotence sont : NANOG, SOX2, OCT4 (également connu comme POU5F1 ou OCT3/4), REX1, FLI1, DNMT3L/DPPA4 etc. (Asprer and Lakshmiathy, 2015; Richards et al., 2004; Sato et al., 2003). L'expression de ces gènes est commune aux mES et aux hES. Il est à remarquer que SOX2 est également exprimé dans le neur ectoderme est n'est donc pas spécifique de la pluripotence. Les trois gènes NANOG, SOX2 et OCT4 forme le noyau du circuit de régulation de la pluripotence (« core ES cell

regulatory circuitry ») (Jaenisch and Young, 2008). Ce circuit fonctionne comme une boucle de rétrocontrôle positif qui promeut le maintien de la pluripotence. En effet, ce réseau de facteurs de transcription induit l'expression d'un groupe de gènes nécessaire au maintien de la pluripotence, tout en bloquant l'expression de nombreux facteurs de transcription nécessaires à la différenciation dans chacun des trois feuilletts embryonnaires (Figure 10). Ce réseau de facteurs de transcription constitue les facteurs intrinsèques de la pluripotence, par opposition aux facteurs extrinsèques qui sont constitués par les signaux d'origine extracellulaire.

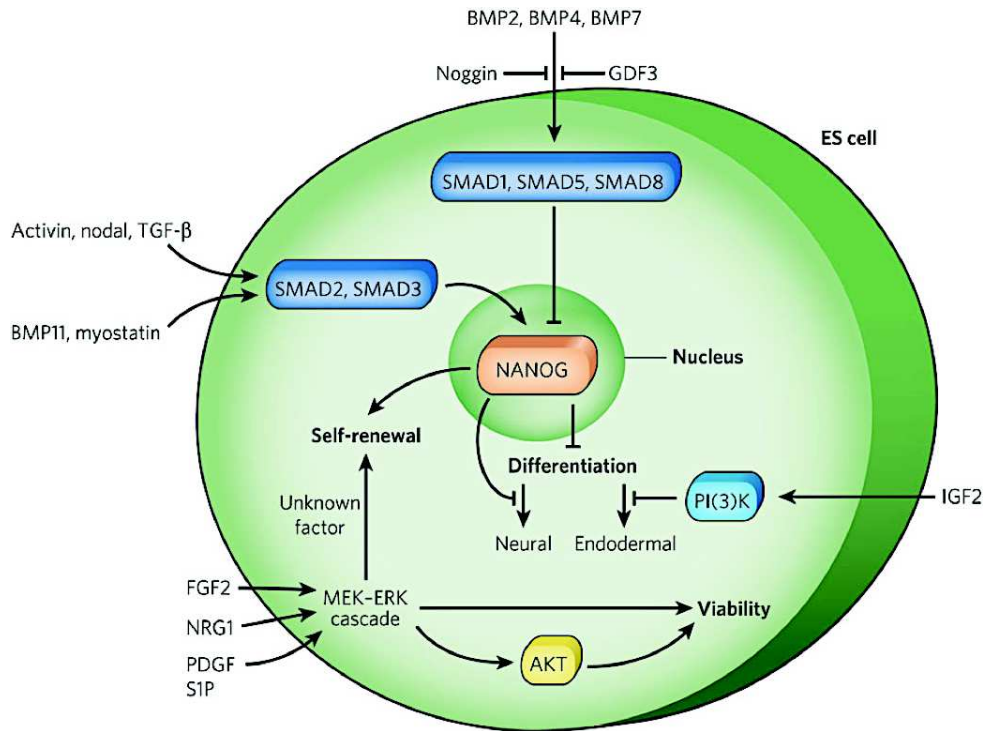


**Figure 12 : Le noyau du circuit de régulation de la pluripotence (OCT4, NANOG et SOX2) (A) Boucle de rétrocontrôle positif du circuit et (B) Fonctionnement du circuit dans le maintien de la pluripotence et la répression de la différenciation (d'après (Jaenisch and Young, 2008)).**

### 1.3.5 Les facteurs extrinsèques de la pluripotence

En plus des facteurs intrinsèques, des facteurs extrinsèques jouent un rôle essentiel dans le maintien de la pluripotence (Figure 11). En effet, l'expression des facteurs intrinsèques que nous venons de détailler ne suffisent pas à maintenir la pluripotence : si une CSP est placée dans un milieu de culture dépourvu des facteurs extrinsèques de la pluripotence, elle va se différencier.

Deux voies de signalisation jouent un rôle particulièrement important dans le maintien de la pluripotence (Pera and Tam, 2010). La voie SMAD2/SMAD3 est essentielle pour le maintien de l'expression du facteur de transcription NANOG. Activine et Nodal, membres de la superfamille des transforming growth factor –  $\beta$  (TGF- $\beta$ ), sont clés pour l'activation de cette cascade de signalisation (Vallier et al., 2009; Xu et al., 2008). A l'opposé, la voie de signalisation SMAD1/SMAD5 favorise la différenciation des CSP et doit être inhibée, ce qui peut être obtenu par l'ajout de Noggin au milieu de culture (Pera et al., 2004; Xu et al., 2005). L'autre voie de signalisation essentielle pour le maintien de la pluripotence des CSP humaines est la voie ERK/MAPK qui peut être activée par FGF2 (Xu et al., 2005). Ces deux voies peuvent se potentialiser mutuellement dans leur action dans le maintien de la pluripotence (Greber et al., 2007; Vallier et al., 2005; Xiao et al., 2006).



**Figure 13 : Schéma représentant les voies de signalisation extrinsèques pour le maintien de la pluripotence dans les CSP humaines (d'après (Pera and Tam, 2010)).**

Ce point sera repris plus loin, dans le chapitre consacré aux facteurs de croissance nécessaires pour le maintien des CSP.

### 1.3.6 L'expression de marqueurs membranaires de la pluripotence

Les cellules souches pluripotentes humaines expriment les glycosphingolipides suivants stage-specific embryonic antigen 3 et 4 (SSEA3, SSEA4), mais pas SSEA1 – contrairement aux mES (Wright and Andrews, 2009). Les protéoglycannes sulfates kératines TRA-1-60 et TRA-1-81 sont également exprimés par les cellules souches embryonnaires humaines (Henderson et al., 2002). Il est à noter que les marqueurs

membranaires exprimés diffèrent selon les espèces, mais également au cours de l'évolution de l'embryon (Tableau 1) (Henderson et al., 2002; Wright and Andrews, 2009).

		2-8 cellules	Morula	ICM
HUMAIN	SSEA1	-	+	-
	SSEA3	-	-	++
	SSEA4	-	-	++
	TRA-1-60	-	-	+++
	TRA-1-81	-	-	++
SOURIS	SSEA1	+	+	++
	SSEA3	+	+	-
	SSEA4	+	++	-
	TRA-1-60	*	-	-
	TRA-1-81	-	-	-

Tableau 1 : Représentation des différents marqueurs membranaires selon l'espèce et le temps (d'après (Henderson et al., 2002)).

### 1.3.7 La formation de corps embryoïdes, ou embryoid bodies (EB)

Lorsque les ES prolifèrent en trois dimensions *in vitro*, elles s'organisent en un amas sphérique appelé le corps embryoïde (Kurosawa, 2007). Les cellules se différencient alors en un mélange 3D de cellules issues des trois feuillets embryonnaires (Figure 12). Il s'agit d'une différenciation *in vitro* localement organisée (tissu cardiaque avec activité pace-maker, tube neural, épithélium digestif, etc.) mais globalement désordonnée et



n'aboutissant pas à un développement embryonnaire harmonieux. La production d'EB fait partie des méthodes de différenciation non dirigée (voir Partie 3 Chapitre 1).

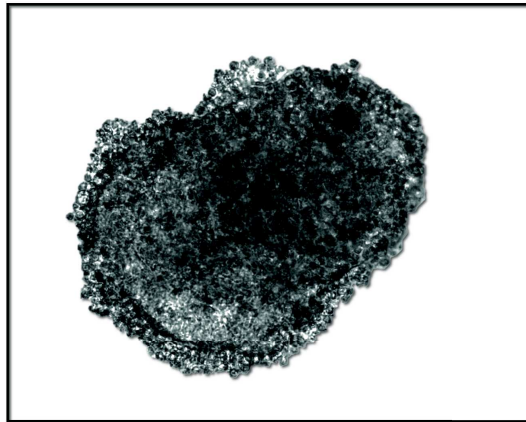


Figure 14 : Image d'un corps embryotaire après 2 jours de différenciation, à partir des iPS iPCD02-30.

### **1.3.8 La formation de tératomes dans un animal immunodéprimé**

Il s'agit du « gold standard » pour la caractérisation de la pluripotence. Nous décrivons ce procédé ultérieurement dans un chapitre dédié (voir Partie 3 Chapitre 1). La formation de tératomes fait également partie des méthodes de différenciation non dirigée.

## **2. Cellules souches pluripotentes embryonnaires**

Afin de comprendre d'où proviennent les cellules souches pluripotentes embryonnaires, il est indispensable de rappeler les différentes étapes de la première semaine du développement embryonnaire (Figure 13).

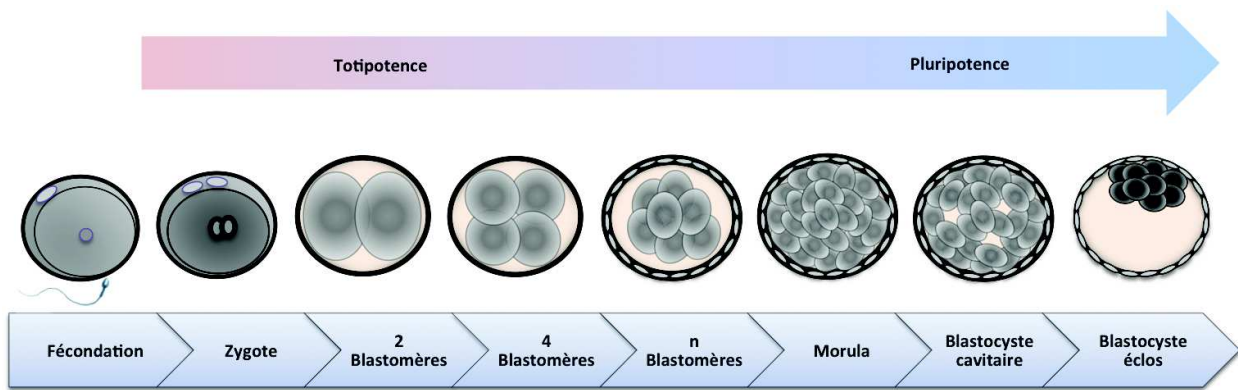


Figure 15 : Représentation de la première semaine du développement humain.

La fécondation, la fusion de deux gamètes, est une étape clé du développement embryonnaire. Le spermatozoïde, gamète mâle, et l'ovocyte de deuxième ordre, gamète femelle, se rencontrent dans l'ampoule tubaire des trompes de Fallope. Le spermatozoïde pénètre l'ovocyte II, puis les deux pronuclei fusionnent. C'est l'amphimixie (ou caryogamie) qui marque la fin de la fécondation. Il en résulte la formation d'un zygote à  $2n$  chromosomes. Il s'ensuit une succession de divisions cellulaires qui constituent la phase de la segmentation, laquelle a lieu en même temps que la migration du jeune embryon dans la cavité utérine. Les deux premières cellules ou blastomères apparaissent dès 24 heures après la fusion des deux pronuclei. Les divisions cellulaires s'opèrent avec un certain asynchronisme, donnant ainsi un œuf à 3, 4, 8, etc. et enfin 64 blastomères. Les premières divisions ont lieu à un moment précis de la segmentation : J2 pour le stade à 4 blastomères, J3 pour le stade à 8 blastomères et J4 pour le stade à 64 blastomères. Les divisions successives entraînent la diminution de taille des cellules mais pas de l'œuf qui est cloisonné dans la zone pellucide inextensible. Dès 16 blastomères, l'aspect de l'embryon ressemble à celui d'une petite mûre : ce stade est appelé *morula*. A 64 blastomères, l'embryon mesure toujours environ  $200\ \mu\text{m}$ . Les cellules en périphérie commencent à s'aplatissent et se resserrent alors que les cellules centrales forment un amas de cellules sphériques. Les cellules périphériques se divisent plus rapidement que les cellules centrales dans un rapport de 10:1. Elles délimitent la morula pour former le trophoblaste, une paroi

cellulaire épithéliale. Les cellules centrales, qui étaient bien individualisées au stade de morula, vont adhérer fortement les unes aux autres via l'induction de la molécule d'adhésion E-cadhérine, brouillant les limites intercellulaires sous le microscope : c'est le stade de compaction. Entre le quatrième et le cinquième jour, une cavité se forme, le blastocèle. Il s'agit de liquide extracellulaire accumulé dans les espaces intercellulaires de la morula grâce à un gradient de pression osmotique. Les cellules centrales sont repoussées à un pôle et forment la masse cellulaire interne ou bouton embryonnaire qui sera à l'origine de l'embryon à proprement parler et d'une partie des annexes embryonnaires. Le trophoblaste, ou masse cellulaire externe ou encore trophoctoderme, sera à l'origine des annexes embryonnaires. A la fin du cinquième jour, le blastocyste éclot au pôle anti-embryonnaire. Cette éclosion est également appelée hatching. Le blastocyste se libère de la zone pellucide afin de pouvoir s'implanter dans l'utérus maternel, ce qui clos la première semaine du développement embryonnaire humain.

Au cours de son développement, l'œuf fécondé acquiert et perd plusieurs propriétés de différenciation que sont la totipotence, la pluripotence, la multipotence et l'unipotence. La totipotence est la propriété des blastomères au cours des premières divisions embryonnaires à donner un être vivant, c'est à dire l'embryon et les annexes embryonnaires ; donc de générer les jumeaux monozygotes ou vrais jumeaux. Il s'agit d'un état éphémère qui disparaît rapidement après le stade 4-8 blastomères. La pluripotence est la propriété d'une cellule de se différencier dans n'importe quel type cellulaire d'un des trois feuillets embryonnaires : endoderme, mésoderme, ectoderme. Cette propriété apparaît dans les cellules de la masse cellulaire interne au stade de blastocyste, entre le cinquième et sixième jour de développement. La multipotence est la capacité à générer plusieurs types cellulaires constitutifs d'un tissu. C'est une propriété des cellules souches adultes. Enfin, l'unipotence est la propriété à générer un seul type de cellule qui caractérise certains types de cellules souches spécialisées telles que les cellules satellites du muscle.

La pluripotence est une particularité éphémère de l'embryon. La fixation artificielle de cette propriété *in vitro* a été une avancée remarquable qui a abouti aux cellules souches embryonnaires.

Les premiers à avoir réalisé la dérivation des cellules souches pluripotentes sont Evans et Kaufman, dans la souris en 1981 (Evans and Kaufman, 1981). Après avoir mis en culture la masse cellulaire interne de blastocystes de souris, ils ont réussi à isoler les cellules souches pluripotentes (mES). Ils ont ensuite caractérisé les différentes lignées obtenues par leur caryotype, leur capacité d'auto-renouvellement (au moins trente passages) et la pluripotence (formation de tératomes et de corps embryonnaires). C'est en 1998 que Thomson (Thomson et al., 1998) réalise pour la première fois la dérivation de cellules souches pluripotentes humaines, par mise en culture de la masse cellulaire interne d'embryons humains de cinq à sept jours (Figure 14). Tout comme la dérivation des cellules souches pluripotentes de primates (Thomson et al., 1995), la masse cellulaire interne, ou inner cell mass (ICM) est d'abord séparée du trophoctoderme, puis mise en culture sur des fibroblastes embryonnaires de souris (MEF) qui servent de cellules nourricières, ou feeders, aux cellules souches pluripotentes. A l'instar de l'équipe d'Evans et Kaufman, Thomson a ensuite caractérisé les cellules souches embryonnaires humaines (hES) obtenues par leur caryotype, leur capacité d'auto-renouvellement (trente-deux passages pour une des lignées obtenues) et leur capacité à se différencier dans les trois feuillets embryonnaires par la formation de tératomes. Ce protocole de dérivation robuste est encore utilisé. Une variante du protocole est le remplacement des MEF par des fibroblastes humains (hFF) comme cellules nourricières (Assou et al., 2009; Peura et al., 2007). Les hES obtenues sont ensuite amplifiées par passage mécanique ou enzymatique et peuvent être cryoconservées dans de l'azote liquide.

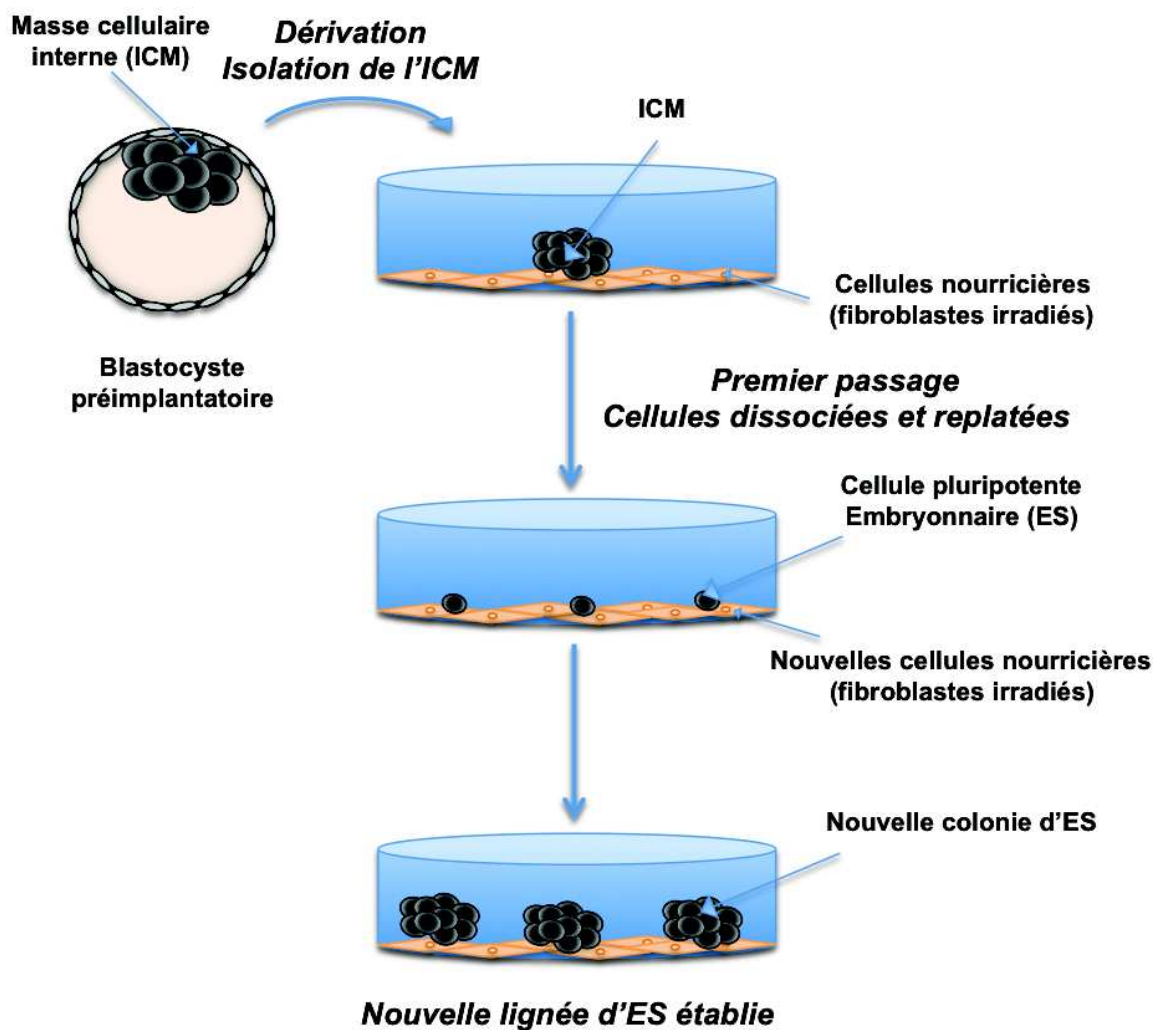


Figure 16 : Protocole de dérivation d'une lignée ES (adapté de (Odorico et al., 2001)).

En France, depuis la révision de la loi de la bioéthique de 2013<sup>1</sup>, la recherche sur les cellules souches embryonnaires humaines est autorisée mais encadrée par l'Agence de la Biomédecine (ABM). Les embryons humains utilisés pour obtenir les hES sont issus d'embryons surnuméraires cryoconservés obtenus par fécondation *in vitro* dans le cadre de procréation médicalement assistée. Ils ne font plus partie d'un projet parental ou sont porteurs d'une maladie génétique grave identifiée lors d'un dépistage préimplantatoire. Le couple concerné doit après information donner son consentement écrit. Ce consentement

<sup>1</sup> LOI n° 2013-715 du 6 août 2013 tendant à modifier la loi n° 2011-814 du 7 juillet 2011 relative à la bioéthique en autorisant sous certaines conditions la recherche sur l'embryon et les cellules souches embryonnaires

doit être confirmé à l'issue d'un délai de réflexion de trois mois. Tant que les recherches n'ont pas débuté le couple peut révoquer sans motif ce consentement.

L'utilisation des hES pose des problèmes éthiques, en raison de la nécessité de détruire un embryon humain pour l'obtention de ces cellules. La loi française n'autorise de recherche sur l'embryon humain que s'il peut y avoir des progrès médicaux majeurs et s'il n'est pas possible d'utiliser d'autres types cellulaires.

Récemment un nouveau type de cellules souches pluripotentes est apparu : les cellules souches pluripotentes induites ou iPS (Figure 15). Celles-ci permettent de s'affranchir de l'utilisation d'embryons pour obtenir des cellules avec les mêmes propriétés de pluripotence et d'auto-renouvellement que les hES.

### 3. induced Pluripotent Stem cells : iPS

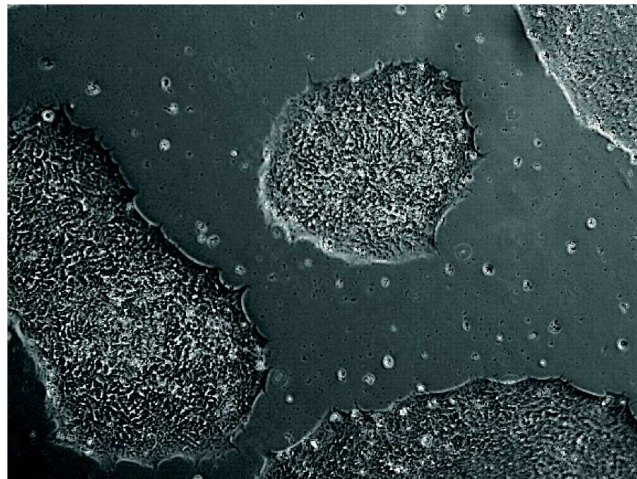


Figure 17 : Image de colonies iPS (iPCD-02-30).

### 3.1 Les précurseurs de la reprogrammation

En 1952, Robert Briggs et Thomas King (Briggs and King, 1952) furent les premiers à transférer un noyau de blastula dans un œuf énucléé et à observer un développement normal de *Rana pipiens*. Bien que ce soit la première reprogrammation, il faudra attendre 1962 et les travaux de Sir John Gurdon sur le transfert de noyau de cellules somatiques (Somatic cell nucleus transfert, SCNT) (Gurdon, 1962; Gurdon and Uehlinger, 1966) pour avoir une démonstration définitive de la possibilité de reprogrammer une cellule adulte, spécialisée, vers une cellule totipotente. Le noyau d'une cellule intestinale de grenouille est transféré dans un oocyte énucléé et permet la formation d'un têtard normal (Gurdon, 1962). Ces travaux ont été récompensés par le prix Nobel de médecine en 2012. Cette technique a ensuite été appliquée à d'autres espèces, notamment les mammifères (Baranska and Koprowski, 1970; Graham, 1969). La porte de la reprogrammation nucléaire s'était ouverte.

En 1997, l'équipe de Sir Ian Wilmut réussit à reprogrammer une cellule somatique de mammifère – une cellule de la mamelle d'une brebis - par la technique SCNT en utilisant un oocyte, et à obtenir un clone vivant issu de cette reprogrammation : Dolly la brebis. Ainsi, l'état de pluripotence éphémère avait été réobtenu à partir d'une cellule adulte de mammifère. Les travaux de Shinya Yamanaka (voir ci-dessous) ont permis de s'affranchir de l'utilisation d'un oocyte par la reprogrammation directe d'une cellule adulte avec un nombre limité de facteurs (Takahashi and Yamanaka, 2006).

### 3.2 Découverte des iPS

C'est en 2006 que le japonais S. Yamanaka et son post-doctorant Kazutoshi Takahashi ont mis au point une technique simple de reprogrammation cellulaire (Takahashi and Yamanaka, 2006). Ils ont démontré que l'expression transitoire de quatre gènes dans des cellules somatiques différenciées suffit à induire un état de pluripotence et donc de les reprogrammer en cellules pluripotentes induites ou iPS.

La découverte se déroule comme suit. Yamanaka et son équipe ont sélectionné vingt-quatre facteurs de transcription candidats à l'induction de la pluripotence. Ces candidats ont été choisis sur leur capacité à maintenir l'état de pluripotence chez les cellules ES, par exemple OCT4 (Nichols et al., 1998; Niwa et al., 2000), NANOG (Chambers et al., 2003) ou SOX2 (Avilion et al., 2003), c-MYC (Cartwright et al., 2005), KLF-4 (Li et al., 2005). L'expression de la totalité des vingt-quatre gènes candidats dans des fibroblastes murins adultes ou embryonnaires a permis l'émergence de plusieurs clones de cellules reprogrammées ou iPS. Cette expression des gènes est réalisée par des vecteurs rétroviraux. L'équipe japonaise a ensuite tenté de réduire la liste des vingt-quatre gènes à un nombre minimal (Tableau 2). Elle n'a retenu que dix gènes candidats pour ensuite terminer avec une combinaison de quatre gènes : OCT3/4, SOX2, c-MYC et KLF4, appelé le cocktail de Yamanaka.

<b>No.</b>	<b>Symbole</b>	<b>Accession number</b>	<b>Candidats</b>
<b>1.</b>	Ecat1	AB211060	
<b>2.</b>	Dppa5 (Esg1)	NM_025274	
<b>3.</b>	Fbxo15	NM_015798	Liste des 10
<b>4.</b>	Nanog	AB093574	Liste des 10
<b>5.</b>	ERas	NM_181548	Liste des 10



<b>6.</b>	Dnmt3l	NM_019448	
<b>7.</b>	Ecat8	AB211061	
<b>8.</b>	Gdf3	NM_008108	
<b>9.</b>	Sox15	NM_009235	
<b>10.</b>	Dppa4	NM_028610	
<b>11.</b>	Dppa2	NM_028615	Liste des 10
<b>12.</b>	Fthl17	NM_031261	
<b>13.</b>	Sall4	NM_175303	
<b>14.</b>	Oct3/4 (Pou5f1)	NM_013633	Liste des 4
<b>15.</b>	Sox2	NM_011443	Liste des 4
<b>16.</b>	Rex1 (Zfp42)	NM_009556	
<b>17.</b>	Utf1	NM_009482	
<b>18.</b>	Tcl1	NM_009337	Liste des 10
<b>19.</b>	Dppa3 (Stella)	NM_139218	
<b>20.</b>	Klf4	NM_010637	Liste des 4

<b>21.</b>	$\beta$ -catenin	NM_007614	Liste des 10
<b>22.</b>	c-Myc	NM_010849	Liste des 4
<b>23.</b>	Stat3	NM_213659	
<b>24.</b>	Grb2	NM_008163	

**Tableau 2 : Vingt-quatre gènes candidats à la pluripotence, réduite à dix candidats puis quatre (adapté de (Takahashi and Yamanaka, 2006)).**

Les clones qui émergent au cours de la reprogrammation sont identifiés puis qualifiés pour leur pluripotence - comme le sont des ES - grâce à leur morphologie, leur activité télomérase, l'expression des marqueurs de pluripotence, leur capacité à se différencier *in vitro* et *in vivo* via les tératomes.

L'année suivante, S. Yamanaka a appliqué avec succès son cocktail de reprogrammation à des fibroblastes humains (Takahashi et al., 2007). Depuis, de nombreuses espèces animales ont été reprogrammées.

Le prix Nobel de médecine 2012 (Jaenisch, 2012) a été remis conjointement à John Gurdon pour son travail de reprogrammation nucléaire avec le premier clonage chez le xénope (Gurdon et al., 1958) et à Shinya Yamanaka pour sa découverte des iPS (Takahashi and Yamanaka, 2006).

### 3.3 Rôle des facteurs de transcription dans la reprogrammation

Des quatre facteurs, c-MYC agit comme un activateur précoce de la reprogrammation, avant que les régulateurs pluripotents soient activés. En effet, c-MYC joue un rôle majeur dans la répression des gènes spécifiques de fibroblastes (Sridharan et al., 2009). Cependant, c-MYC est un protooncogène, sa dérégulation intervient dans de nombreuses tumeurs chez l'homme (Adhikary and Eilers, 2005). En effet, il contribue à l'activation de la chromatine, à la prolifération cellulaire et au passage de l'initiation à l'élongation de la transcription (Rahl et al., 2010). Son utilisation dans la reprogrammation soulève donc un problème pour la fabrication d'iPS de grade clinique, i.e. utilisable dans une stratégie thérapeutique chez l'homme. Dès 2010, l'équipe de Yamanaka a réussi à générer des iPS à partir de fibroblastes murins et humains sans l'utilisation de MYC (Nakagawa et al., 2008; 2010). De plus il apparaît que cette stratégie, bien que moins efficace, formerait moins de colonies mal reprogrammées.

OCT4 appartient à la famille des facteurs de transcription des porteurs du domaine POU (Ryan and Rosenfeld, 1997) et est capable de se lier à des séquences octamériques spécifiques. Dès la fin des années 90, ce facteur est connu pour avoir un rôle déterminant dans la formation de l'état de pluripotence chez les embryons. En effet, la perte de son expression chez les embryons préimplantatoires conduit à la formation de trophectoderme (Nichols et al., 1998) et à l'apoptose des cellules germinales primordiales (Kehler et al., 2004). De plus, ce facteur est nécessaire mais suffisant pour induire la pluripotence dans les cellules souches neurales (Kim et al., 2009b).

SOX2 est un des facteurs de transcription de la famille SOX, SRY-related HMG-box, car son domaine HMG (high mobility group) est semblable à celui de SRY (Sex determining region on Y) (Schepers et al., 2002). Il a généralement besoin de la coopération d'un autre facteur de transcription pour pouvoir ensuite se lier à l'ADN et ainsi permettre l'expression des gènes cibles de ce complexe. Lorsque SOX2 se lie à OCT4, ce sont les gènes de la pluripotence qui sont alors activés, tels que NANOG, OCT4, SOX2, FGF4, UTF1, FBX15, LEFTY1, NMYC1, REX1, TCF3, DPPA4, TRP53 et REST (Boyer et al., 2005; Chew et al., 2005;

Kondoh and Kamachi, 2010; Kuroda et al., 2005; Loh et al., 2006; Nakatake et al., 2006; Nishimoto et al., 1999; Okumura-Nakanishi et al., 2005; Rizzino, 2009; Rodda et al., 2005; Tokuzawa et al., 2003; Tomioka et al., 2002).

KLF4 appartient à la famille des facteurs Krüppel-like (Klfs) (Dang et al., 2000). Ce sont des protéines à doigt de zinc qui contiennent une séquence d'acides aminés similaire chez la drosophile. Cette protéine est surtout distribuée au niveau gastro-intestinal. KLF4 joue un rôle important dans l'activation de nombreux gènes impliqués dans la progression du cycle cellulaire par l'activation des points de contrôle et dans le maintien de l'intégrité de l'ADN (McConnell et al., 2007). D'autre part, KLF4 a une activité redondante avec KLF2 et KLF5, et forment à eux trois un corps régulateur clé pour l'auto-renouvellement des cellules ES et plus particulièrement les gènes OCT4, NANOG, SOX2 (Jiang et al., 2008).

Il est à noter que les facteurs OCT4, SOX2 et KLF4 sont indispensables à la plupart des protocoles de reprogrammation, tandis que c-MYC joue le rôle d'activateur de la reprogrammation et est dispensable (Nakagawa et al., 2008).

### **3.4 Le déroulement moléculaire de la reprogrammation**

La reprogrammation passe par trois grandes étapes : l'initiation, la maturation et la stabilisation (David and Polo, 2014) (Figure 16). L'initiation consiste en un ensemble d'événements morphologiques et métaboliques. Il s'agit d'événements associés à une transition mésenchymato-épithéliale (Samavarchi-Tehrani et al., 2010) ainsi qu'à la perte de l'expression des gènes de différenciation (par exemple la perte de la signature somatique tels que les marqueurs cellulaires fibroblastiques (Mikkelsen et al., 2008)), des changements métaboliques (réactivation de l'activité télomérase), l'augmentation de la prolifération, l'inhibition de l'apoptose et de la sénescence. L'étape suivante est la maturation. Elle consiste en l'acquisition de l'expression des marqueurs de la pluripotence (NANOG, ou OCT4 endogène), mais toujours sous la dépendance des transgènes. Enfin, la dernière étape est la stabilisation (Ho et al., 2011). Il s'agit de l'acquisition complète de la pluripotence, la réactivation du chromosome X (Stadtfield et al., 2008b), l'élongation des

télomères (Marion et al., 2009), la perte de la mémoire épigénétique (Kumar et al., 2013; Papp and Plath, 2013) et l'indépendance des transgènes pour l'auto-renouvellement des cellules pluripotentes.

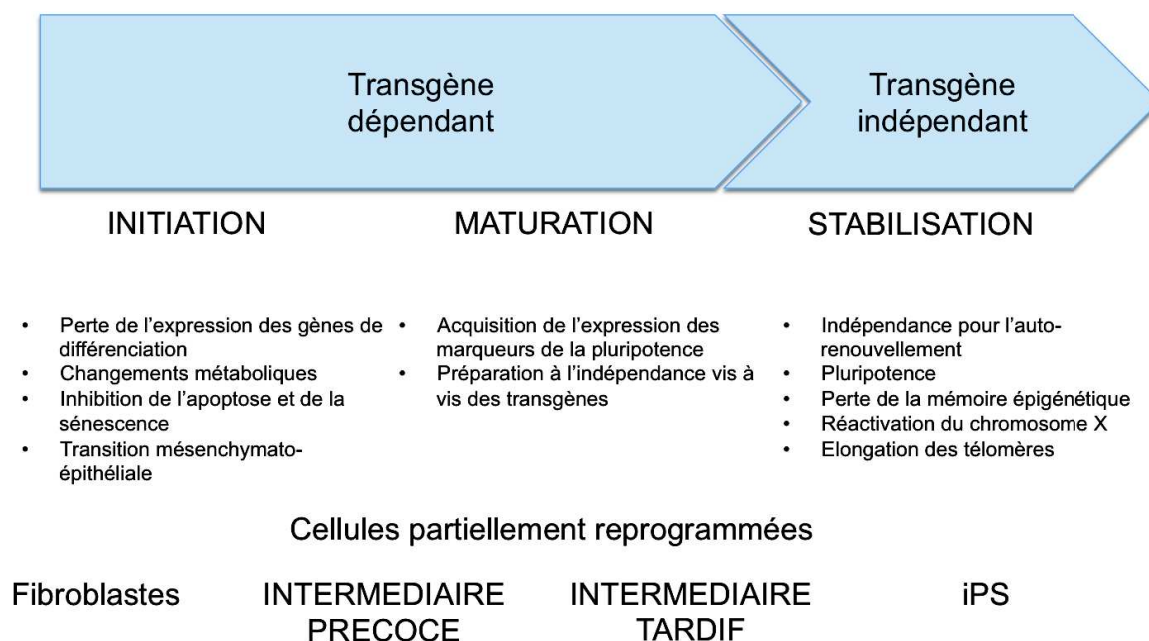


Figure 18 : Les phases de la reprogrammation (adapté de (David and Polo, 2014; Papp and Plath, 2013)).

### 3.5 Différentes espèces reprogrammées

Depuis les fibroblastes murins en 2006, différentes espèces ont été reprogrammées. Dès 2007, l'équipe de Yamanaka (Takahashi et al., 2007) a réussi à reprogrammer des fibroblastes humains à partir d'une simple biopsie de peau. Depuis, de nombreuses autres espèces ont été reprogrammées (Tableau 3). En effet, outre l'intérêt scientifique, la reprogrammation en iPS de différentes espèces pourrait permettre le développement et l'application de médecine régénérative, le développement de modèles animaux pour diverses maladies humaines (drépanocytose (Hanna et al., 2007) ; Anémie de Fanconi (Raya et al., 2009)) afin de pouvoir mieux les étudier et les traiter, pour le développement de la médecine vétérinaire (Cebrian-Serrano et al., 2013), et pour la biotechnologie des animaux,

notamment la production d'animaux génétiquement modifiés qui synthétisent des protéines thérapeutiques, qui ont un taux de croissance amélioré ou qui sont résistants à certaines maladies (Kumar et al., 2015).

<b>Espèce</b>	<b>Année</b>	<b>Intérêt scientifique</b>	<b>publication</b>
<b>Souris</b>	2006	Découverte des iPS	(Takahashi and Yamanaka, 2006)
<b>Humain</b>	2007	- médecine régénérative - étude du développement	(Takahashi et al., 2007)
<b>Cheval</b>	2011	Médecine vétérinaire	(Nagy et al., 2011)
<b>Lapin</b>	2010	Modèle expérimental pour médecine humaine	(Honda et al., 2010)
<b>Chien</b>	2011	Médecine vétérinaire	(Luo et al., 2011)
<b>Cochon</b>	2010	Modèle expérimental pour médecine humaine	(West et al., 2010)
<b>Rhésus Singe</b>	2008	Modèle expérimental pour médecine humaine	(Liu et al., 2008)
<b>Rat</b>	2009	Modèle expérimental pour médecine humaine	(Liao et al., 2009)

<b>Bœuf – Bovin</b>	2011	Médecine vétérinaire	(Han et al., 2011) (Deng et al., 2012)
<b>Mouton – Ovin</b>	2011	Médecine vétérinaire	(Bao et al., 2011) (Ren et al., 2011)
<b>Chauve-souris</b>	2014	Médecine vétérinaire	(Mo et al., 2014)
<b>Rhinocéros</b>	2011	Conservation d'espèce en voie d'extinction	(Friedrich Ben-Nun et al., 2011)
<b>Léopard des neiges</b>	2012	Conservation d'espèce en voie d'extinction	(Verma et al., 2012)

Tableau 3 : Liste non exhaustive des espèces dont les cellules ont pu être reprogrammées par la technologie des iPS.

### 3.6 Différents cocktails de reprogrammation

C'est en testant des combinaisons de vingt-quatre gènes candidat que S. Yamanaka et K. Takahashi ont découvert en 2006 le cocktail historique de la reprogrammation (Takahashi and Yamanaka, 2006). Cette combinaison (OCT3/4, SOX2, KLF4 et c-MYC) est également connue sous l'acronyme OSKM. L'expression transitoire de ces quatre facteurs permet le changement dans l'expression des protéines et des modifications épigénétiques profondes des cellules somatiques (Hewitt and Garlick, 2013; Schmidt and Plath, 2012).

J.A. Thomson a, dès 2007, proposé un autre cocktail de reprogrammation : OCT3/4, SOX2, NANOG et LIN28 (Yu et al., 2007), dont l'anagramme est OSLN. L'équipe américaine

a démontré que LIN28 influence la fréquence de reprogrammation. En effet, cette protéine n'est pas requise pour l'initiation de la reprogrammation ni pour l'expansion des cellules reprogrammées. Cependant les deux isoformes LIN28A et LIN28B, toutes deux protéines liées à l'ARN, jouent un rôle, l'isoforme B précoce et la A tardive, dans la reprogrammation (Zhang et al., 2016).

L'équipe de Jean-Marc Lemaître a ensuite combiné ces deux cocktails : OCT3/4, SOX2, KLF4, c-MYC, NANOG et LIN28. Ce nouveau cocktail de six facteurs a permis de rajeunir des cellules centenaires et sénescents (Lapasset et al., 2011), ce que ne permet pas le cocktail de Yamanaka ne permet pas leur reprogrammation.

Dès la découverte du cocktail de Yamanaka, de nombreuses équipes se sont employées à remplacer ces facteurs afin de pouvoir augmenter l'efficacité de la reprogrammation. Le tableau ci-après (Ho et al., 2011) regroupe les différents ARN et protéines capables de substituer les facteurs de reprogrammation de Yamanaka.

<b>Facteur original</b>	<b>Facteur de substitution</b>	<b>Mécanisme d'action</b>	<b>Références</b>
<b>c-Myc</b>	n-Myc ou L-Myc	Homologues qui auraient une fonction redondante	(Nakagawa et al., 2008; 2010)
	Wnt3a	c-Myc est une cible de la voie Wnt	(Marson et al., 2008)
	miR-291-3p, -294 ou -295	miR290 est une cible de c-Myc	(Judson et al., 2009)
<b>Klf4</b>	Klf1, Klf2 ou Klf5	Homologues qui auraient une	(Nakagawa et



		fonction redondante	al., 2008)
	Esrrb ou Esrrg	Esrrb régule l'expression de Klf4, a des cibles de liaison similaires à Oct4, Sox 2 et Klf4 et forme un complexe protéique avec Oct4	(Chen et al., 2008; Feng et al., 2009a; van den Berg et al., 2008)
<b>Sox2</b>	Sox1, Sox3, Sox15 ou Sox18	Homologues qui auraient une fonction redondante	(Nakagawa et al., 2008)
	Nanog	Nanog a des cibles de liaison similaires à Sox2	(Ichida et al., 2009)
<b>Sox2 et c-Myc</b>	Smad7	Inhibe la voie de signalisation Tgf $\beta$ et promeut la transition épithélio-mésenchymateuse	(Ichida et al., 2009; Li et al., 2010; Maherli and Hochedlinger, 2009)
<b>Oct4</b>	Nr5a1 ou Nr5a2	Lie le promoteur et l'enhancer d'Oct4 et régule son expression	(Heng et al., 2010)

Tableau 4 : Liste des ARN et protéines capables de remplacer les facteurs de reprogrammation (d'après (Ho et al., 2011)).

Il est également possible de remplacer les facteurs de reprogrammation par des molécules chimiques. L'utilisation de ces produits peut être résumée dans le tableau

suivant (Feng et al., 2009b; Maherali and Hochedlinger, 2008).

<b>Molécules, Facteurs</b>	<b>Mode d'action</b>	<b>Rôle dans la reprogrammation</b>	<b>Références</b>
BIX-01294	Inhibiteur de la méthyltransférase	Améliore la reprogrammation, possibilité de remplacer Sox2	(Shi et al., 2008a; 2008b)
BayK8644	Agoniste de canal Calcium type-L	Améliore la reprogrammation avec BIX-01294 (x15)	(Shi et al., 2008a)
RG108	Inhibiteur de l'ADN méthyltransférase	Améliore la reprogrammation avec BIX-01294 (x30)	(Shi et al., 2008a)
AZA (5-azacytidine) shRNA contre Dnmt1	Inhibiteur de l'ADN méthyltransférase	Augmente l'efficacité et promeut la reprogrammation complète	(Huangfu et al., 2008a; Mikkelsen et al., 2008)
Dexamethasone	Glucocorticoïde	Améliore l'efficacité de AZA	(Huangfu et al., 2008a)
Acide Valproïque	Inhibiteur de l'histone désacétylase	Améliore la reprogrammation (x100)	(Huangfu et al., 2008a; 2008b)

TSA (Trichostatin A)	Inhibiteur de l'histone désacétylase	Améliore la reprogrammation	(Huangfu et al., 2008a)
SAHA (suberoylanilide hydroxamic acid)	Inhibiteur de l'histone désacétylase	Améliore la reprogrammation	(Huangfu et al., 2008a)
PD0325901 + CHIR99021	Inhibiteur respectivement de MEK et GSK3	Améliore la reprogrammation	(Silva et al., 2008)
Wnt3a	Signalisation cellulaire	Améliore la reprogrammation	(Marson et al., 2008)
siRNA contre p53 et Utf1 cDNA	Respectivement suppresseur de tumeur et cofacteur spécifique d'ES	Améliore la reprogrammation (x100)	(Zhao et al., 2008)

Tableau 5 : Liste des facteurs et molécules chimiques capables d'améliorer la reprogrammation (adapté de (Feng et al., 2009b; Maherali and Hochedlinger, 2008)).

### 3.7 Différentes méthodes de reprogrammation

Il existe plusieurs vecteurs permettant d'exprimer les facteurs nécessaires à la reprogrammation. La technique historique de Yamanaka utilise des rétrovirus (Takahashi and Yamanaka, 2006). Mais le gros désavantage de cette méthode est que le virus est

intégratif. Il y a donc un risque non négligeable d'endommager l'information génétique des cellules à reprogrammer. De plus, l'efficacité de la reprogrammation est très faible, comprise entre 0,05 et 0,1 %.

De nombreuses équipes ont travaillé afin d'augmenter l'efficacité de la reprogrammation. Par exemple, des lentivirus ont été utilisés avec succès (Sommer et al., 2009). D'autres techniques de reprogrammation, bien que fonctionnelles, montrent un taux de reprogrammation très faible, comme les protéines, les adénovirus, les transfections de plasmides, les vecteurs épisomaux.

Outre l'efficacité de la reprogrammation, sa sécurité est également recherchée. En effet, l'intégration de transgènes peut interrompre ou activer un gène de façon aléatoire. Ce qui peut alors compromettre les applications des iPS produites comme le criblage de drogues, l'étude des voies de signalisation dans les maladies et compromettre la sécurité des cellules pour des applications cliniques.

Découvert dans les années cinquante, le virus sendai (SeV) se montre être un outil remarquable (Nakanishi and Otsu, 2012). Il appartient à la famille des paramyxovirus, il est non intégratif et se réplique dans le cytoplasme des cellules qu'il infecte. Il n'y a donc pas de problèmes d'intégration (Fusaki et al., 2009) et il permet une forte expression de gènes exogènes. Ce virus présente donc toutes les caractéristiques pour être considéré comme un outil majeur dans la reprogrammation. Cependant, si les quatre facteurs de reprogrammation sont sur quatre vecteurs différents, les différences d'expression des gènes peuvent altérer la qualité des iPS fabriquées (Carey et al., 2011).

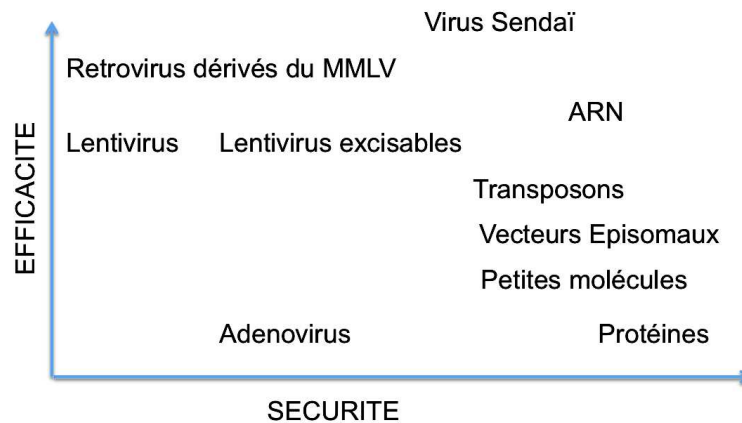


Figure 19 : Evaluation des techniques de reprogrammation en fonction de leur efficacité et de leur sécurité (adapté de (Gonzalez et al., 2011)).

Concernant les différentes techniques qui existent pour reprogrammer des cellules en iPS, il existe des différences non seulement au niveau de l'efficacité mais également au niveau de la sécurité de reprogrammation (Figure 17). A ce jour l'utilisation des virus de Sendai semble être le compromis le plus acceptable entre sécurité et efficacité. Il faudra néanmoins améliorer cette technique afin de pouvoir obtenir des iPS de grade clinique utilisables en thérapie humaine.

### 3.8 Différentes cellules à reprogrammer

En 2006, l'équipe de Yamanaka a reprogrammé des fibroblastes de souris (Takahashi and Yamanaka, 2006) et en 2007 ce sont également des fibroblastes humains qu'ils ont utilisés (Takahashi et al., 2007).

Une biopsie de peau permet de produire des lignées de fibroblastes (Figure 18). Chez l'homme, une biopsie de petite taille (punch) est réalisée à la face interne du bras et son diamètre est de quelques millimètres. La biopsie est disséquée et mise dans un milieu de culture favorisant la dérivation des fibroblastes. Après quelques jours, des cellules

émergent des morceaux de peau et prolifèrent rapidement. Il faut alors s'assurer qu'il s'agit bien de fibroblastes – et non pas de kératinocytes; l'expression de Vimentine est caractéristique de des fibroblastes et permet de les discriminer des kératinocytes (Figure 19). La lignée de fibroblaste est caractérisée sur le plan génétique par un caryotype et congelée.

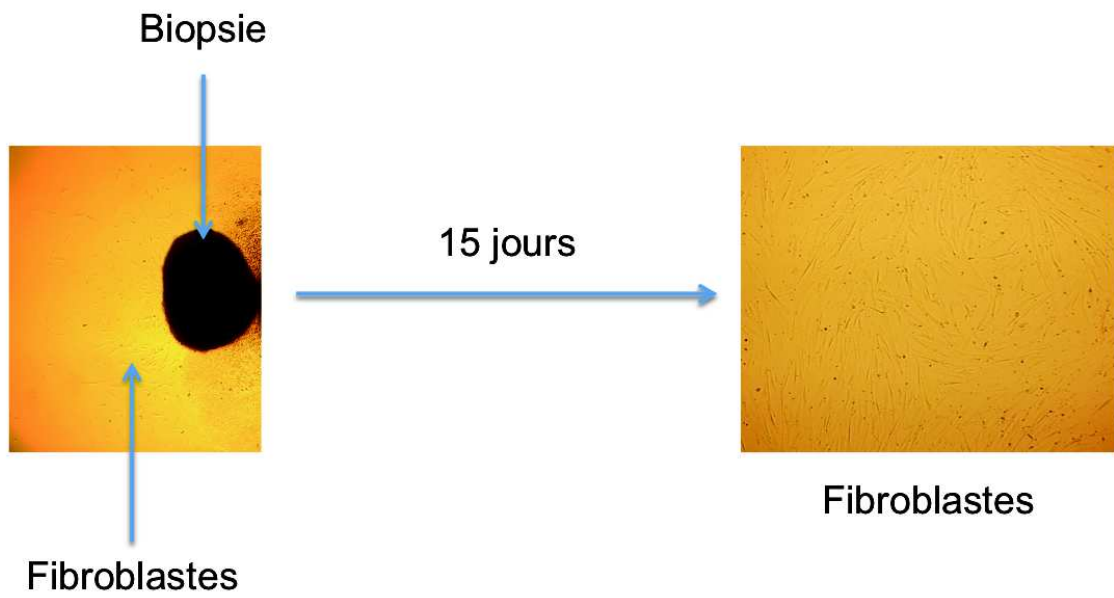


Figure 20 : Dérivation de fibroblastes, lignée FMO2.

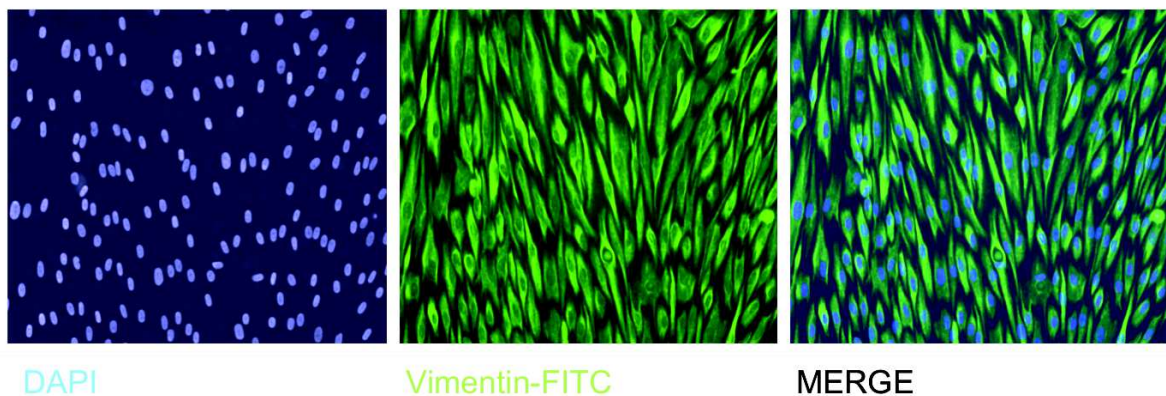


Figure 21 : Immunofluorescence permettant la mise en évidence de l'expression de vimentine par les fibroblastes dérivés.

Depuis les publications initiales sur les iPS, de nombreux types cellulaires différents ont été reprogrammés avec succès, à partir de types cellulaires plus ou moins faciles à reprogrammer, et plus ou moins faciles à obtenir (Figure 20). En effet, la biopsie pour obtenir des fibroblastes reste un acte légèrement invasif et il peut laisser une cicatrice. Par comparaison, les cellules du sang peuvent également se reprogrammer et s'obtiennent par une simple prise de sang (Eminli et al., 2009). D'autres cellules, plus atypiques, ont également été reprogrammées telles que des cellules fœtales humaines (Spinelli et al., 2013), des cellules du follicule pileux (Petit et al., 2012), des cellules de l'estomac et du foie (Aoi et al., 2008), des cellules pancréatiques (Stadtfield et al., 2008a), des lymphocytes (Hanna et al., 2008), des kératinocytes (Aasen et al., 2008), des progéniteurs neuraux (Eminli et al., 2008; Kim et al., 2008), des cellules de la pulpe dentaire (Tamaoki et al., 2010), des cellules issues du liquide amniotique (Li et al., 2009), des astrocytes (Ruiz et al., 2010). Cette liste des cellules reprogrammées avec succès à ce jour n'est pas exhaustive.

Il est à noter qu'il existe un type cellulaire se reprogrammant par l'expression transitoire d'un seul facteur. En effet, la surexpression d'OCT4 permet la reprogrammation des cellules souches neurales en iPS (Kim et al., 2009a).

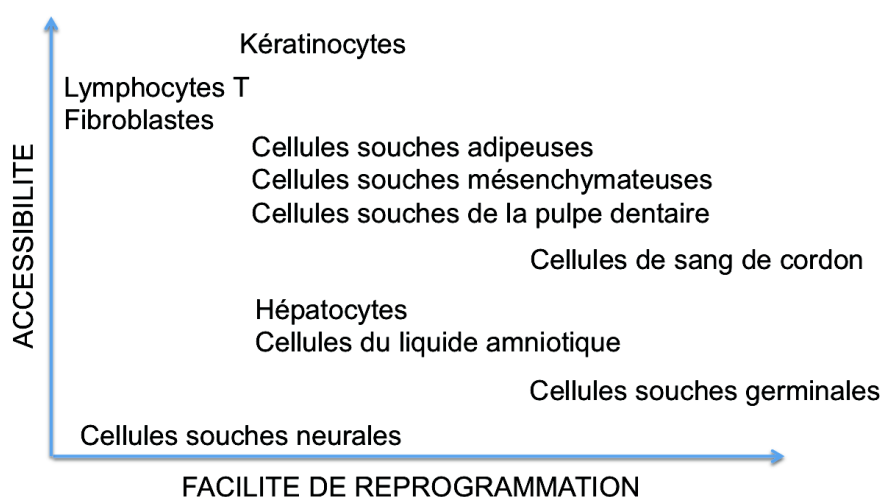


Figure 22 : Evaluation de la reprogrammation des cellules en fonction de leur accessibilité et de leur facilité à se reprogrammer (adapté de (Gonzalez et al., 2011)).

En résumé, pour fabriquer des iPS, plusieurs choix s'offrent aux scientifiques concernant les cellules à reprogrammer, les facteurs de reprogrammation et la méthode d'induction (Figure 21).

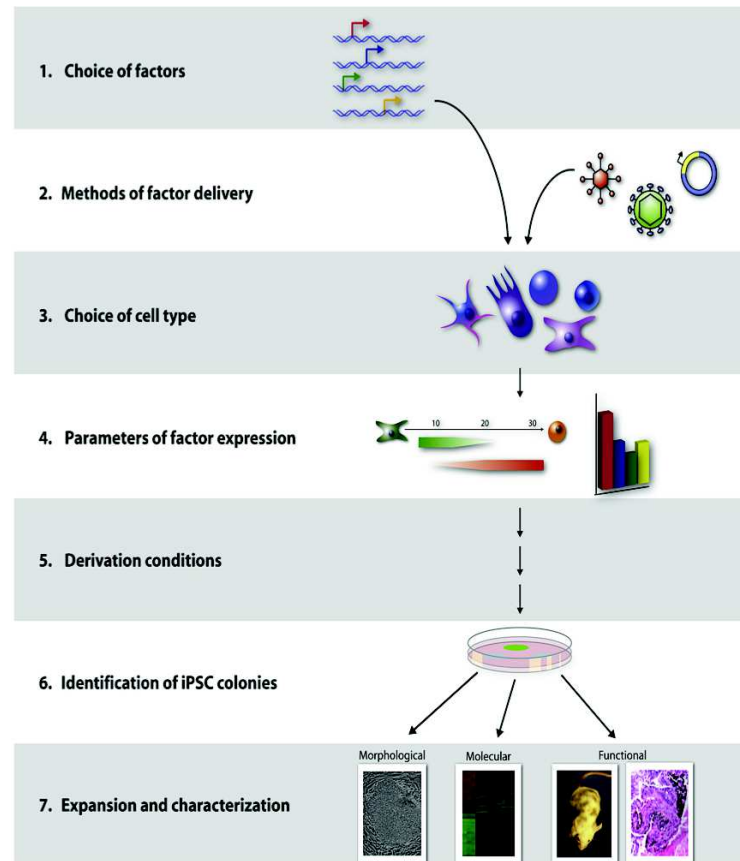


Figure 23 : Schéma résumant les différents choix possibles à chaque étape dans le but d'obtenir des iPS (adapté de (Maherali and Hochedlinger, 2008)).

### 3.9 Les applications des iPS

Les applications des iPS sont nombreuses et diverses. Il est possible de les regrouper en neuf catégories :

- étude du déterminisme cellulaire



- étude de la pluripotence
- étude du développement humain
- étude de la physiologie des tissus et organes
- étude des maladies génétiques *in vitro*
- outil de rajeunissement cellulaire
- outil de screening de candidats-médicaments
- source de cellules pour la médecine régénératrice et pour la production d'organes humains.

Nous avons détaillé ces 9 applications dans un article de revue.

**Article 1** : « Human induced pluripotent stem cells : A disruptive innovation »

Publié dans *Curr Res Transl Med* en 2016.



Available online at  
**ScienceDirect**  
[www.sciencedirect.com](http://www.sciencedirect.com)

Elsevier Masson France  
**EM|consulte**  
[www.em-consulte.com](http://www.em-consulte.com)



General review

## Human induced pluripotent stem cells: A disruptive innovation



J. De Vos<sup>a,b,c,d,e,f,\*</sup>, J. Bouckenheimer<sup>a,b,g</sup>, C. Sansac<sup>a,b,g</sup>, J.-M. Lemaître<sup>a,b</sup>, S. Assou<sup>a,b,c,\*</sup>

<sup>a</sup> CHU Montpellier, Institute for Regenerative Medicine and Biotherapy, Hôpital Saint-Eloi, 34000 Montpellier, France

<sup>b</sup> INSERM, U1183, 34000 Montpellier, France

<sup>c</sup> Université de Montpellier, UFR de Médecine, 34000 Montpellier, France

<sup>d</sup> Institut de Biologie Computationnelle, 34000 Montpellier, France

<sup>e</sup> CHU Montpellier, SAFE-IPS Reprogramming Platform, Institute of Research in Biotherapy, 34000 Montpellier, France

<sup>f</sup> CHU Montpellier, Unit for Cellular Therapy, Hospital Saint-Eloi, 34000 Montpellier, France

<sup>g</sup> Université de Montpellier, UFR de Pharmacie, 34000 Montpellier, France

### ARTICLE INFO

#### Article history:

Received 16 February 2016

Accepted 8 April 2016

Available online 25 May 2016

#### Keywords:

Stem cells

Induced pluripotent stem cells

Cell reprogramming

Pluripotency

Trans-differentiation

Regenerative medicine

### ABSTRACT

This year (2016) will mark the 10th anniversary of the discovery of induced pluripotent stem cells (iPSCs). The finding that the transient expression of four transcription factors can radically remodel the epigenome, transcriptome and metabolome of differentiated cells and reprogram them into pluripotent stem cells has been a major and groundbreaking technological innovation. In this review, we discuss the major applications of this technology that we have grouped in nine categories: a model to study cell fate control; a model to study pluripotency; a model to study human development; a model to study human tissue and organ physiology; a model to study genetic diseases in a dish; a tool for cell rejuvenation; a source of cells for drug screening; a source of cells for regenerative medicine; a tool for the production of human organs in animals.

© 2016 Elsevier Masson SAS. All rights reserved.

### 1. Introduction

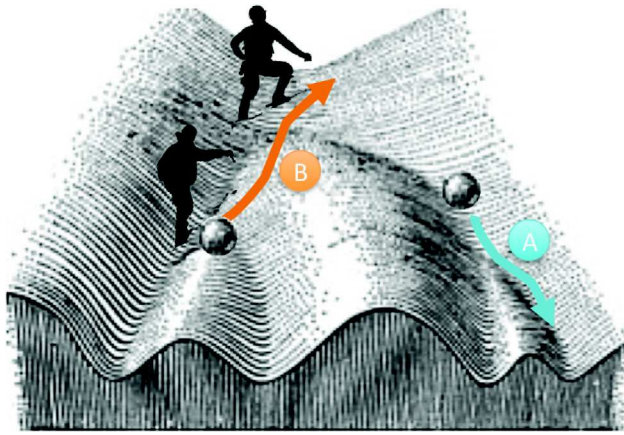
Pluripotency is the capacity of a cell to differentiate into any cell type. In human cells, this feature is observed only in the blastocyst inner cell mass (ICM), during a very short time-window in early embryo development. Remarkably, it has been possible to ‘freeze’ this very transient developmental feature by identifying the culture conditions required to maintain pluripotency *in vitro*. This allowed the derivation of embryonic stem cells (ESCs) that self-renew indefinitely in a Petri dish, while maintaining their capacity to differentiate into all varieties of cells that form human tissues [1,2]. ESCs are pluripotent stem cells (PSC) and have opened the door to many applications, including mimicking normal tissue development *in vitro* or producing differentiated cells for regenerative medicine. However, ESCs have two main drawbacks. The derivation of human ESC lines is hampered by ethical issues concerning the destruction of human embryos. Moreover, it is exceedingly difficult to derive ESC lines with a specific genotype

and almost impossible to obtain an autologous ESC line from a patient.

Normal human development entails the progressive specialization of undifferentiated cells into differentiated, mature cells. This process is tightly regulated by various transcription factors and non-coding RNAs as well as epigenetic remodeling. The inherent irreversibility of differentiation is required for maintaining the delicate architecture of multicellular organisms, and is schematized by Waddington (Fig. 1). However, somatic cell nuclear transfer (SCNT) studies have demonstrated that the epigenetic status of a differentiated cell can be reversed, even in mammals [3,4]. Similarly, the cell fate of a differentiated somatic cell can be brought back toward pluripotency upon fusion with one ECS [5]. Starting from these and other findings, Shinya Yamanaka and Kazutoshi Takahashi demonstrated that the transient overexpression of only four transcription factors (OCT4, SOX2, KLF4 and cMYC) can profoundly modify the transcriptome, epigenome and metabolome of differentiated cells and reprogram them into induced pluripotent stem cells (iPSCs), a new PSC type [6–8]. These results were rapidly reproduced in several other laboratories [9–12] and Shinya Yamanaka was awarded the Nobel Prize of Medicine in 2012 for this discovery [13]. Although there have been discussions on whether iPSCs are equivalent to ES cells and some epigenetic and genetic defects have been detected in iPSCs (depending on the reprogramming protocol used), the emerging

\* Corresponding authors. Institute for Research in Regenerative Medicine and Biotherapy, Hôpital Saint-Eloi, 80, Avenue Augustin-Fliche, 34295 Montpellier cedex 5, France. Fax: +33 4 67 33 78 88.

E-mail addresses: [john.devos@inserm.fr](mailto:john.devos@inserm.fr) (J. De Vos), [said.assou@inserm.fr](mailto:said.assou@inserm.fr) (S. Assou).

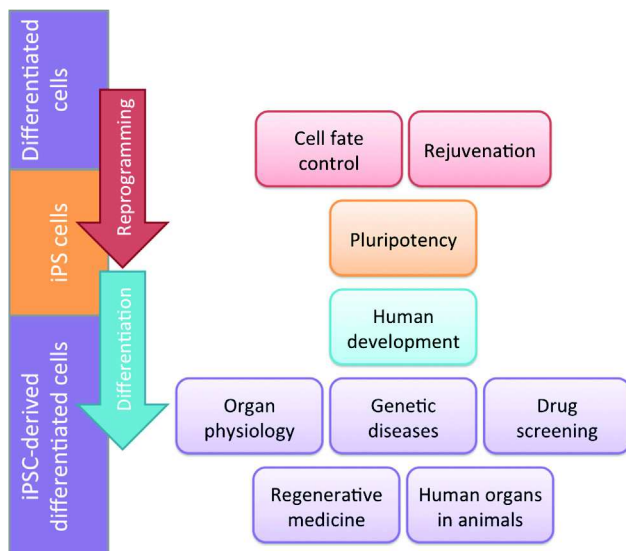


**Fig. 1.** Applications of the iPSC technology. The Waddington's diagram revisited [96]. A. Cell differentiation is represented by a metal ball rolling down a valley that branches into smaller valleys. This illustrates the irreversibility of the differentiation process during development and adulthood whereby a differentiated cell cannot spontaneously dedifferentiate (roll upwards) or trans-differentiate into another cell type (roll to an adjacent valley). B. Cell reprogramming is the process by which a specialized cell is de-differentiated into a pluripotent cell at the summit of the Waddington diagram.

agreement is that iPSCs are a good model to study pluripotency and may be identical or nearly identical to ES cells [14–17]. This year will mark the 10th anniversary of this discovery. The aim of this review is to discuss the most important present and future applications of iPSCs (Fig. 2).

## 2. Understanding cell fate control

The cell phenotype is remarkably stable despite all environmental perturbations and stochastic variability in gene expression. Hence, after commitment and differentiation, the identity of a tissue is definitive and robust. This stability is absolutely necessary. One can easily imagine the catastrophe if skin cells, for instance, could easily turn into blood cells or vice versa! To



**Fig. 2.** Applications of the iPSC technology. The basic research and medical applications of human iPSCs can be divided in nine categories: (1) cell fate control, (2) cell rejuvenation, (3) pluripotency, (4) organ development, (5) organ physiology, (6) genetic diseases, (7) drug screening, (8) regenerative medicine, (9) human organs in animals. These categories are depicted as boxes displayed side by side with the main steps of iPSC generation and differentiation: reprogramming, pluripotency maintenance, differentiation and differentiated cells/tissues.

avoid this, the cell epigenome and transcriptome – including miRNAs – can withstand random perturbations of their milieu and maintain their identity. This is also true during development where differentiation leads to stereotyped phenotypic outcomes, a process termed canalization [18]. It was initially thought that differentiation was irreversible in mammals [19], but SCNT studies demonstrated that differentiation could be completely reversed also in mammals. Compared with SCNT, cell reprogramming and iPSC generation offer a more convenient tool to manipulate and study cell fate control. Currently, much effort is focused on unraveling the precise molecular mechanisms underlying cell reprogramming, ultimately to understand cell fate control [20–24].

Some earlier reports showed that a specialized cell type could trans-differentiate into another one upon expression of specific transcription factors [25,26]. Accordingly, iPSCs can be seen as the ultimate demonstration of cell fate plasticity through modulation of transcription factor expression. For this reason, the iPSC technology has significantly boosted the field of trans-differentiation directly [27] (transient expression of the “Yamanaka’s cocktail” of transcription factors favors trans-differentiation) and indirectly, by changing the cell fate paradigm [28,29].

Moreover, the iPSC technology also contributed technically to facilitate and optimize cell identity manipulation. The initial iPSC lines were obtained by transient expression of transcription factors from murine retroviruses [6,7]. Many improvements and/or modifications of the initial protocol have been described. For example, the use of non-integrated vectors, particularly Sendai viruses, is now the preferred technique. In the future, this improvement could be applied also to the field of trans-differentiation. Techniques that avoid the use of viruses have also been developed for efficient iPSC generation, such as the use of synthetic mRNAs [30], cell-permeant proteins [31] or small molecule compounds [32,33]. Again, such approaches may be transferred to the field of trans-differentiation.

## 3. Cell rejuvenation

Human iPSCs can be obtained from somatic cells of aged donors and even centenarians, and, importantly, the crucial markers of ageing are reversed by cell reprogramming [34,35]. For instance, telomere length in fibroblasts derived from aged donors is extended after cell reprogramming. Even cell senescence can be reversed, as indicated by the disappearance of senescence-associated heterochromatic foci (SAHF), inhibition of p53/p21CIP1 and p16INK4A/pRB activation, disappearance of senescence-associated  $\beta$ -galactosidase (SA- $\beta$ -Gal) activity and restoration of cell proliferation [34]. In the field of ageing, the possibility of extensively erasing markers of ageing is a major step toward the goal of rejuvenating tissues or even individuals [36,37].

Moreover, iPSCs are a unique tool to study age-related diseases in vitro [38]. For instance iPSCs have been used in the study of laminopathy-associated lamin A mutations and telomere diseases [39–41].

## 4. Studying pluripotency

As iPSCs are nearly identical, if not identical to ESCs, they can be used to study pluripotency instead of ESCs, thus circumventing the difficulties and ethical concerns concerning the donation/use of human embryos for research. Indeed, iPSCs have replaced ESCs for assessing chromatin modifications [42,43], nuclear genome organization [44], metabolome [45], cytoplasmic complexity [46], early human development [47] and for devising new differentiation protocols or studying organ development (see below).

## 5. Tissue and organ development and physiology

The key property of iPSCs is their ability to differentiate *in vitro* into any of the cell types derived from the three germ layers upon sequential addition of growth factors, small factors, co-cultured cells or specific matrices. Finding the optimum combination of factors to direct differentiation relies on the precise understanding of normal development sequences in human or model mammals, and requires mimicking the conditions of embryonic and fetal human development. For human cells, this field of research started with the discovery of human ESCs almost twenty years ago, and now human iPSCs are often used instead of ESCs. By reproducing cell differentiation in a dish, iPSCs provide a mean to study human tissue development and physiology. This is a major breakthrough. Indeed, before iPSC discovery, human developmental biology relied on animal models and on the study of fragments of aborted human fetuses, or biopsies of adult tissues, most of which would not survive nor grow *in vitro*. Protocols to differentiate human iPSCs into a large array of cell types have been published and some selected examples are given in Table 1. For instance, it is now possible to generate and maintain *in vitro* beating cardiomyocytes starting from human iPSCs. This opened the way to study cardiac development, including the consequences of perturbing specific signaling pathways or genetic regulations. The technology is sufficiently reliable for large scale production of functional cardiomyocytes for electrophysiology and drug testing applications (see below) [48].

However, a major limitation of the current iPSC systems is that most protocols differentiate PSCs in 2D monolayers, while human tissues are organized in 3D structures. In 2D systems, cell-to-cell contacts, mechanical forces and signaling cascades as well the cell diversity of a growing tissue are imperfect and limited. As a consequence, also the cell signaling pathways that are normally required for proper differentiation cannot be fully reproduced in 2D culture systems. This could partly explain the common paradox that PSC lines derived from different donors (i.e., healthy individuals with different genotypes) differentiate into specific tissues with variable efficiencies [49,50]. One likely explanation is that these 2D *in vitro* differentiation systems lack some of the signals required to promote proper cell differentiation. Consequently, the buffering systems that may compensate the effects of genetic diversity in differentiation pathways also are defective [17].

To circumvent these 2D limitations, recent work focused on adapting 2D PSC differentiation protocols to 3D culture systems that allow the development of structures mimicking organogenesis (commonly termed “organoids”) [51]. This strategy has already been used to model the development of kidney [52],

stomach [53], retinal tissue [54], cerebellum [55] or brain tissue [56]. Organoids display complex morphogenetic patterning and include multiple cell types that are organized like during human development. Therefore, organoids represent a formidable tool for understanding human development. Nevertheless, *in vitro* 3D models are still imperfect as they lack capillaries, which are major players in normal tissue development and tissue homeostasis. Interestingly, it is possible to incorporate capillaries and blood cells in developing organoids by mixing differentiating iPSCs with mesenchymal and endothelial cells and transplanting this cell mixture in highly vascularized tissues, such as the kidney capsule or the mesentery, of immuno-deficient mice. This approach was initially developed by Takebe et al. to produce vascularized “liver buds” [57]. This system restores the physiological relationship between the cells of interest and endothelial cells, and also allows addressing specific issues, such as the precise level of oxygen or nutrients found *in vivo*. An even more sophisticated approach to generate vascularized 3D human tissues from PSCs in animals is described in the chapter “Human organs in animals”.

## 6. Modeling human genetic diseases in a dish

The use of PSCs to study human tissue development and physiology started with the discovery of human ESCs [2,58]. However, iPSCs much simplified the study of human tissue development and physiology in genetically altered cells. Indeed, the genetic background of human ESC lines cannot be chosen in advance because it is the result of the combination of two haploid gametes. By contrast, the donor (and thus the genotype) of the differentiated cells used to derive iPSC lines can be selected at will. For instance, a genetic disease can now be modeled *in vitro* by deriving iPSCs from a patient with that disease. Many examples illustrate this powerful application of this technology. Electrophysiological cardiomyocyte defects, such as those observed in the long-QT syndrome, can be reproduced *in vitro* using patient-specific differentiated iPSCs. This offers the unique possibility of screening small compounds directly on beating cardiomyocytes with the disease genotype [48]. Even more complex diseases, such as Rett syndrome, an autism spectrum disorder, can be, at least partially, reproduced *in vitro*. Analysis of neuronal networks showed that neurons differentiated *in vitro* from iPSCs derived from patients with Rett syndrome have fewer synapses, reduced spine density and electrophysiological defects [59]. Moreover, besides preclinical studies in animal models, iPSCs can serve as an *in vitro* model of a human diseased tissue to test gene therapy approaches using human cells before their clinical transfer [60].

## 7. Drug screening

Drug discovery is a complex and time-consuming process that still largely relies on primary or immortalized cells and animals testing [61]. iPSCs present several important advantages for drug discovery. Indeed, they are diploid cells with unlimited proliferation capacity; they can differentiate into a wide array of tissues, and can readily be genetically manipulated. In addition, the iPSC genotype (normal or pathological) can be selected by choosing the cell donor. Therefore, normal or patient-specific iPSCs may offer a relevant, scalable and cost-effective model for drug discovery. There are several example of small molecule high-throughput screens (HTS) using iPSCs that identified novel candidate therapeutic compounds for genetic diseases, including Rett syndrome [59], Parkinson's disease [62] and FGFR3 skeletal dysplasia [63]. Moreover, iPSCs are disease-relevant cell models that can be used to validate small compounds identified by HTS in

**Table 1**

Selected references on the use of iPSCs for modeling normal human tissue development and physiology.

Human tissue	Reference
Neurons	[82]
Retinal pigment epithelium (RPE)	[83]
Skin	[84]
Melanocytes	[85]
Cardiomyocytes	[48]
Skeletal muscle	[86]
T lymphocytes	[87]
Red blood cells	[88]
Endothelial cells	[89]
Cartilage	[90]
Adipocytes	[91]
Kidney	[52]
Bronchial epithelium	[92]
Liver	[93]
Thyroid	[94]
Pancreatic beta cells	[95]

**Table 2**

The main, past and on-going, clinical trials using human ESCs or iPSCs.

PSC-derived	Cell type	Trial start date–estimated completion	Identifier or reference	Sponsor
Oligodendrocyte progenitors (GRNOPC1)	hESC	2010–2011	NCT01217008	Geron Corp.
Retinal pigmented epithelial	hESC	2011–2013	NCT01345006	Advanced Cell Technology
Retinal pigmented epithelial	hESC	2011–2014	NCT01344993	Advanced Cell Technology
Retinal pigmented epithelial	hESC	2011–2014	NCT01469832	Advanced Cell Technology
Retinal pigmented epithelial	hESC	2012–2014	NCT01625559	CHA Bio&Diostech
Retinal pigmented epithelial	hESC	2012–2016	NCT01674829	CHA Bio&Diostech
Retinal pigmented epithelial	hESC	2013–2015	NCT01691261	Pfizer/University College London
Human embryonic stem cell-derived CD15+ ISL-1+ progenitors	hESC	2013–2017	NCT02057900	AP-HP/France
Autologous iPSC-cell-derived retinal pigment epithelium (RPE)	hiPSC	2014–2015	[77]	RIKEN, Japan

Based on the <https://clinicaltrials.gov> web site and cited references.

immortalized cells. For instance, Naryshkin et al. used iPSCs to test three small compounds that promote the inclusion of exon 7 in SMN2 mRNA and that they discovered in immortalized cells. They could show that one of these compounds significantly increased SMN protein levels in motor neurons differentiated from iPSCs derived from patients with spinal muscular atrophy type I and II [64].

Finally, iPSCs can also be used to test the potential toxicity on organs, such as heart, brain or lung, or the teratogenicity of a compound, thus avoiding or strongly limiting the cumbersome, expensive and ethically sensitive animal testing step [65,66].

## 8. Regenerative medicine

The aim of regenerative medicine is to restore or establish normal function in a damaged organ by replacing or regenerating cells, tissues or the whole organ [67]. Cellular therapy is one approach to achieve this goal, but it has been hampered by the critical issue of the cell source. As iPSCs can virtually produce any cell type in vitro, they represent an invaluable asset. Moreover, iPSCs provide something ESCs cannot, namely the possibility of producing autologous cells, thus alleviating the need of finding a HLA-compatible cell donor and of immunosuppression.

Many studies in animal models have demonstrated the feasibility of the clinical applications of PSC-derived cells. PSCs have been used for long-term glycemic control in diabetic immune-competent mice [68], for rescuing hepatic function in animals with severe liver damage [69], to improve contractility after myocardial infarction [70], to protect the degenerating retina in mice homozygous for the *rd1* mutation [71], to restore vision in *Crx*-deficient mice [72], and to completely restore amphetamine-induced rotation behavior and improve akinesia in rodent models of Parkinson's disease [73].

Based on PSC promise for cellular therapy and the success in animal models, several clinical trials have been started world-wide using human ESCs or iPSCs (Table 2). Based on the published and unpublished data available to date, no serious adverse event associated with the injection of differentiated ESCs or iPSCs has been reported [74–77]. Moreover, some encouraging results have been observed in patients with retinal diseases after treatment with hESC-derived retinal pigment epithelium [76].

However, many obstacles still need to be overcome before PSC-based cellular therapy will become a routine clinical practice. These hurdles include the difficulty of differentiating PSCs in vitro beyond the fetal stage for many tissue types, the challenge for the injected cells to survive in vivo and to integrate the organ to regenerate its function [78], and the high costs of autologous iPSC-based treatments. Another significant issue is the control of iPSC

genetic integrity, which contributed to the suspension of the first clinical trial based on cells differentiated from iPSCs [77,79].

## 9. Human organs in animals

As an alternative to PSC-based cellular therapy, human organs could be grown in animals [80]. The proof-of-concept of this approach was the demonstration that it is possible to grow a rat pancreas in mice by injecting wild type rat iPSCs in *Pdx1*<sup>-/-</sup> mouse blastocysts (*Pdx1*<sup>-/-</sup> mice are born without pancreas). Rat iPSCs colonize minimally each mouse organ, but for the pancreas niche, only the rat iPSC-derived cells contribute to pancreas formation because of mouse *Pdx1* absence. These chimeric animals survive after birth by relying on their rat pancreas, whereas non-chimeric *Pdx1*<sup>-/-</sup> mouse pups rapidly die after birth because of pancreas agenesis. These results have led to the hypothesis that similar results could be obtained using pigs as carriers and human iPSCs as donor cells. For instance, by injecting human iPSCs in *Pdx1*<sup>-/-</sup> pig blastocysts, it may be possible to produce a human pancreas in pigs. Other organs, such as kidneys and heart, could also be grown using this approach [81]. Nonetheless, many obstacles will need to be surmounted before this hypothesis becomes reality, including the need of an inter-specific signaling compatibility high enough to support the development of a human organ in the pig cellular environment, the risk of pig cell contribution (such as blood cells, endothelial cells or nervous cells) to the human organ, the risk of transmitting zoonoses to the human recipient, and of course, serious ethical concerns [80].

## 10. Conclusion

The iPSC technology has clearly demonstrated that it is possible to alter cell fate in the most extreme possible way, i.e., by driving dedifferentiation back to pluripotency. It is a disruptive technology because it is now theoretically possible to produce any cell type in vitro. The breadth of applications based on iPSCs (catalogued as nine applications in this review) illustrates the major impact of cell reprogramming in biology and medicine. In addition, the discovery of iPSCs significantly boosted the related field of directed trans-differentiation. Most iPSC applications discussed here are already widely implemented; while others, such as drug screening, regenerative medicine and production of human organs in animals are still in their early days.

## Disclosure of interest

The authors declare that they have no competing interest.

## Acknowledgement

Supported by grants from the Agence nationale de la recherche (INGESTEM), the association “Gueules cassées” and the University Hospital of Montpellier (AOI 2013-A00892-43).

## References

- [1] Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. *Nature* 1981;292(5819):154–6.
- [2] Thomson JA, Itskovitz-Eldor J, Shapiro SS, et al. Embryonic stem cell lines derived from human blastocysts. *Science* 1998;282(5391):1145–7.
- [3] Gurdon JB, Elsdale TR, Fischberg M. Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. *Nature* 1958;182(4627):64–5.
- [4] Wilmut I, Schnieke AE, McWhir J, Kind AJ, Campbell KH. Viable offspring derived from fetal and adult mammalian cells. *Nature* 1997;385(6619):810–3.
- [5] Cowan CA, Atienza J, Melton DA, Eggan KC. Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. *Science* 2005;309(5739):1369–73.
- [6] Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006;126(4):663–76.
- [7] Takahashi K, Tanabe K, Ohnuki M, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 2007;131(5):861–72.
- [8] Okita K, Ichisaka T, Yamanaka S. Generation of germline-competent induced pluripotent stem cells. *Nature* 2007;448(7151):313–7.
- [9] Wernig M, Meissner A, Foreman R, et al. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* 2007;448(7151):318–24.
- [10] Yu J, Vodyanik MA, Smuga-Otto K, et al. Induced pluripotent stem cell lines derived from human somatic cells. *Science* 2007;318(5858):1917–20.
- [11] Maherali N, Sridharan R, Xie W, et al. Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* 2007;1(1):55–70.
- [12] Park I-H, Zhao R, West JA, et al. Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* 2008;451(7175):141–6.
- [13] Science AAFTAO. Nobel Prize in physiology or medicine. Reprogrammed cells earn biologists top honor. *Science* 2012;338(6104):178–9.
- [14] Yamanaka S. Induced pluripotent stem cells: past, present, and future. *Cell Stem Cell* 2012;10(6):678–84.
- [15] Choi J, Lee S, Mallard W, et al. A comparison of genetically matched cell lines reveals the equivalence of human iPSCs and ESCs. *Nat Biotechnol* 2015;33(11):1173–81.
- [16] Vallier L. Putting induced pluripotent stem cells to the test. *Nat Biotechnol* 2015;33(11):1145–6.
- [17] Rouhani F, Kumasaka N, de Brito MC, Bradley A, Vallier L, Gaffney D. Genetic background drives transcriptional variation in human induced pluripotent stem cells. *PLoS Genet* 2014;10(6):e1004432.
- [18] Hornstein E, Shomron N. Canalization of development by microRNAs. *Nat Genet* 2006;38:S20–4.
- [19] McGrath J, Solter D. Inability of mouse blastomere nuclei transferred to enucleated zygotes to support development in vitro. *Science* 1984;226(4680):1317–9.
- [20] Cyranoski D. Stem cells: the black box of reprogramming. *Nature* 2014;516(7530):162–4.
- [21] Samavarchi-Tehrani P, Golipour A, David L, et al. Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* 2010;7(1):64–77.
- [22] Cheloufi S, Elling U, Hopfgartner B, et al. The histone chaperone CAF-1 safeguards somatic cell identity. *Nature* 2015;528(7581):218–24.
- [23] Soufi A, Garcia MF, Jaroszewicz A, Osman N, Pellegrini M, Zaret KS. Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* 2015;161(3):555–68.
- [24] Onder TT, Kara N, Cherry A, et al. Chromatin-modifying enzymes as modulators of reprogramming. *Nature* 2012;483(7391):598–602.
- [25] Tapscott SJ, Davis RL, Thayer MJ, Cheng PF, Weintraub H, Lassar AB. MyoD1: a nuclear phosphoprotein requiring a Myc homology region to convert fibroblasts to myoblasts. *Science* 1988;242(4877):405–11.
- [26] Xie H, Ye M, Feng R, Graf T. Stepwise reprogramming of B cells into macrophages. *Cell* 2004;117(5):663–76.
- [27] Zhu S, Rezvani M, Harbell J, et al. Mouse liver repopulation with hepatocytes generated from human fibroblasts. *Nature* 2014;508(7494):93–7.
- [28] Karagiannis P, Yamanaka S. The fate of cell reprogramming. *Nat Methods* 2014;11(10):1006–8.
- [29] Morris SA, Cahan P, Li H, et al. Dissecting engineered cell types and enhancing cell fate conversion via cellnet. *Cell* 2014;158(4):889–902.
- [30] Warren L, Manos PD, Ahfeldt T, et al. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell* 2010;7(5):618–30.
- [31] Lee J, Sayed N, Hunter A, et al. Activation of innate immunity is required for efficient nuclear reprogramming. *Cell* 2012;151(3):547–58.
- [32] Hou P, Li Y, Zhang X, et al. Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science* 2013;341(6146):651–4.
- [33] Zhao Y, Zhao T, Guan J, et al. A XEN-like state bridges somatic cells to pluripotency during chemical reprogramming. *Cell* 2015;163(7):1678–91.
- [34] Lapasset L, Milhavel O, Prieur A, et al. Rejuvenating senescent and centenarian human cells by reprogramming through the pluripotent state. *Genes Dev* 2011;25(21):2248–53.
- [35] Frobel J, Hemedal H, Lenz M, et al. Epigenetic rejuvenation of mesenchymal stromal cells derived from induced pluripotent stem cells. *Stem Cell Reports* 2014;3(3):414–22.
- [36] Wahlestedt M, Norddahl GL, Sten G, et al. An epigenetic component of hematopoietic stem cell aging amenable to reprogramming into a young state. *Blood* 2013;121(21):4257–64.
- [37] Oh J, Lee YD, Wagers AJ. Stem cell aging: mechanisms, regulators and therapeutic opportunities. *Nat Med* 2014;20(8):870–80.
- [38] Lemey C, Milhavel O, Lemaître J-M. iPSCs as a major opportunity to understand and cure age-related diseases. *Biogerontology* 2015;16(4):399–410.
- [39] Liu G-H, Suzuki K, Qu J, et al. Targeted gene correction of laminopathy-associated LMNA mutations in patient-specific iPSCs. *Cell Stem Cell* 2011;8(6):688–94.
- [40] Batista LF, Artandi SE. Understanding telomere diseases through analysis of patient-derived iPSC cells. *Curr Opin Genet Dev* 2013;23(5):526–33.
- [41] Nissan X, Blondel S, Navarro C, et al. Unique preservation of neural cells in Hutchinson-Gilford progeria syndrome is due to the expression of the neural-specific miR-9 microRNA. *Cell Rep* 2012;2(1):1–9.
- [42] Wu T, Liu Y, Wen D, et al. Histone variant H2A.X deposition pattern serves as a functional epigenetic mark for distinguishing the developmental potentials of iPSCs. *Cell Stem Cell* 2014;15(3):281–94.
- [43] Villasante A, Piazzolla D, Li H, Gomez-Lopez G, Djabali M, Serrano M. Epigenetic regulation of Nanog expression by Ezh2 in pluripotent stem cells. *Cell Cycle* 2011;10(9):1488–98.
- [44] de Wit E, Bouwman BAM, Zhu Y, et al. The pluripotent genome in three dimensions is shaped around pluripotency factors. *Nature* 2013;501(7466):227–31.
- [45] Carey BW, Finley LWS, Cross JR, Allis CD, Thompson CB. Intracellular  $\alpha$ -ketoglutarate maintains the pluripotency of embryonic stem cells. *Nature* 2015;518(7539):413–6.
- [46] Ramirez JM, Bai Q, Pequignot M, et al. Side scatter intensity is highly heterogeneous in undifferentiated pluripotent stem cells and predicts clonogenic self-renewal. *Stem Cells Dev* 2013;22(12):1851–60.
- [47] Bai Q, Assou S, Haouzi D, et al. Dissecting the first transcriptional divergence during human embryonic development. *Stem Cell Rev Rep* 2012;8(1):150–62.
- [48] Moretti A, Bellin M, Welling A, et al. Patient-specific induced pluripotent stem-cell models for long-QT syndrome. *N Engl J Med* 2010;363(15):1397–409.
- [49] Boulting GL, Kiskinis E, Croft GF, et al. A functionally characterized test set of human induced pluripotent stem cells. *Nat Biotechnol* 2011;29(3):279–86.
- [50] Kajiwara M, Aoi T, Okita K, et al. Donor-dependent variations in hepatic differentiation from human-induced pluripotent stem cells. *Proc Natl Acad Sci U S A* 2012;109(31):12538–43.
- [51] Huch M, Koo B-K. Modeling mouse and human development using organoid cultures. *Development* 2015;142(18):3113–25.
- [52] Takasato M, Er PX, Chiu HS, et al. Kidney organoids from human iPSC cells contain multiple lineages and model human nephrogenesis. *Nature* 2015;526(7574):564–8.
- [53] McCracken KW, Catá EM, Crawford CM, et al. Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. *Nature* 2014;516(7531):400–4.
- [54] Zhong X, Gutierrez C, Xue T, et al. Generation of three-dimensional retinal tissue with functional photoreceptors from human iPSCs. *Nat Commun* 2014;5:4047.
- [55] Muguruma K, Nishiyama A, Kawakami H, Hashimoto K, Sasai Y. Self-organization of polarized cerebellar tissue in 3D culture of human pluripotent stem cells. *Cell Rep* 2015;10(4):537–50.
- [56] Lancaster MA, Renner M, Martin C-A, et al. Cerebral organoids model human brain development and microcephaly. *Nature* 2013;501(7467):373–9.
- [57] Takebe T, Sekine K, Enomura M, et al. Vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature* 2013;499(7459):481–4.
- [58] Odorico JS, Kaufman DS, Thomson JA. Multilineage differentiation from human embryonic stem cell lines. *Stem Cells* 2001;19(3):193–204.
- [59] Marchetto MCN, Carromeu C, Acab A, et al. A model for neural development and treatment of rett syndrome using human induced pluripotent stem cells. *Cell* 2010;143(4):527–39.
- [60] Cereso N, Pequignot MO, Robert L, et al. Proof of concept for AAV2/5-mediated gene therapy in iPSC-derived retinal pigment epithelium of a choroideremia patient. *Mol Ther Methods Clin Dev* 2014;1:14011.
- [61] McNeish J. Embryonic stem cells in drug discovery. *Nat Rev Drug Discov* 2004;3(1):70–80.
- [62] Ryan SD, Dolatabadi N, Chan SF, et al. Isogenic human iPSC Parkinson's model shows nitrosative stress-induced dysfunction in MEF2-PGC1 $\alpha$  transcription. *Cell* 2013;155(6):1351–64.
- [63] Yamashita A, Morioka M, Kishi H, et al. Statin treatment rescues FGFR3 skeletal dysplasia phenotypes. *Nature* 2014;513(7519):507–11.
- [64] Naryshkin NA, Weetall M, Dakka A, et al. Motor neuron disease. SMN2 splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. *Science* 2014;345(6197):688–93.
- [65] Peric D, Barragan I, Giraud-Triboulet K, et al. Cytostatic effect of repeated exposure to simvastatin: a mechanism for chronic myotoxicity revealed by the use of mesodermal progenitors derived from human pluripotent stem cells. *Stem Cells* 2015;33(10):2936–48.

- [66] Liang P, Lan F, Lee AS, et al. Drug screening using a library of human induced pluripotent stem cell-derived cardiomyocytes reveals disease specific patterns of cardiotoxicity. *Circulation* 2013;127(16):1677–91.
- [67] Mason C, Dunnill P. A brief definition of regenerative medicine. *Regen Med* 2008;3(1):1–5.
- [68] Vegas AJ, Veisoh O, Gürtler M, et al. Long-term glycemic control using polymer-encapsulated human stem cell-derived beta cells in immune-competent mice. *Nat Med* 2016;22(3):306–11.
- [69] Tolosa L, Caron J, Hannoun Z, et al. Transplantation of hESC-derived hepatocytes protects mice from liver injury. *Stem Cell Res Ther* 2015;6(1):246.
- [70] Bellamy V, Vanneaux V, Bel A, et al. Long-term functional benefits of human embryonic stem cell-derived cardiac progenitors embedded into a fibrin scaffold. *J Heart Lung Transplant* 2015;34(9):1198–207.
- [71] Sun J, Mandai M, Kamao H, et al. Protective effects of human iPSC-derived retinal pigmented epithelial cells in comparison with human mesenchymal stromal cells and human neural stem cells on the degenerating retina in rd1 mice. *Stem Cells* 2015;33(5):1543–53.
- [72] Lamba DA, Gust J, Reh TA. Transplantation of human embryonic stem cell-derived photoreceptors restores some visual function in *crx*-deficient mice. *Cell Stem Cell* 2009;4(1):73–9.
- [73] Kriks S, Shim J-W, Piao J, et al. Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. *Nature* 2011;480(7378):547–51.
- [74] Scott CT, Magnus D. Wrongful termination: lessons from the geron clinical trial. *Stem Cells Transl Med* 2014;3(12):1398–401.
- [75] Menasche P, Vanneaux V, Hagège A, et al. Human embryonic stem cell-derived cardiac progenitors for severe heart failure treatment: first clinical case report. *Eur Heart J* 2015;36(30):2011–7.
- [76] Schwartz SD, Regillo CD, Lam BL, et al. Human embryonic stem cell-derived retinal pigment epithelium in patients with age-related macular degeneration and Stargardt's macular dystrophy: follow-up of two open-label phase 1/2 studies. *Lancet* 2015;385(9967):509–16.
- [77] Garber K. RIKEN suspends first clinical trial involving induced pluripotent stem cells. *Nat Biotechnol* 2015;33(9):890–1.
- [78] Hamdi H, Planat-Benard V, Bel A, et al. Long-term functional benefits of epicardial patches as cell carriers. *Cell Transplant* 2014;23(1):87–96.
- [79] Bai Q, Desprat R, Klein B, Lemaitre J-M, De Vos J. Embryonic stem cells or induced pluripotent stem cells? A DNA integrity perspective. *Curr Gene Ther* 2013;13(2):93–8.
- [80] Rashid T, Kobayashi T, Nakauchi H. Revisiting the flight of Icarus: making human organs from PSCs with large animal chimeras. *Cell Stem Cell* 2014;15(4):406–9.
- [81] Normile D. Stem cell research. Chimeric embryos may soon get their day in the sun. *Science* 2013;340(6140):1509–10.
- [82] Chambers SM, Fasano CA, Papapetrou EP, Tomishima M, Sadelain M, Studer L. Highly efficient neural conversion of human ES and iPSC cells by dual inhibition of SMAD signaling. *Nat Biotechnol* 2009;27(3):275–80.
- [83] Kanemura H, Go MJ, Nishishita N, et al. Pigment epithelium-derived factor secreted from retinal pigment epithelium facilitates apoptotic cell death of iPSC. *Sci Rep* 2013;3:2334.
- [84] Shalom-Feuerstein R, Serron L, Aberdam E, et al. Impaired epithelial differentiation of induced pluripotent stem cells from ectodermal dysplasia-related patients is rescued by the small compound APR-246/PRIMA-1MET. *Proc Natl Acad Sci U S A* 2013;110(6):2152–6.
- [85] Nissan X, Larribere L, Saidani M, et al. Functional melanocytes derived from human pluripotent stem cells engraft into pluristratified epidermis. *Proc Natl Acad Sci U S A* 2011;108(36):14861–6.
- [86] Chal J, Oginuma M, Tanoury Al Z, et al. Differentiation of pluripotent stem cells to muscle fiber to model Duchenne muscular dystrophy. *Nat Biotechnol* 2015;33(9):962–9.
- [87] Smith MJ, Webber BR, Mohtashami M, Stefanski HE, Zúñiga-Pflücker JC, Blazar BR. In vitro T-cell generation from adult, embryonic, and induced pluripotent stem cells: many roads to one destination. *Stem Cells* 2015;33(11):3174–80.
- [88] Kobari L, Yates F, Oudrhiri N, et al. Human induced pluripotent stem cells can reach complete terminal maturation: in vivo and in vitro evidence in the erythropoietic differentiation model. *Haematologica* 2012;97(12):1795–803.
- [89] Orlova VV, van den Hil FE, Petrus-Reurer S, Drabsch Y, Dijke Ten P, Mummery CL. Generation, expansion and functional analysis of endothelial cells and pericytes derived from human pluripotent stem cells. *Nat Protoc* 2014;9(6):1514–31.
- [90] Umeda K, Oda H, Yan Q, et al. Long-term expandable SOX9+ chondrogenic ectomesenchymal cells from human pluripotent stem cells. *Stem Cell Reports* 2015;4(4):712–26.
- [91] Mohsen-Kanson T, Hafner A-L, Wdziekonski B, et al. Differentiation of human induced pluripotent stem cells into brown and white adipocytes: role of Pax3. *Stem Cells* 2014;32(6):1459–67.
- [92] Wong AP, Chin S, Xia S, Garner J, Bear CE, Rossant J. Efficient generation of functional CFTR-expressing airway epithelial cells from human pluripotent stem cells. *Nat Protoc* 2015;10(3):363–81.
- [93] Hannan NRF, Segeritz C-P, Touboul T, Vallier L. Production of hepatocyte-like cells from human pluripotent stem cells. *Nat Protoc* 2013;8(2):430–7.
- [94] Kurmann AA, Serra M, Hawkins F, et al. Regeneration of thyroid function by transplantation of differentiated pluripotent stem cells. *Cell Stem Cell* 2015;17(5):527–42.
- [95] Rezanian A, Bruin JE, Arora P, et al. Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. *Nat Biotechnol* 2014;32(11):1121–33.
- [96] Waddington CH. Principles of embryology. New York; 1956.

## 4. Techniques de culture

En culture, les cellules souches pluripotentes sont des cellules de petite taille (10-15  $\mu\text{m}$ ), adhérentes, formant des colonies compactes. Les CSP sont des cellules exigeantes pour leur culture. En effet, l'état de pluripotence est un équilibre fragile. Si les CSP entrent en différenciation, la pluripotence est perdue car le phénomène de différenciation est irréversible. Un soin quotidien doit donc être apporté aux cellules afin d'apporter les conditions optimales aux CSP pour maintenir leur état indifférencié.

C'est la combinaison délicate du milieu, de la matrice (ou du feeder) et des facteurs de croissance qui permet de maintenir l'état de pluripotence *in vitro*.

### 4.1 Les milieux de culture

Il existe de nombreux milieux de culture sur le marché pour maintenir l'état de pluripotence. Le milieu initialement décrit par James Thomson contenait du sérum de veau foetal (SVF) (Thomson et al., 1998). Afin de s'affranchir de la variabilité des lots de sérum, le SVF fut d'abord remplacé par un équivalent synthétique du SVF, le KnockOut Serum Replacement (KO-SR) (Amit et al., 2000; Koivisto et al., 2004). Plus tard, dans le but de s'affranchir des composants d'origine animale dans l'optique d'une culture de grade clinique, l'équipe de James Thomson a proposé un milieu défini et dépourvu de protéines d'origine animale, le milieu TSER (Ludwig et al., 2006). Plus récemment, ce milieu qui a fait l'objet d'une simplification, et est devenu le milieu E8, et est maintenant largement utilisé dans la communauté scientifique s'intéressant aux CSP (Chen et al., 2011). De nombreux autres milieux commerciaux sont également disponibles et une revue a été consacrée récemment aux modalités de culture des CSP dépourvues de protéines animales (Desai et al., 2015).



## 4.2 Les feeders et les matrices extracellulaires

Les CSP ne peuvent être cultivées sur plastique. Pour promouvoir la survie des CSP et le maintien de la pluripotence, la surface de culture doit être recouverte soit d'un tapis cellulaire (« feeder » ou cellules nourricières), soit d'une matrice protéique. Les CSP ont initialement été cultivées en système de co-culture avec des feeders. Historiquement, les premières matrices utilisées par Evans et Kauffman (Evans and Kaufman, 1981) étaient des fibroblastes d'embryons de souris, ou mouse embryonic fibroblast (MEF), irradiés ou mitomycinés pour empêcher leur prolifération. Pour éviter la transmission d'agents pathogènes animaux vers des cellules humaines, d'autres cellules nourricières d'origine humaine sont ont été utilisées, telles que des cellules musculaires adultes, des cellules de la peau fœtale, des cellules de la trompe utérine adulte, des cellules fibroblastiques placentaires, des cellules stromales de la moelle, des cellules de l'endomètre, des fibroblastes de prépuce (Hovatta et al., 2003). Ces derniers ont longtemps été utilisés par notre laboratoire.

Afin d'optimiser les cultures, il a fallu s'affranchir des cellules nourricières. Sont donc apparues des matrices protéiques plus ou moins complexes : le Matrigel<sup>®</sup>, le Geltrex<sup>®</sup>, la laminine, la vitronectine, etc. Le Matrigel<sup>®</sup> (Hughes et al., 2010) est un mélange complexe extrait de cellules de la tumeur murine Engelbreth-Holm-Swarm, un sarcome de souris contenant des constituants de lame basale, telles que la laminine et le collagène IV, ainsi que des facteurs de croissance. Le Geltrex<sup>®</sup> est produit similaire. Les laminines, glycoprotéines de la lame basale, ou la vitronectine, glycoprotéine extracellulaire, peuvent constituer, à l'état pur ou en combinaison, des matrices définies pour la culture des CSP (Braam et al., 2008; Rodin et al., 2014; 2010).

## 4.3 Les facteurs de croissance

Outre la surface et le milieu de culture utilisés pour maintenir l'état de pluripotence, certains facteurs de croissance sont également nécessaires.

Les cellules souches embryonnaires de souris et les CSP ne partagent pas les mêmes besoins en facteur de croissance. La pluripotence des ES murines peut-être maintenue par l'addition du Leukemia Inhibitory Factor (LIF). Par contraste, les cellules souches pluripotentes humaines sont indépendantes de cette cytokine mais bénéficient de la présence de basic Fibroblast Growth Factor (bFGF, FGF2) et d'Activine A ou de TGF $\beta$  pour le maintien de leur pluripotence (Greber et al., 2007) (voir ci-dessus le chapitre sur les facteurs extrinsèques de la pluripotence 1.3.5 et la Figure 11). Ces différences pourraient ne pas être des différences d'espèce, mais des différences de niveau de pluripotence (Dodsworth et al., 2015).

## 5. Limites des iPS

Les iPS présentent cependant un certain nombre de limites et d'inconvénients potentiels qu'il convient de connaître lors de leur manipulation.

### 5.1 Différences entre iPS et ES

Les cellules iPS sont-elles complètement pluripotentes ? Sont-elles l'équivalent des ES, ne différant que par leur origine ? En effet, des différences entre les cellules ES et iPS sont décrites dans la littérature. La génération de souris chimères par injection de cellules pluripotentes dans l'embryon au stade de blastocyste est plus difficile avec les iPS qu'avec les ES (Boland et al., 2009; Zhao et al., 2009). De plus, il existe des différences dans la signature transcriptomique et épigénomique des ES et des iPS, suggérant que les iPS pourraient conserver une part de l'épigénome de leur cellule d'origine (Chin et al., 2009; Ghosh et al., 2010; Kim et al., 2010; Marchetto et al., 2009; Ohi et al., 2011; Polo et al., 2010; Stadtfeld et al., 2010). Cependant, ces aberrations épigénétiques ne seraient pas

nécessairement dues à la reprogrammation (Tiemann et al., 2016). De plus, les différences observées sont moins nombreuses que les similarités. Enfin, ces différences ne pourraient être que la conséquence d'importantes variations du transcriptome et de l'épigénome entre les différentes lignées de cellules pluripotentes du fait des variations des génotypes, ou la conséquence du biais de la reprogrammation ou de culture cellulaire (Bai et al., 2015; Baker, 2012; Barrero and Izpisua Belmonte, 2011; Carey et al., 2011), voire tout simplement des différences inter-laboratoires (Newman and Cooper, 2010). ES et iPS, sont-elles équivalentes ? Des travaux récents montrent que les différences interlignées en raison des différences de génotype sont bien plus importantes que les différences séparant les iPS des ES, suggérant que ces deux types cellulaires sont très similaires (Choi et al., 2015; Kyttilä et al., 2016; Rouhani et al., 2014).

## **5.2 Difficultés de la mise en place des conditions de production de grade clinique**

L'autre limite importante à l'utilisation des iPS, et des CSP en général, est imposée par les règles de bonne conduite en clinique. Les trois étapes qui sont la sécurité, la qualité et l'efficacité des cellules issues d'hiPS dans le cadre de médecine régénératrice doivent, a fortiori, de par le caractère pluripotent des hiPS, être scrupuleusement respectées afin de ne pas engager la santé des patients en soignant leur pathologie. En effet, en l'absence de barrière immunologique, réintroduire ses propres cellules pluripotentes induites à un patient pourrait avoir des conséquences dramatiques, par exemple la formation de tératomes.

D'autre part, la réglementation exige que les cellules utilisées en thérapie doivent être de grade clinique, c'est à dire que la production des iPS et des cellules différenciées qui en sont issues respectent les bonnes pratiques de fabrication du médicament (BPF) (Seki and Fukuda, 2015). Ce qui soulève de nombreux points sensibles quant à l'utilisation des cellules dérivées d'hiPS. Quelle est la meilleure technique de reprogrammation à

utiliser ? Conserver l'intégrité génomique des cellules à reprogrammer tout en obtenant un rendement de reprogrammation correct, c'est le challenge que doivent accomplir les plateformes de reprogrammation des laboratoires pharmaceutiques. Et que faire de c-MYC ? Ce facteur de transcription entre dans la plupart des cocktails de reprogrammation efficaces mais c'est également un oncogène impliqué dans la physiopathologie de nombreux cancers. Quelles cellules reprogrammer ? Des fibroblastes issus d'une biopsie de peau (Nakagawa et al., 2008; Takahashi et al., 2007) ou de la muqueuse buccale (Miyoshi et al., 2010), des cellules de sang de cordon (Nishishita et al., 2011), des cellules du sang périphérique (Loh et al., 2009) ? Une approche peu invasive semble être la plus appropriée pour les applications cliniques. Comment se passer des protéines animales en culture ? Afin de supprimer le risque de transmission inter-espèces de zoonoses, les produits d'origine animale doivent être supprimés de toutes les étapes de reprogrammation, de culture, de différenciation. L'utilisation de fragments recombinants de protéines de la lame basale pourrait être une solution adaptée (Miyazaki et al., 2012). Les milieux de culture ne doivent pas non plus contenir de sérum animal, ce qui explique la multiplication de milieux de cultures définis et dépourvus de protéines animales (xeno-free) (Chen et al., 2011). Enfin, faut-il privilégier une utilisation d'iPS allogéniques ou autologues en clinique ? L'utilisation autologue iPS en médecine régénérative constituerait un grand pas dans vers l'affranchissement du problème de rejet immunologique et de la nécessité d'un traitement immunosuppresseur. Cependant, l'utilisation d'iPS autologues pose le problème d'un coût considérable. D'autre part, dans une démarche de correction d'une maladie génique, est-il préférable de modifier génétiquement les iPS issues de patients porteurs de maladie génétique et s'affranchir des incompatibilités immunologiques, ou d'utiliser des iPS allogéniques issues de donneurs sains pour éviter d'intervenir sur le génome des cellules qui seront injectées ?

Bien qu'un essai clinique ait déjà débuté au Japon (Reardon and Cyranoski, 2014), les problématiques évoquées ici ne sont pas encore tout à fait résolues. En particulier, des altérations génomiques des iPS utilisées dans l'essai ont imposé son arrêt.

### 5.3 L'apparition d'anomalies génomiques

Les CSP peuvent acquérir des anomalies génomiques au cours de leur culture. L'apparition de ces anomalies est en particulier liée à la technique de passage.

Le passage des CSP a lieu tous les 4-7 jours, lorsque les cellules s'approchent de la confluence. En effet, la confluence provoque la différenciation irréversible des CSP. Le passage consiste à détacher les CSP de leur support et de les remettre dans de nouveaux puits en présence d'un milieu et d'une matrice renouvelée. Il est possible de classer les nombreuses techniques de passage en deux grands groupes : le passage mécanique et le passage par dissociation cellulaire. Au cours du passage mécanique, les parties de colonies de CSP macroscopiquement pluripotentes sont découpées au scalpel ou à la pipette Pasteur effilée, afin de transférer les fragments dans une autre boîte de culture. Le passage par dissociation cellulaire fait intervenir une enzyme (trypsine, TrypLE, collagénase, etc.) ou un chélateur du calcium tel que l'EDTA pour séparer la colonie en cellules isolées (« single-cell ») avant de les transférer dans une nouvelle boîte de culture.

Notre équipe et d'autres ont constaté que les CSP accumulent plus facilement des anomalies génétiques en cas d'usage de passages enzymatiques à l'aide de trypsine ou de TrypLE (Bai et al., 2015; 2013; Mitalipova et al., 2005).

Notre équipe s'est également intéressée à l'analyse des modifications phénotypiques et transcriptionnelles des cellules au cours des passages enzymatiques et après acquisition d'anomalies génétiques par puces à ADN et séquençage des ARNm. Sur le plan du phénotype, nous avons remarqué que les CSP, qui présentent initialement une très grande sensibilité à la dissociation cellulaire – celle-ci entraînant une apoptose massive – s'adaptent rapidement au passage enzymatique et deviennent résistantes à l'induction de l'apoptose après dissociation cellulaire des colonies. Ces études ont mis également en évidence un profil transcriptionnel particulier après adaptation, et tout particulièrement l'induction massive d'une famille d'ARN longs non codants (long non coding RNA/lncRNA) que nous avons appelé ADAPTs. Ces lncRNA jouent un rôle fonctionnel dans l'adaptation des CSP à la culture en condition de passage enzymatique puisque leur répression bloque

l'adaptation. Ces résultats font l'objet d'un manuscrit en cours de rédaction (non présenté dans ce mémoire de thèse). Par ailleurs, nous avons rédigé une revue sur le rôle des lncRNA dans le développement embryonnaire humain, les CSP et la fécondation *in vitro*.

**Article 2** : « Long non coding RNA in human early embryonic development and in assisted reproductive technologies »

Accepté à *Hum Reprod Update*

1 **Long non-coding RNAs in human early embryonic**  
2 **development and in assisted reproductive technologies**

3

4 Julien Bouckenheimer<sup>1,2,3,\*</sup>, Said Assou<sup>1,2,3,\*</sup>, Sébastien Riquier<sup>1,2,3</sup>, Cyrielle  
5 Hou<sup>1,2,3</sup>, Nicolas Philippe<sup>1,2,3</sup>, Caroline Sansac<sup>1,2,3</sup>, Thierry Lavabre-Bertrand<sup>3</sup>,  
6 Thérèse Commes<sup>1,2,3,5</sup>, Jean-Marc Lemaître<sup>1,2,6,#</sup>, Anthony Boureux<sup>1,2,3</sup>, John  
7 De Vos<sup>1,2,3,5,6,7,#</sup>

8

9 <sup>1</sup> CHRU Montpellier, Institute for Regenerative Medicine and Biotherapy,  
10 Saint-Eloi Hospital, Montpellier, F34000 France; <sup>2</sup> INSERM, U1183,  
11 Montpellier, F34000 France; <sup>3</sup> Université de Montpellier, F34000 France; <sup>4</sup>  
12 Coretec, Montpellier; <sup>5</sup> Institut de Biologie Computationnelle, Montpellier,  
13 F34000 France; <sup>6</sup> Stem Cell Core Facility SAFE-iPSC, INGESTEM, Saint-Eloi,  
14 Hospital Montpellier, F34000 France; <sup>7</sup> CHU Montpellier, Department of Cell  
15 and Tissue Engineering, Saint-Eloi Hospital, Montpellier, F34000 France;

16 \* Contributed equally to this manuscript.

17

18 **Key words:** lncRNA, human embryo, pluripotency, human development, long  
19 non-coding RNA.

20

21 **Running title:** LncRNAs in early human embryo development

22

23

24 **# Corresponding authors:**

25 Jean-Marc Lemaître, Unit INSERM U1183, Hôpital Saint-Eloi, 80 Avenue

26 Augustin Fliche, 34295 Montpellier Cedex 5, France. Email: jean-

27 marc.lemaitre@inserm.fr

28 John De Vos, Department of Cell and Tissue Engineering, Hôpital Saint-Eloi,

29 80 Avenue Augustin Fliche, 34295 Montpellier Cedex 5, France. Email:

30 john.devos@inserm.fr

31

32



33 **ABSTRACT**

34 *Background:*

35 Human long non-coding RNAs (lncRNAs) are an emerging new category of  
36 transcripts with increasingly documented functional roles during development.  
37 lncRNAs and roles during human early embryo development have recently  
38 begun to be unravelled.

39 *Objective and rationale:*

40 This review recapitulates the most recent knowledge on lncRNAs and focuses  
41 on their expression patterns and role during early human embryo  
42 development and in pluripotent stem cells (PSC). Public mRNAseq data were  
43 used to illustrate these expression signatures.

44 *Search methods:*

45 The PubMed and EMBASE databases were first interrogated using specific  
46 terms, such as “lncRNAs” (or “long non-coding RNA”), to get an extensive  
47 overview on lncRNAs up to February 2016, and then using “human lncRNAs”  
48 AND “embryo”, “development”, or “pluripotent stem cells” to focus on lncRNAs  
49 involved in human embryo development or in PSC.

50 Recently published RNA-seq data from human oocytes and pre-implantation  
51 embryos (including single-cell data), PSC and a panel of normal and  
52 malignant adult tissues were used to describe the specific expression patterns  
53 of some lncRNAs in early human embryos.

54 *Outcomes:*

55 The existence and the crucial role of lncRNAs in many important biological  
56 phenomena in each branch of the life tree are now well documented. The  
57 number of identified lncRNAs is rapidly increasing and has already

58 outnumbered that of protein-coding genes. Differently from small non-coding  
59 RNAs, a variety of mechanisms of action have been proposed for lncRNAs.  
60 The functional role of lncRNAs has been demonstrated in many biological and  
61 developmental processes, including cell pluripotency induction, X inactivation  
62 or gene imprinting. Analysis of RNA-seq data highlights that lncRNA  
63 abundance changes significantly during human early embryonic development.  
64 This suggests that they could represent candidate biomarkers for developing  
65 non-invasive tests for oocyte or embryo quality. Finally, some of these  
66 lncRNAs are also expressed in human cancer tissues, suggesting that  
67 reactivation of an embryonic lncRNA program may contribute to human  
68 malignancies.

69 *Wider implications:*

70 lncRNAs are emerging potential key players in gene expression regulation.  
71 Analysis of RNA-seq data from human pre-implantation embryos identified  
72 lncRNA signatures that are specific to this critical step. We anticipate that  
73 further studies will show that these new transcripts are major regulators of  
74 embryo development. These findings might also be used to develop new  
75 tests/treatments for improving the pregnancy success rate in *in vitro*  
76 fertilization procedures or for regenerative medicine applications involving  
77 PSC.

## 78 **INTRODUCTION**

### 79 **Definition**

80 There is no official consensus concerning the sequence, structural, or  
81 biochemical features that might be used to clearly define long non-coding  
82 RNAs (lncRNAs). Nevertheless, transcripts that do not code for a protein and

83 are longer than 200 nucleotides are usually considered to be lncRNAs (Cabili  
84 et al., 2011; Derrien et al., 2012; Pang et al., 2006; Ulitsky et al., 2011). This  
85 size cut-off was arbitrarily chosen and is used to differentiate lncRNAs from  
86 smaller non-coding RNAs, such as microRNAs (miRNAs), small interfering  
87 RNAs (siRNAs), piwi-interacting RNAs (piRNAs), transfer RNAs (tRNAs) and  
88 small nucleolar RNAs (snoRNAs). Other non-coding, structural RNAs include  
89 ribosomal RNAs (rRNAs) and nuclear RNAs that are often longer than 200  
90 nucleotides, but that are not considered to be lncRNAs (Housman and Ulitsky,  
91 2016; Mattick and Rinn, 2015; Morris and Mattick, 2014; Taft et al., 2010).  
92 Moreover, compared with protein-coding genes, lncRNA genes are usually  
93 shorter in length (Derrien et al., 2012; Pauli et al., 2012; Ulitsky et al., 2011)  
94 and have fewer exons (Cabili et al., 2011; Derrien et al., 2012; Pauli et al.,  
95 2012; Ravasi, 2005; Zhu et al., 2009), although there are some exceptions.  
96 Finally, the assumption that lncRNAs are strictly non-coding transcripts has  
97 been recently challenged by several groups who reported that some lncRNAs  
98 may encode short peptides with important functional roles (Ruiz-Orera et al.,  
99 2014).

## 100 **A brief history of the discovery of lncRNAs**

101 The discovery of lncRNAs challenged one of the major biology dogmas.  
102 Indeed, in the common description of the human genome landscape, the ~21  
103 500 protein-coding genes (as indexed in Uniprot (Apweiler, 2004)) are  
104 drowned in a vast ocean of repetitive and non-transcribed regions (Lander et  
105 al., 2001). However, many tissue and cell transcriptome analyses brought  
106 evidences that in our genome also some non-coding DNA regions are  
107 transcribed into non-coding RNAs that may be functional. Consequently, a

108 more complex definition of what is a gene is needed, because a seemingly  
109 silent zone of the genome can be transcribed, leading to the regulated  
110 production of non-coding RNAs, such as small regulatory RNAs or functional  
111 lncRNAs. Examples of functional lncRNAs are rapidly accumulating, proving  
112 their role in many mechanisms involved in cell regulation, signalling pathways,  
113 or cell identity in specific tissues or cell types.

114 Historically, the human genome has been roughly separated in coding and  
115 non-coding regions and only protein-coding genes and some structural non-  
116 coding RNAs, such as tRNAs and rRNAs, have been considered as essential  
117 for the cell functions. Most of the non-coding genome was thought to be junk  
118 DNA, without any noticeable function. RNA biology was dominated by the  
119 concept that gene transcription and translation are the main regulatory  
120 processes, thus positioning proteins as the key functional regulators (Morris  
121 and Mattick, 2014). However, the roles ascribed to RNA evolved with the  
122 technological development. Messenger RNAs (mRNA) were discovered in the  
123 1960s by Jacob, Meselson and Brenner. They showed that mRNAs are  
124 translated into proteins and have key functions within cells (Brenner et al.,  
125 1961). In the 1970s, RNA expression in a specific cell or tissue could be  
126 studied only by Northern blot analysis, a complex and time-consuming  
127 procedure (Alwine et al., 1977). The development of cloning methods and  
128 particularly of the Polymerase Chain Reaction (PCR) by Mullis in the 1980s  
129 allowed a better characterization of RNAs and their expression levels within  
130 specific tissues and cells (Mullis, 1990). Meanwhile, other structural RNA  
131 classes were discovered, such as rRNAs, tRNAs and heteronuclear RNAs  
132 (hnRNAs) (Kim et al., 1974; Warner et al., 1966). In the 1990s, the first cDNA

133 arrays were developed, followed by tiling microarrays that can assess the  
134 expression of thousands of genes in a single experiment and even study gene  
135 expression genome-wide, outside protein coding regions (Schena et al.,  
136 1995). Microarrays revolutionized gene expression studies. For the first time,  
137 researchers did not have to know precisely what to look for to investigate  
138 gene expression in specific conditions, diseases, or drug effects. These  
139 techniques were greatly informative to understand cell functionality, but  
140 initially they did not allow the detection of splicing variants, and most often  
141 relied on known protein-coding genes. While the international human genome  
142 sequencing programme was close to obtain the first complete sequence,  
143 researchers were confronted with an unexpected difficulty. Indeed, they did  
144 not know how to define a gene or the gene portion that is in fine translated. It  
145 was thought that RNA maps could help finding an answer to this issue.  
146 Improved sequencing technologies solved this problem. Several international  
147 research programs, such as the Functional Annotation of the Mammalian  
148 Genome (FANTOM) project, using tiling microarray (Bertone et al., 2004) and  
149 sequencing techniques (Carninci et al., 2005) showed that most of the  
150 mammalian genome is actively transcribed into RNAs, although most of them  
151 do not code for proteins. Thousands of non-coding RNAs, including short  
152 (<200nt) and long (>200nt) RNAs, were identified. Furthermore, these  
153 projects showed that alternative splicing is very common, including in non-  
154 coding sections of the genome. These outcomes forced to widen the definition  
155 of a gene and to consider that non-coding RNAs may not simply be junk DNA  
156 or transcriptional noise (ENCODE Project Consortium et al., 2007; Gerstein et  
157 al., 2007; Kapranov et al., 2007). The advent of Next Generation Sequencing

158 (NGS) and the increasing accessibility to this technology because of falling  
159 costs enabled deep sequencing of the transcriptome (RNA-seq) of many  
160 cell/tissue types. Its rapid progress, as well as that of bioinformatics since the  
161 late 2000s paved the way for truly understanding transcription on a genome-  
162 wide scale and in different conditions. Particularly, NGS offered the possibility  
163 to compare the expression pattern of coding and non-coding RNAs (Dinger et  
164 al., 2008a; Mortazavi et al., 2008; Wang et al., 2009).

165 All these techniques helped describing and classifying non-coding RNAs in  
166 three subfamilies: (1) small structural RNAs that include small nuclear RNAs  
167 involved in splicing regulation, small nucleolar RNAs involved in ribosome  
168 biogenesis, rRNAs and tRNAs involved in translation, (2) small regulatory  
169 RNAs, such as miRNAs, piRNAs, snoRNAs that are involved in cell  
170 regulation, and (3) lncRNAs (Bartel, 2009; Carninci et al., 2005; Mattick, 2006;  
171 Rinn and Chang, 2012; Seto et al., 2007). This class includes non-coding  
172 RNAs that present similarities with mRNAs in length, structure and  
173 transcriptional regulation. Indeed, many functional non-coding transcripts  
174 were studied before being designed as lncRNAs. They were simply  
175 considered to be large functional RNAs with no coding function (Nagano et  
176 al., 2008; Pandey et al., 2008). The first described functional lncRNA was  
177 H19, discovered long before NGS advent and the first description of lncRNAs  
178 (Brannan et al., 1990). H19 is abundantly expressed in many mesoderm- and  
179 endoderm-derived tissues and regulates a network of co-expressed imprinted  
180 genes that modulate foetal and post-natal growth (Gabory et al., 2010).  
181 Indeed, H19 expression is regulated by genomic imprinting and is therefore  
182 expressed only by the maternal allele. Another lncRNA identified before their

183 discovery is XIST, which is transcribed only from the inactive X chromosome  
184 in mammalian female cells (Brown et al., 1991; Penny et al., 1996). XIST is  
185 essential for X inactivation and is pivotal for mammalian dosage  
186 compensation of the X chromosome (see below).

187 Here, we review the current knowledge on lncRNAs with a focus on their role  
188 in early human embryo development. Recent published mRNA sequencing  
189 (mRNAseq) data indicate that in the early human embryo, like in other  
190 vertebrate embryos, lncRNA abundance significantly changes during each  
191 stage of development, suggesting key functional roles. Therefore, lncRNAs  
192 could represent novel candidate markers for the development of non-invasive  
193 tools to assess embryo competency.

## 194 **METHODS**

### 195 **Article selection**

196 A summary of the general knowledge on lncRNAs was compiled based on  
197 seminal articles on this topic. A systematic review of the current literature in  
198 the English language on lncRNAs in mammalian, including human early  
199 embryo development was performed. Articles were retrieved from journal  
200 databases, such as PubMed (<http://www.ncbi.nlm.nih.gov/sites/entrez>) and  
201 EMBASE (<https://www.elsevier.com/solutions/embase-biomedical-research>)  
202 up to February 2016, using the keywords 'long non-coding RNA' or 'lncRNA',  
203 and 'embryo', 'development', or 'pluripotent stem cells'. 2845 articles were  
204 retrieved of which 211 were included in this review.

### 205 **RNA-Sequencing data retrieval**

206 Paired- or single-end RNA-seq data were downloaded from the NCBI SRA  
207 database (130 datasets corresponding to 5 SRA projects)

208 (<http://www.ncbi.nlm.nih.gov/sra>), or from the ENCODE project portal (36  
209 datasets from 18 ENCODE projects) (<https://www.encodeproject.org/>). The  
210 maximum read length was 121nt for the entire cohort and the minimum read  
211 length was 90nt, but for 8 RNA-seq datasets from the PRJNA189204 human  
212 embryo Bioproject that had a minimum read length of 49nt (Xue et al., 2013).  
213 The dataset list is provided as supplemental data (Table S1).

## 214 **RNA-Sequencing Data Analysis**

215 Reads were mapped to the indexed Ensembl GRCh38 human genome using  
216 CRAC v2.5.0 (<http://crac.gforge.inria.fr/>) (Philippe et al., 2013) with a kmer  
217 parameter of 22nt (-k 22) and the -bam option. Aligned reads were then  
218 counted against the Ensembl annotation file GRCh38.83 (Ensembl API  
219 version 66, <http://www.ensembl.org/index.html>) using featureCounts v14.6  
220 with the default unstranded option (-s 0). RNA-seq data analysis focused on  
221 the following Ensembl annotation categories: i) coding transcripts that  
222 included all protein-coding genes and ii) long non-coding transcripts that  
223 regrouped pseudogenes, sense and antisense overlapping transcripts,  
224 lncRNAs and lincRNAs. Small non-coding RNAs, such as microRNAs  
225 (miRNAs), were removed from the analysis, as well as ribosomal RNAs  
226 (rRNAs), which are not classified as lncRNAs.

## 227 **RESULTS AND DISCUSSION**

### 228 **LncRNA classification**

229 Non-coding RNAs can be arbitrarily separated in two main classes according  
230 to their length: small (< 200nt) and long (> 200nt, i.e. lncRNAs) (St. Laurent et  
231 al., 2015; Tuck and Tollervey, 2013). LncRNAs can then be further classified  
232 in two main subclasses according to their genome localization: intergenic and



233 intragenic (Fig. 1A). Currently, these lncRNA subclasses are not linked to a  
234 specific function or a mode of action. Moreover, no general mechanism of  
235 lncRNA biogenesis has been described so far, differently from small non-  
236 coding RNAs, such as piRNAs or miRNAs (Kim et al., 2009).

237

### 238 *Intragenic lncRNAs*

239 This lncRNA subclass can be further subdivided in several categories  
240 depending on the lncRNA position relative to the associated protein-coding  
241 gene (Marques et al., 2013):

242 a) Sense lncRNAs overlap with protein-coding regions and are transcribed  
243 from the same strand. Most of them do not encode proteins; however, several  
244 lncRNAs have both coding and non-coding potential. For instance, the  
245 intragenic lncRNA steroid receptor RNA activator (SRA) was among the first  
246 lncRNAs found to encode a small peptide (Chooniedass-Kothari et al., 2004;  
247 Denoeud et al., 2007; Djebali et al., 2008; Kapranov et al., 2005).

248 b) Antisense lncRNAs, usually called Natural Antisense Transcripts (NAT),  
249 also overlap with protein-coding regions, but are transcribed from the opposite  
250 strand (Schwartz et al., 2008).

251 Sense and antisense lncRNAs can overlap with exonic or intronic regions or  
252 both. NATs exist in virtually every living species, including eukaryotes,  
253 prokaryotes and even viruses (Britto-Kido et al., 2013; Campbell et al., 2014).  
254 However, no evolutionarily conserved NAT has been indexed (He et al., 2008;  
255 Katayama et al., 2005; Wahlestedt, 2013; Wood et al., 2013).

256 c) Bidirectional lncRNAs in which the transcription start site is located at less  
257 than 1kb away from its neighbouring protein-coding gene. They are usually  
258 transcribed in the opposite direction (Knauss and Sun, 2013).

259 d) Intronic lncRNAs are the predominant non-coding RNAs in the human  
260 genome and include totally intronic RNAs (TINs), which represent more than  
261 70% of all non-coding RNAs (except rRNAs) in human cells (Nakaya et al.,  
262 2007), and partially intronic RNAs (PINs), which are generally unspliced  
263 (Louro et al., 2009; St Laurent et al., 2012).

264

265 *Intergenic lncRNAs:*

266 Long intergenic (or intervening) lncRNAs, called lincRNAs, do not overlap with  
267 protein-coding regions (introns or exons). Therefore, their expression profile  
268 and the effects of their overexpression or down-regulation are simpler to  
269 interpret (Hacisuleyman et al., 2014; Hangauer et al., 2013). Many functional  
270 lincRNAs display mRNA-like stability (Clark et al., 2012) and active  
271 transcription marks, such as the K4H36 chromatin signature at their promoter  
272 when expressed (Guttman et al., 2009), and cell-specific expression profiles  
273 (Jia et al., 2010; Khalil et al., 2009).

274 LincRNAs are not clearly evolutionarily conserved; however, a recent study  
275 identified more than 2500 lincRNAs that are expressed in three tetrapod  
276 species and originated 90 million years ago (Necsulea et al., 2014). Most  
277 lincRNAs overlap with repetitive elements, such as SINEs or LINEs (Kelley  
278 and Rinn, 2012; Ulitsky et al., 2011), that seem to play significant roles within  
279 lincRNAs, notably by helping homologous base-pairing with other RNAs  
280 (Carrieri et al., 2012; Gong and Maquat, 2011).

281 This subclass includes also very long intergenic non-coding RNAs that are  
282 called vlincRNAs or macroRNAs. Their length can vary from 50kb to 1Mb and  
283 they can also derive from pseudogenes (Hackermüller et al., 2014; Lazorthes  
284 et al., 2015; Milligan and Lipovich, 2014; St Laurent et al., 2013).

285

### 286 **Mechanisms of action**

287 Many lncRNAs are non-functional by-products of pervasive transcription (De  
288 Santa et al., 2010; Ebisuya et al., 2008; Struhl, 2007). However, an increasing  
289 number of studies demonstrated that some lncRNAs have specific functions  
290 and that they can be major actors in the biology of a cell (Huarte and Marín-  
291 Béjar, 2015; Khorkova et al., 2015).

292 lncRNA mechanisms of action are numerous due to their diverse structural  
293 conformations, biochemical properties and specific subcellular localization  
294 (Wang and Chang, 2011). Most lncRNAs with documented functions have  
295 been found in the nucleus, but evidence of functional cytoplasmic lncRNAs is  
296 increasing. Nuclear lncRNAs are involved in many processes, including  
297 epigenetic, transcriptional, post-transcriptional and translational regulation  
298 mechanisms (Guttman and Rinn, 2012; Rinn and Chang, 2012; Ulitsky and  
299 Bartel, 2013). On the other hand, cytoplasmic lncRNA functions are restricted  
300 essentially to post-transcriptional and translational regulations and they act  
301 through the modulation of mRNA stability and protein modifications (Rashid et  
302 al., 2016). Their functions can be divided in five main categories (not mutually  
303 exclusive) that will be briefly described below and that are summarized in Fig.  
304 1B. Examples illustrating these different mechanisms of action are given in  
305 Table 1.

306

307 *Molecular scaffolds*

308 Most lncRNAs interact with DNA, RNA, proteins, or a combination of these  
309 molecules. lncRNAs tertiary structure leads to specific allosteric  
310 conformations, forming functional domains with high affinity for their targets  
311 (Guttman and Rinn, 2012). These molecular scaffolds can bind to several cell  
312 components to form ribonucleoprotein complexes and guide their interaction  
313 partners towards specific genomic loci thanks to their RNA sequence  
314 (Hacisuleyman et al., 2014; Quinn et al., 2014; Tsai et al., 2010). This type of  
315 mechanism highlights the importance of lncRNA structure and sequence for  
316 the formation, stability and localization of these complexes (He et al., 2015;  
317 Johnsson et al., 2014; Rouskin et al., 2013; Smith et al., 2013; Wan et al.,  
318 2014).

319 Such scaffolding generally takes place in the nucleus, where lncRNAs form  
320 protein complexes that modulate transcription. Well-known examples of  
321 proteins interacting with scaffolding lncRNAs include the chromatin modifiers  
322 PRC1 (Bonasio et al., 2014; Yap et al., 2010), PRC2 and LSD1 (Rinn et al.,  
323 2007), the Trithorax Group (TrxG) of proteins (Rinn et al., 2007), the DNMT  
324 protein family (Di Ruscio et al., 2013; Martianov et al., 2007; Schmitz et al.,  
325 2010), transcription factors (such as CBP, NPAS4, CREB and SRF),  
326 ribosomal complexes and also RNA polymerase II (Chisholm et al., 2012;  
327 West et al., 2014). lncRNAs can recruit these proteins and directly bring them  
328 to gene regulatory regions, for instance promoters to activate transcription.  
329 They can also act as structural platforms to which several proteins can bind  
330 and form specific structural entities (Sunwoo et al., 2008), such as nuclear

331 speckles (MALAT1) (Tripathi et al., 2010), paraspeckles (NEAT1/MEN $\beta$ )  
332 (Clemson et al., 2009; Murthy and Rangarajan, 2010; Souquere et al., 2010)  
333 and polycomb bodies (TUG1) (Yang et al., 2011b).

334

### 335 *Guides*

336 The primary nucleotide sequence of lncRNAs recognizes specific genomic  
337 regions through direct DNA binding, while their secondary and tertiary  
338 structures allow the binding of cellular components. Therefore, they can guide  
339 specific proteins or assembled scaffolds to a specific target locus that these  
340 molecules could not reach on their own. Well-described examples are  
341 HOTAIR (Schorderet and Duboule, 2011), AIR (Nagano et al., 2008),  
342 KCNQ1OT1 (Pandey et al., 2008), lincRNA-p21 (Huarte et al., 2010), HOTTIP  
343 (Wang et al., 2011) and XIST (Penny et al., 1996).

344

### 345 *Enhancers*

346 Enhancer lncRNAs (eRNAs) operate on neighbouring genes (*in cis*) or over  
347 long distances (*in trans*) by binding to promoters and inducing their  
348 transcription activity (Darrow and Chadwick, 2013; Fanucchi et al., 2013;  
349 Kaikkonen et al., 2013; Mousavi et al., 2013; Ørom et al., 2010). Many  
350 examples of eRNAs can be found in the literature (De Santa et al., 2010; Han  
351 et al., 2007; Kim et al., 2010; Lai et al., 2013; Lam et al., 2013; Léveillé et al.,  
352 2015; Melo et al., 2013; Ørom and Shiekhattar, 2013; Taft et al., 2009).

353 Most of lncRNAs have an enhancer-like function *in cis*, as indicated by the  
354 finding that their depletion leads to decreased expression of neighbouring  
355 protein-coding genes (Ørom and Shiekhattar, 2013). For instance, it has been

356 shown that a lncRNA interacts with the Mediator transcriptional activator  
357 complex to regulate its chromatin localization and to increase its enzymatic  
358 activity (Lai et al., 2013). Another eRNA facilitates the transition of paused  
359 RNA polymerase II into productive elongation by sequestering NELF, a  
360 negative elongation factor (Schaukowitch et al., 2014).

361 Among the lncRNAs that act *in trans*, CSR is engaged in long-range RNA  
362 interactions that facilitate the chromosomal looping-mediated association with  
363 super-enhancer regions (Pefanis et al., 2015). HOTTIP is another long-range  
364 eRNA that drives HOX genes transcription by getting in proximity of its target  
365 genes via chromosomal looping (Wang et al., 2011). eRNAs are also involved  
366 in cancer progression, particularly in acute leukaemia (Trimarchi et al., 2014)  
367 and in colorectal cancer (McClelland et al., 2016).

368

#### 369 *Decoys*

370 Another way of regulating gene expression by lncRNAs is endogenous  
371 competition that takes place preferentially in the cytoplasm (Salmena et al.,  
372 2011; Tay et al., 2014). Specifically, such lncRNAs act as molecular decoys  
373 by competitive binding to other cellular components, such as miRNAs, RNAs  
374 or proteins (transcription factors) in order to prevent their activity at their  
375 targeted action sites (Hung et al., 2011; Johnsson et al., 2014; Wang et al.,  
376 2013; Willingham et al., 2005). These lncRNAs are called competing  
377 endogenous RNAs (ceRNAs). They can act as artificial miRNA sponges that  
378 pair with and sequester miRNAs from their mRNA targets (Ebert et al., 2007),  
379 as observed in plants and viruses (Cazalla et al., 2010; Franco-Zorrilla et al.,  
380 2007). Of note, some of these miRNA sponges form remarkable structures

381 termed circular RNAs (ENCODE Project Consortium et al., 2007; Guo et al.,  
382 2014; Hansen et al., 2013; Memczak et al., 2013). The cytoplasmic lncRNA  
383 MD1 inhibits two important miRNAs (miR-133 and 135) and thus regulates  
384 myogenic progenitor cells during muscle differentiation (Cesana et al., 2011).  
385 This competition mechanism is also involved in cancer (Karreth et al., 2011).  
386 H19 is also found in the cytoplasm where it antagonizes let-7 microRNA  
387 subfamilies (Kallen et al., 2013) and is also a primary miRNAs precursor that  
388 negatively regulates the expression of specific mRNAs (Gao et al., 2012;  
389 Keniry et al., 2012). Similarly, the ceRNA GAS5 targets the glucocorticoid  
390 receptor and prevents its binding to regulatory elements (Kino et al., 2010).  
391 PANDA is another cytoplasmic ceRNA that interacts with the transcription  
392 factor NF- $\kappa$ B to limit the expression of pro-apoptotic genes (Hung et al., 2011).  
393 At the post-transcriptional level, lncRNAs with snoRNA ends interact with the  
394 FOX2 transcription factor and alter splicing patterns (Yin et al., 2012),  
395 whereas the UV-induced lncRNA GADD7 binds to the transcriptional  
396 repressor TDP-43, leading to the degradation of CDK6 mRNA in the  
397 cytoplasm (Liu et al., 2012). NEAT1 can titrate specific mRNAs by  
398 sequestering them within nuclear paraspeckles from where they are not  
399 released for translation until NEAT1 is degraded or down-regulated (Chen and  
400 Carmichael, 2009).

401 This titration mechanism of miRNAs should be effective only in specific cases  
402 because lncRNA level is generally lower than that of miRNAs and their mRNA  
403 targets (Bosson et al., 2014; Denzler et al., 2014). Thus, lncRNA  
404 concentration within the cell has to exceed a certain level of expression to  
405 allow their role as miRNA sponges.

## 406 **Controversy about lncRNA functional relevance**

407 lncRNAs have also been called transcripts of unknown function (TUF)  
408 (Gingeras, 2007) or transcripts of unknown coding potential (TUCP) (Cabili et  
409 al., 2011) that are localized in transcriptionally active regions (Li et al., 2007).  
410 Indeed, there are different views concerning lncRNA function. Although some  
411 lncRNAs, such as H19 or XIST, are now accepted as bona fide actors in cell  
412 regulation, the fraction of functional lncRNAs is still the subject of heated  
413 discussions (Kowalczyk et al., 2012). One view is that most of the non-coding  
414 transcriptome results from spurious transcription and could even be just  
415 transcriptional noise (Clark et al., 2011; De Santa et al., 2010; Ebisuya et al.,  
416 2008; Ponjavic et al., 2007; Struhl, 2007). This view is supported by the fact  
417 that many of these transcripts are expressed at very low levels (less than a  
418 single copy per cell for some lncRNAs) and seem to be less conserved than  
419 protein-coding RNAs (Cabili et al., 2011; Ravasi, 2005; Schultes et al., 2005).  
420 However, pooling cells for transcriptome sequencing could dilute a potentially  
421 strong signal localized in only few cells within the overall population. Thus,  
422 some lncRNAs that show an overall low expression level could in fact be  
423 tremendously overexpressed in a limited number of cells. Alternatively,  
424 lncRNA expression could be limited to a specific time during development, or  
425 be transient to regulate a precise cell mechanism (Clark et al., 2015; Derrien  
426 et al., 2012; Palazzo and Gregory, 2014).

427 Similarly, the lack of lncRNA conservation remains controversial. The  
428 absence of sequence conservation might suggest that lncRNAs do not have a  
429 large functional role (Ponjavic et al., 2007); however, lncRNAs may in fact be  
430 structurally conserved (Johnsson et al., 2014). Analyses on the structural



431 conservation of lncRNAs clearly support the hypothesis of an evolutionarily  
432 driven process (Guttman et al., 2009, 2010; Nitsche et al., 2015; Rands et al.,  
433 2014; Smith et al., 2013). Moreover, the overall number of lncRNAs increases  
434 with the complexity of the organism (Taft et al., 2007), an observation that  
435 could support lncRNA functionality (Dinger et al., 2008b; Guttman et al., 2011;  
436 Mattick, 2004). Indeed, the list of annotated lncRNAs with proven function is  
437 steadily increasing, with roles in pathways as diverse as cell regulation,  
438 apoptosis, drug resistance, cell identity specification, pluripotency and many  
439 other processes (Altschul et al., 1990; Huarte and Marín-Béjar, 2015;  
440 Khorkova et al., 2015; Pauli et al., 2011; Ponting et al., 2009; Rinn and  
441 Chang, 2012; Ulitsky and Bartel, 2013).

#### 442 **lncRNA expression in human pre-implantation embryos**

##### 443 *lncRNAs and early embryo development*

444 Cell lineage specification and differentiation during embryo development is  
445 achieved through the precise spatio-temporal activation of key genes in  
446 embryonic and progenitor cells (Ben-Tabou de-Leon and Davidson, 2007;  
447 Mallo and Alonso, 2013). Some lncRNAs associated with embryonic  
448 development have been recently identified by RNA-seq of animal and human  
449 pre-implantation embryos (from pronuclear to blastocyst stage). In frog  
450 (Paranjpe et al., 2013), mouse (Hamazaki et al., 2015; Zhang et al., 2014)  
451 and bovine early embryos (Caballero et al., 2014), maternal mRNAs peak in  
452 oocytes, whereas spliced lncRNAs are predominantly detected at the other  
453 early stages. Knockdown of some specific lncRNAs in bovine embryos affects  
454 embryo growth, resulting in increased developmental rates and larger  
455 blastocysts. In mice, siRNA-mediated silencing of a promoter-associated

456 lncRNA specifically expressed in 2-cell stage embryos results in embryonic  
457 lethality. This phenotype is rescued by overexpressing the protein that  
458 interacts with this lncRNA (Hamazaki et al., 2015). Importantly, all these  
459 studies highlight the huge number of lncRNAs expressed during embryo  
460 development.

461 Analysis of RNA-seq data of single cells from human pre-implantation  
462 embryos and of human embryonic stem cells (hESCs) (Yan et al., 2013)  
463 showed the expression of 3405 lncRNAs of which 2733 are potential novel  
464 intergenic non-coding RNAs. Interestingly, the expression of many of these  
465 lncRNAs is developmental-stage specific. Two other studies provided  
466 developmental-stage specific RNA-Seq data in human pre-implantation  
467 embryos (Blakeley et al., 2015; Xue et al., 2013), but did not specifically  
468 analyse the lncRNA profiles during embryo development.

#### 469 *A strong lncRNA signature in pre-implantation human embryos*

470 To illustrate the lncRNA expression profile of the first stages of human embryo  
471 development, we performed a meta-analysis of RNA-seq expression data by  
472 assembling the expression profiles of the two studies that included all  
473 embryonic stages during the first week of development (Xue et al., 2013; Yan  
474 et al., 2013) including six human mature oocytes, three pronuclear embryos,  
475 five zygotes, eight 2-cell stage embryos, twenty 4-cell stage embryos, thirty-  
476 one 8-cell stage embryos, nineteen morulae and thirty late blastocysts, as well  
477 as embryonic stem cells, foetal and adult normal tissues, and cancer samples  
478 (Supplemental Table S1). The analysis of the proportion of lncRNA and  
479 protein-coding mRNA reads shows that lncRNAs are present from the  
480 beginning of human embryo development and become the predominant

481 category of transcripts after embryonic gene activation (EGA) that occurs at  
482 the 4-cell to 8-cell stage (Fig. 2A).

483 Comparison with foetal RNA and adult tissues indicates that some of these  
484 lncRNAs are shared with foetal and adult tissues, while others are specific to  
485 early embryos (Fig. 2B). On the other hand, some foetal/adult tissues display  
486 lncRNAs that are not present in the oocyte and early embryo samples. These  
487 findings indicate that early human embryos and differentiated cells  
488 (foetal/adult tissues) have quite distinct lncRNA profiles.

489 As some lncRNAs are detected at specific and crucial steps of early  
490 embryonic development, they could represent candidate markers of EGA and  
491 embryo competency. Taken together, these results support the notion that  
492 lncRNAs are part of the dynamic changes in transcript expression that occur  
493 during human early embryo development and that contribute to the dramatic  
494 morphological changes during development.

#### 495 *Re-expression of pre-implantation embryo lncRNAs in cancer*

496 lncRNAs are involved in many essential physiological processes, from cell  
497 cycle control and apoptosis to epigenetic regulation. Therefore, deregulation  
498 of lncRNA expression could have dramatic consequences on cell  
499 homeostasis. Several lncRNAs exhibit oncogenic properties: HULC (Du et al.,  
500 2012) and HEIH in hepatocellular carcinoma (Yang et al., 2011a); HOTAIR in  
501 breast, gastric, colorectal and cervical cancer (Hajjari and Salavaty, 2015);  
502 MALAT1 in liver, pancreas, breast, colon and prostate carcinoma (Zhou et al.,  
503 2015); SRA in steroid-dependent tumours as breast and uterus cancer or  
504 ANRIL in prostate cancers (Hauptman and Glavač, 2013). Conversely, other  
505 lncRNAs have oncosuppressive features, such as MEG3 in pituitary adenoma

506 and meningioma, or lincRNA-p21, PTENP1 and GAS5 in prostate and breast  
507 cancer.

508 It has been proposed that some cancers arise through cell de-differentiation  
509 towards more immature cell states or to a state reminiscent of pluripotent  
510 stem cells (Ben-Porath et al., 2008; Schoenhals et al., 2009). Therefore, we  
511 investigated whether selected lncRNAs that are up-regulated in pre-  
512 implantation embryos but not in adult tissue are re-expressed in three cancer  
513 types (lung, colon and bladder) using publicly available RNA-seq data. Our  
514 analysis indicates that AFAP1-AS1 is present in lung cancer as well as in  
515 oocytes and up to the 8-cell embryo stage, but not in normal adult tissues  
516 (Fig. 2B). AFAP1-AS1 has been reported to be overexpressed in several  
517 cancers, such as hepatocellular carcinoma (Zhang et al., 2016), lung cancer  
518 (Zeng et al., 2016) and esophageal squamous cell carcinoma, even  
519 chemoresistant forms (Zhou et al., 2016), and it is often associated with poor  
520 prognosis. Similarly, the oncogene HOTAIR is overexpressed in colon and  
521 bladder cancer and is detected also at the 8-cell embryo stage and morula,  
522 but not at the blastocyst stage. AACSP1 is expressed in early embryos  
523 (oocyte to 8-cell stage) and then is reactivated in colon cancer. LUCAT1 is  
524 expressed at the blastocyst stage and upregulated in lung cancer compared  
525 with matched normal samples.

526

## 527 **LncRNAs and pluripotent stem cells**

### 528 *Pluripotency maintenance*

529 Pluripotency is the capacity of a cell to differentiate into any of the body cell  
530 types (De Los Angeles et al., 2015). For instance, hESCs can generate more

531 than 200 specialized cell types that derive from the three primary germ layers  
532 (ectoderm, endoderm and mesoderm) (Ramirez, 2010). Therefore,  
533 pluripotency maintenance and stem cell differentiation must be precisely  
534 regulated. Many transcriptions factors, such as OCT4 or NANOG, chromatin  
535 modifiers and small non-coding RNAs form a complex network for  
536 pluripotency regulation (De Los Angeles et al., 2015). Several recent studies  
537 showed that lncRNAs also are involved in this process (Guttman et al., 2011;  
538 Rosa and Brivanlou, 2013). Indeed, lncRNAs have been described as  
539 “regulators of regulators” or “pivots of pluripotency” that maintain pluripotency  
540 or act as a barrier against differentiation (Huo and Zambidis, 2013; Jia et al.,  
541 2013).

542 The first data that linked lncRNAs to ESC regulation were obtained in mice by  
543 microarray (Dinger et al., 2008b) and Chip-seq experiments (Guttman et al.,  
544 2009). They showed that mouse ESCs have specific lncRNA signatures that  
545 are lost during their differentiation. Guttman et al. discovered more than  
546 thousand new lncRNAs in mouse ESCs (Guttman et al., 2010). Many of these  
547 lncRNAs regulate and are regulated by pluripotency-associated transcription  
548 factors (Ng et al., 2012; Sheik Mohamed et al., 2010). Moreover, knock-down  
549 of pluripotency-associated lncRNAs causes profound transcriptome variations  
550 in mouse ESCs (Guttman et al., 2011), often resulting in loss of pluripotency  
551 and differentiation, as observed upon silencing of key pluripotency  
552 transcription factors (Sheik Mohamed et al., 2010).

553 Like many other lncRNAs, pluripotency-associated lncRNAs act as scaffolds  
554 to guide epigenetic regulatory complexes (Mondal and Kanduri, 2012) that are  
555 required for transcriptional regulation in ESCs. These complexes include

556 chromatin remodellers and histone modifiers, such as Polycomb and Tritorax  
557 complexes (Tsai et al., 2010). For instance, Tcl1 Upstream Neuron-  
558 Associated lincRNA (TUNA) forms a complex with ribonucleoproteins and the  
559 nucleolar phosphoprotein Nucleolin to guide them to the Nanog, Sox2 and  
560 Fgf4 promoters (Lin et al., 2014). TUNA depletion decreases the expression  
561 level of these pluripotency factors and inhibits mouse ESC self-renewal  
562 capacity and also differentiation into neural cells. Similarly, the lncRNAs ES1,  
563 ES2, ES3 are implicated in pluripotency and neural differentiation regulation in  
564 human cells (Ng et al., 2012). The lncRNAs GOMAFU (also called MIAT) and  
565 AK141205 are direct targets of OCT4 and NANOG, respectively, and their  
566 depletion in mouse ESCs leads to important transcriptomic changes and  
567 pluripotency perturbation (Sheik Mohamed et al., 2010).

568 Interestingly, many of these pluripotency-associated lncRNAs modulate the  
569 activity of epigenetic regulators. For instance, lncRNAs can regulate PRC2  
570 function directly (Dey et al., 2014; Kaneko et al., 2013; Zhao et al., 2010) or  
571 indirectly through interaction with related cofactors, as it is the case for the  
572 lncRNAs MEG3 or ANRIL (Kaneko et al., 2014; Shen et al., 2009; Yap et al.,  
573 2010). LncRNAs also interact with Tritorax-related factors, such as WDR5 that  
574 is very active in ESCs and regulates transcription by promoting H3K4me3  
575 deposition. Notably, lincRNA-1592 and lincRNA-1552 regulate and are  
576 regulated by pluripotency transcription factors in a positive feedback loop  
577 (Yang et al., 2014). Regulator of Reprogramming (RoR) is another important  
578 lncRNA for the maintenance of hESC self-renewal capacity. This lncRNA acts  
579 as a ceRNA in the cytoplasm where it maintains pluripotency by endogenous

580 competition (Wang et al., 2013) with miR-145, which normally inhibits  
581 translation of pluripotency transcription factors (Xu et al., 2009).

582

### 583 *Pluripotency induction (iPSCs)*

584 Induced Pluripotent Stem Cells (iPSCs) are pluripotent stem cells generated  
585 directly from adult somatic cells, such as fibroblasts, by overexpressing four  
586 specific transcription factors that reprogram adult cells into pluripotent stem  
587 cells (Takahashi et al., 2007). Loewer et al. found that some lncRNAs are up-  
588 regulated in iPSCs compared with ESCs (Loewer et al., 2010). However, it is  
589 not clear whether this difference is linked to intrinsic differences between  
590 iPSCs and ESCs, or to distinct features of their reprogramming protocol.  
591 Importantly, the lncRNA RoR is essential for cell reprogramming. Indeed, its  
592 depletion by shRNA-mediated knockdown leads to an important decrease of  
593 iPSC colonies, while its overexpression promotes iPSC colonies formation.  
594 Specifically, RoR depletion causes p53 up-regulation (Zhang et al., 2013),  
595 leading to induction of apoptosis and oxidative stress and ultimately to cell  
596 reprogramming defects. Pluripotency is regulated also by linc-p21, another  
597 lncRNA that represses p53 (Bao et al., 2015; Huarte et al., 2010). Differently  
598 from RoR, linc-p21 reduces reprogramming efficiency by acting on the  
599 conversion from pre-iPSCs to iPSCs that is required to establish bona fide  
600 pluripotent stem cells and during which the presence of the exogenous  
601 reprogramming factors is not needed any longer (Hanna et al., 2010). Linc-  
602 p21 forms a complex with the ribonucleoprotein Heterogeneous Nuclear  
603 Ribonucleoprotein K (HNRNPK) that maintains repressive H3K9me3 marks  
604 and CpG methylation at the promoter of core pluripotency transcription

605 factors. Interestingly, linc-p21 is also found in the cytoplasm where it acts as a  
606 post-transcriptional modulator by repressing the translation of beta catenin  
607 and JUNB mRNAs by imperfect base pairing, forming a complex that  
608 enhances the interaction between mRNAs and cytoplasmic translational  
609 repressors (Rashid et al., 2016).

610

### 611 **Regulation of genomic imprinting by lncRNAs**

612 Genomic imprinting is a crucial developmental epigenetic process that results  
613 in the mono-allelic expression of certain genes, in a parent-of-origin  
614 dependent manner (see Barlow and Bartolomei for a general review on this  
615 phenomenon (Barlow and Bartolomei, 2014). These genes are thus called  
616 imprinted genes. Currently, about 100 imprinted genes have been identified.  
617 Many of them are essential for embryo development and growth (especially  
618 during gestation and post-partum) and genetic mutations at imprinted regions  
619 can cause various pathologies, particularly growth disorders (Azzi et al., 2014)  
620 and cancers (Kim et al., 2015). Most imprinted genes are grouped in large  
621 clusters with coordinated expression (Abramowitz and Bartolomei, 2012). It  
622 has been hypothesized that the regulation of these clusters is directed by  
623 lncRNAs that operate locally *in cis* by interacting with chromatin modifiers, or  
624 *in trans* by controlling gene expression (Autuoro et al., 2014; Kanduri, 2016).  
625 Moreover, many imprinted genes share some common features with  
626 lncRNAs, for example their tissue and temporal specific expression,  
627 suggesting that they could be regulated by lncRNAs, or that they share similar  
628 regulating mechanisms in differentiated tissues (Prickett and Oakey, 2012).



629 KCNQ1OT1 is a lncRNA that is transcribed in the opposite direction of the  
630 protein-coding gene KCNQ1 (potassium voltage-gated channel, KQT-like  
631 subfamily, member 1). KCNQ1 can be silenced by KCNQ1OT1 (Kanduri,  
632 2011; Thakur et al., 2004). This lncRNA is expressed only from the paternal  
633 allele (Khoueiry et al., 2008) and contributes to the repression of transcription  
634 of at least four neighbouring genes via interaction with chromatin modifiers,  
635 such as the PRC2 complex (Magistri et al., 2012), and with histone-modifying  
636 enzymes, such as histone methyltransferases and deacetylases (Otte and  
637 Kwaks, 2003). Mouse Air (Sleutels et al., 2002) (or Airn for Antisense Igf2r  
638 RNA Noncoding) shares some characteristics with KCNQ1OT1. It is  
639 expressed only from the paternal allele and its expression silences  
640 neighbouring protein-coding genes present on the maternal Insulin-like growth  
641 factor 2 receptor (Igf2r) locus. Igf2r is expressed from the maternal  
642 chromosome (Barlow and Bartolomei, 2014), but can be repressed by Air by  
643 limiting RNA polymerase II access to the Igf2r promoter. An orthologue of Air,  
644 named AIR, was found in human cells, but its function is unknown because  
645 IGF2R is expressed by both paternal and maternal alleles in humans (Yotova  
646 et al., 2008). H19 is a maternally-expressed lncRNA (Brannan et al., 1990)  
647 that acts in conjunction with the protein-coding Insulin Growth Factor (IGF2)  
648 gene that is expressed from the paternal allele (Smits et al., 2008). H19 is  
649 involved in the regulation of foetal growth (Gabory et al., 2010) by limiting  
650 body weight increase, whereas IGF does the opposite. It also acts as an  
651 oncogene in various cancer types (Matouk et al., 2007) and is deregulated in  
652 many growth disorders (Wojdacz et al., 2008).

653 Mouse orthologues of these lncRNAs have been identified, suggesting an  
654 important functional and evolutionarily conserved role. Finally, it must be  
655 noted that deregulation of these lncRNAs leads to various forms of cancer  
656 (Higashimoto et al., 2006; Li et al., 2014) and are therefore considered as new  
657 druggable targets (Fatemi et al., 2014).

658

### 659 **X chromosome inactivation**

660 Human female cells possess two X chromosomes, whereas male cell possess  
661 one X and one Y. In female cells, the overexpression of X-linked genes is  
662 inhibited via a dosage compensation mechanism, called X chromosome  
663 inactivation (XCI) (Gartler et al., 1992). XCI leads to the random silencing of  
664 one of the two X copies (either maternal or paternal), after embryo  
665 implantation in the uterus wall. Once silenced, the same inactive X  
666 chromosome is transmitted to all daughter cells. XCI is a well-studied  
667 mechanism that helped highlighting the functional role of the mammalian  
668 lncRNA XIST (Brockdorff et al., 1992; Penny et al., 1996). A specific locus on  
669 the X chromosome contains several lncRNAs among which X-inactive specific  
670 transcript (XIST), which is a 19kb transcript, and XIST antisense transcript  
671 (TSIX), which is 40kb-long and negatively regulates XIST (Stavropoulos et al.,  
672 2001). The mechanism whereby XIST transcription is promoted at the X  
673 chromosome to be silenced (Xi) and TSIX transcription at the future active X  
674 chromosome (Xa) is unknown. XIST transcripts form a coat around the Xi and  
675 promote the recruitment of chromatin remodellers, such as PRC2 (Lee, 2012)  
676 that deposits inactive chromatin marks. Another recently identified lncRNA, X-  
677 Active Coating Transcript (XACT), might coat the Xa to protect it from

678 inactivation (Vallot et al., 2013). XACT is very long transcript (252kb) that is  
679 only expressed on the Xa and only in pluripotent stem cells.

680

### 681 **Regulation of developmental gene expression**

682 Beside early embryonic development, lncRNAs have also been involved in the  
683 subsequent steps of embryonic and foetal development, ranging from the  
684 regulation of HOX genes via the lncRNA HOTAIR (Rinn et al., 2007) to cell  
685 differentiation (Fatica and Bozzoni, 2014). For instance, modulation of the  
686 expression of brain-specific lncRNAs affects the differentiation capacity of  
687 neural progenitors. MALAT1 dysregulation influences synapse density and  
688 dendrite growth (Bernard et al., 2010) by controlling the localization of key  
689 neuronal transcription factors, while BDNF-AS controls neurite formation  
690 (Modarresi et al., 2012). Forced overexpression of the lncRNA NKX2.2AS  
691 (Tochitani and Hayashizaki, 2008) in neural stem cells induces differentiation  
692 of oligodendrocytes. LncRNAs are also involved in heart development. The  
693 mouse-specific lncRNA Braveheart (Bvht) controls differentiation of  
694 mesodermal precursors into cardiomyocytes by regulating the expression of  
695 mesoderm posterior basic helix-loop-helix transcription factor 1 (MesP1)  
696 through its interaction with the PRC2 histone modifier (Klattenhoff et al.,  
697 2013). Bvht is required for proper development of cardiomyocytes because its  
698 depletion affects the expression of cardiac-specific genes and prevents  
699 cardiomyocyte maturation. Table 2 lists other examples illustrating the major  
700 contribution of lncRNAs in gene expression regulation during mammalian  
701 development.

702

703 **Potential applicability of lncRNAs for non-invasive testing of**  
704 **oocyte/embryo quality**

705 Assisted reproductive technologies (ART) include the selection of the  
706 embryos with the best developmental potential before their transfer in the  
707 uterus. Currently, the method of choice to evaluate embryo quality is based on  
708 the microscopic observations of various morphological parameters (Ahlström  
709 et al., 2011; Lundin et al., 2001; Racowsky et al., 2010). However,  
710 morphological criteria might be insufficient for predicting embryo quality and  
711 pregnancy outcome (Assou et al., 2011; Mastenbroek et al., 2011) and  
712 adjunctive technologies for the assessment of embryo potential are currently  
713 investigated (Hillier, 2008). These new approaches include the identification of  
714 molecular biomarkers based on the analysis of the oocyte microenvironment  
715 (specifically cumulus cells) using sophisticated and high-performance  
716 technologies, such as quantitative RT-PCR, microarray and RNA sequencing.  
717 Several groups have proposed to use mRNAs expressed in cumulus cells, as  
718 a non-invasive biomarker to predict oocyte quality and embryo competence  
719 and/or pregnancy outcome (Anderson et al., 2009; Assidi et al., 2011; Assou  
720 et al., 2008, 2010; Fragouli et al., 2012; Gebhardt et al., 2011; McKenzie et  
721 al., 2004). Other groups investigated the role of miRNAs in cumulus cells and  
722 tried to link their expression profile in cumulus cells with oocyte maturation  
723 and polycystic ovarian syndrome (PCOS) (Assou et al., 2013; Huang et al.,  
724 2016a; Liu et al., 2015; Tong et al., 2014). Similarly, lncRNA expression  
725 profile might not only reflect the developmental stage of the embryo, but also  
726 that of granulosa and cumulus cells (Burnik Papler et al., 2015). Indeed,  
727 differential cumulus cell expression of specific lncRNAs was reported in

728 germinal vesicle (GV) and mature (MII) oocytes (Yerushalmi et al., 2014) and  
729 in pre-implantation embryos of high and poor quality (Xu et al., 2015).  
730 Moreover, lncRNAs are aberrantly expressed in cumulus cells of patients with  
731 polycystic ovary syndrome (PCOS) compared with women without PCOS,  
732 suggesting that lncRNA expression may influence oocyte competence (Huang  
733 et al., 2016b). Thus, like for mRNAs and miRNAs, the analysis of lncRNAs  
734 expressed in cumulus cells using quantitative RT-PCR might also be used for  
735 the development of a non-invasive test for the selection of the oocytes to be  
736 fertilized and/or of good-quality embryos for transfer during in vitro fertilization  
737 (IVF) procedures. Further studies are required to validate the suitability and  
738 value of cumulus cell lncRNA expression testing in ART. Moreover, as the  
739 expression patterns of lncRNAs in pre-implantation embryos suggest a  
740 potential role of lncRNAs during human early embryonic development,  
741 specific embryonic lncRNAs released in the culture medium after IVF are  
742 potentially attractive biomarkers for the prediction of embryo quality and  
743 pregnancy outcomes.

744

## 745 **CONCLUSION**

746 lncRNAs are a group of recently identified RNA molecules involved in many  
747 important biological processes in virtually all life forms. The total number of  
748 lncRNAs is still unknown, but is steadily increasing and has already  
749 surpassed that of protein-coding genes. They have been recognized as major  
750 regulators of mammalian tissue physiology and disease processes, including  
751 cancer. Moreover, lncRNAs are strongly expressed in the early human  
752 embryo, with a rapidly changing expression pattern, and some of them

753 contribute to critical aspects of early human development. These lines of  
754 evidence show that lncRNAs are still widely unexplored in early human  
755 development and strongly suggest that they may become a mean of  
756 intervention in ART.

757

## 758 **ACKNOWLEDGEMENT**

759 We thank E. Andermarcher for critical reading of this manuscript.

760

## 761 **AUTHORS' CONTRIBUTIONS**

762 JB, SA and JDV wrote the manuscript. JB, SA, SR, NP, CS, TLB, TC, JML  
763 and AB contributed to the meta-analysis. CH contributed to the tables. All  
764 authors read and approved the final manuscript.

765

## 766 **FUNDING STATEMENT**

767 Supported by grants from the Agence Nationale de la Recherche  
768 (INGESTEM), the association “Gueules cassées” and the University Hospital  
769 of Montpellier (AOI 2013-A00892-43).

770

## 771 **DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST**

772 The authors indicate no potential conflicts of interest.

773

## 774 **REFERENCES**

775 Abramowitz, L.K., and Bartolomei, M.S. (2012). Genomic imprinting:  
776 recognition and marking of imprinted loci. *Curr. Opin. Genet. Dev.* 22, 72–78.

777 Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990).  
778 Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.

779 Alwine, J.C., Kemp, D.J., and Stark, G.R. (1977). Method for detection of  
780 specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and  
781 hybridization with DNA probes. *Proc. Natl. Acad. Sci.* *74*, 5350–5354.

782 Anderson, R.A., Sciorio, R., Kinnell, H., Bayne, R. a. L., Thong, K.J., de  
783 Sousa, P.A., and Pickering, S. (2009). Cumulus gene expression as a  
784 predictor of human oocyte fertilisation, embryo development and competence  
785 to establish a pregnancy. *Reprod. Camb. Engl.* *138*, 629–637.

786 Apweiler, R. (2004). UniProt: the Universal Protein knowledgebase. *Nucleic  
787 Acids Res.* *32*, 115D – 119.

788 Assidi, M., Montag, M., Van der Ven, K., and Sirard, M.-A. (2011). Biomarkers  
789 of human oocyte developmental competence expressed in cumulus cells  
790 before ICSI: a preliminary study. *J. Assist. Reprod. Genet.* *28*, 173–188.

791 Assou, S., Haouzi, D., Mahmoud, K., Aouacheria, A., Guillemin, Y., Pantesco,  
792 V., Rème, T., Dechaud, H., De Vos, J., and Hamamah, S. (2008). A non-  
793 invasive test for assessing embryo potential by gene expression profiles of  
794 human cumulus cells: a proof of concept study. *Mol. Hum. Reprod.* *14*, 711–  
795 719.

796 Assou, S., Haouzi, D., De Vos, J., and Hamamah, S. (2010). Human cumulus  
797 cells as biomarkers for embryo and pregnancy outcomes. *Mol. Hum. Reprod.*  
798 *16*, 531–538.

799 Assou, S., Boumela, I., Haouzi, D., Anahory, T., Dechaud, H., De Vos, J., and  
800 Hamamah, S. (2011). Dynamic changes in gene expression during human  
801 early embryo development: from fundamental aspects to clinical applications.  
802 *Hum. Reprod. Update* *17*, 272–290.

803 Assou, S., Al-edani, T., Haouzi, D., Philippe, N., Lecellier, C.-H., Piquemal, D.,  
804 Commes, T., Aït-Ahmed, O., Dechaud, H., and Hamamah, S. (2013).  
805 MicroRNAs: new candidates for the regulation of the human cumulus-oocyte  
806 complex. *Hum. Reprod. Oxf. Engl.* *28*, 3038–3049.

807 Autuoro, J., Pirnie, S., and Carmichael, G. (2014). Long Noncoding RNAs in  
808 Imprinting and X Chromosome Inactivation. *Biomolecules* *4*, 76–100.

809 Azzi, S., Brioude, F., Le Bouc, Y., and Netchine, I. (2014). Human imprinting  
810 anomalies in fetal and childhood growth disorders: clinical implications and  
811 molecular mechanisms. *Curr. Pharm. Des.* *20*, 1751–1763.

812 Bai, Q., Assou, S., Haouzi, D., Ramirez, J.-M., Monzo, C., Becker, F., Gerbal-  
813 Chaloin, S., Hamamah, S., and De Vos, J. (2012). Dissecting the First  
814 Transcriptional Divergence During Human Embryonic Development. *Stem  
815 Cell Rev. Rep.* *8*, 150–162.

816 Bao, X., Wu, H., Zhu, X., Guo, X., Hutchins, A.P., Luo, Z., Song, H., Chen, Y.,  
817 Lai, K., Yin, M., et al. (2015). The p53-induced lincRNA-p21 derails somatic  
818 cell reprogramming by sustaining H3K9me3 and CpG methylation at  
819 pluripotency gene promoters. *Cell Res.* *25*, 80–92.

- 820 Barlow, D.P., and Bartolomei, M.S. (2014). Genomic imprinting in mammals.  
821 Cold Spring Harb. Perspect. Biol. 6.
- 822 Bartel, D.P. (2009). MicroRNAs: Target Recognition and Regulatory  
823 Functions. *Cell* 136, 215–233.
- 824 Ben-Porath, I., Thomson, M.W., Carey, V.J., Ge, R., Bell, G.W., Regev, A.,  
825 and Weinberg, R.A. (2008). An embryonic stem cell-like gene expression  
826 signature in poorly differentiated aggressive human tumors. *Nat. Genet.* 40,  
827 499–507.
- 828 Ben-Tabou de-Leon, S., and Davidson, E.H. (2007). Gene regulation: gene  
829 control network in development. *Annu. Rev. Biophys. Biomol. Struct.* 36, 191.
- 830 Bernard, D., Prasanth, K.V., Tripathi, V., Colasse, S., Nakamura, T., Xuan, Z.,  
831 Zhang, M.Q., Sedel, F., Jourden, L., Couplier, F., et al. (2010). A long  
832 nuclear-retained non-coding RNA regulates synaptogenesis by modulating  
833 gene expression. *EMBO J.* 29, 3082–3093.
- 834 Bertani, S., Sauer, S., Bolotin, E., and Sauer, F. (2011). The noncoding RNA  
835 *Mistral* activates *Hoxa6* and *Hoxa7* expression and stem cell differentiation by  
836 recruiting MLL1 to chromatin. *Mol. Cell* 43, 1040–1046.
- 837 Bertone, P., Stolc, V., Royce, T.E., Rozowsky, J.S., Urban, A.E., Zhu, X.,  
838 Rinn, J.L., Tongprasit, W., Samanta, M., Weissman, S., et al. (2004). Global  
839 identification of human transcribed sequences with genome tiling arrays.  
840 *Science* 306, 2242–2246.
- 841 Blakeley, P., Fogarty, N.M.E., del Valle, I., Wamaitha, S.E., Hu, T.X., Elder,  
842 K., Snell, P., Christie, L., Robson, P., and Niakan, K.K. (2015). Defining the  
843 three cell lineages of the human blastocyst by single-cell RNA-seq.  
844 *Development* 142, 3151–3165.
- 845 Bonasio, R., Lecona, E., Narendra, V., Voigt, P., Parisi, F., Kluger, Y., and  
846 Reinberg, D. (2014). Interactions with RNA direct the Polycomb group protein  
847 SCML2 to chromatin where it represses target genes. *Elife* 3, e02637.
- 848 Bosson, A.D., Zamudio, J.R., and Sharp, P.A. (2014). Endogenous miRNA  
849 and Target Concentrations Determine Susceptibility to Potential ceRNA  
850 Competition. *Mol. Cell* 56, 347–359.
- 851 Brannan, C.I., Dees, E.C., Ingram, R.S., and Tilghman, S.M. (1990). The  
852 product of the H19 gene may function as an RNA. *Mol. Cell. Biol.* 10, 28–36.
- 853 Brenner, S., Jacob, F., and Meselson, M. (1961). An unstable intermediate  
854 carrying information from genes to ribosomes for protein synthesis. *Nature*  
855 190, 576–581.
- 856 Britto-Kido, S. de A., Ferreira Neto, J.R.C., Pandolfi, V., Marcelino-  
857 Guimarães, F.C., Nepomuceno, A.L., Vilela Abdelnoor, R., Benko-Iseppon,  
858 A.M., and Kido, E.A. (2013). Natural Antisense Transcripts in Plants: A



- 859 Review and Identification in Soybean Infected with *Phakopsora pachyrhizi*  
860 SuperSAGE Library. *Sci. World J.* 2013, 1–14.
- 861 Brockdorff, N., Ashworth, A., Kay, G.F., McCabe, V.M., Norris, D.P., Cooper,  
862 P.J., Swift, S., and Rastan, S. (1992). The product of the mouse Xist gene is a  
863 15 kb inactive X-specific transcript containing no conserved ORF and located  
864 in the nucleus. *Cell* 71, 515–526.
- 865 Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M.,  
866 Tonlorenzi, R., and Willard, H.F. (1991). A gene from the region of the human  
867 X inactivation centre is expressed exclusively from the inactive X  
868 chromosome. *Nature* 349, 38–44.
- 869 Burnik Papler, T., Vrtacnik Bokal, E., Maver, A., Kopitar, A.N., and Lovrečić, L.  
870 (2015). Transcriptomic Analysis and Meta-Analysis of Human Granulosa and  
871 Cumulus Cells. *PloS One* 10, e0136473.
- 872 Caballero, J., Gilbert, I., Fournier, E., Gagné, D., Scantland, S., Macaulay, A.,  
873 and Robert, C. (2014). Exploring the function of long non-coding RNA in the  
874 development of bovine early embryos. *Reprod. Fertil. Dev.* 27, 40–52.
- 875 Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., and  
876 Rinn, J.L. (2011). Integrative annotation of human large intergenic noncoding  
877 RNAs reveals global properties and specific subclasses. *Genes Dev.* 25,  
878 1915–1927.
- 879 Campbell, M., Kim, K.Y., Chang, P.-C., Huerta, S., Shevchenko, B., Wang,  
880 D.-H., Izumiya, C., Kung, H.-J., and Izumiya, Y. (2014). A Lytic Viral Long  
881 Noncoding RNA Modulates the Function of a Latent Protein. *J. Virol.* 88,  
882 1843–1848.
- 883 Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N.,  
884 Oyama, R., Ravasi, T., Lenhard, B., Wells, C., et al. (2005). The  
885 transcriptional landscape of the mammalian genome. *Science* 309, 1559–  
886 1563.
- 887 Carrieri, C., Cimatti, L., Biagioli, M., Beugnet, A., Zucchelli, S., Fedele, S.,  
888 Pesce, E., Ferrer, I., Collavin, L., Santoro, C., et al. (2012). Long non-coding  
889 antisense RNA controls Uchl1 translation through an embedded SINEB2  
890 repeat. *Nature* 491, 454–457.
- 891 Cazalla, D., Yario, T., Steitz, J.A., and Steitz, J. (2010). Down-regulation of a  
892 host microRNA by a Herpesvirus saimiri noncoding RNA. *Science* 328, 1563–  
893 1566.
- 894 Cesana, M., Cacchiarelli, D., Legnini, I., Santini, T., Sthandier, O., Chinappi,  
895 M., Tramontano, A., and Bozzoni, I. (2011). A Long Noncoding RNA Controls  
896 Muscle Differentiation by Functioning as a Competing Endogenous RNA. *Cell*  
897 147, 358–369.
- 898 Chaumeil, J., Waters, P.D., Koina, E., Gilbert, C., Robinson, T.J., and Graves,  
899 J.A.M. (2011). Evolution from XIST-independent to XIST-controlled X-

900 chromosome inactivation: epigenetic modifications in distantly related  
901 mammals. *PLoS One* 6, e19040.

902 Chen, L.-L., and Carmichael, G.G. (2009). Altered Nuclear Retention of  
903 mRNAs Containing Inverted Repeats in Human Embryonic Stem Cells:  
904 Functional Role of a Nuclear Noncoding RNA. *Mol. Cell* 35, 467–478.

905 Chisholm, K.M., Wan, Y., Li, R., Montgomery, K.D., Chang, H.Y., and West,  
906 R.B. (2012). Detection of Long Non-Coding RNA in Archival Tissue:  
907 Correlation with Polycomb Protein Expression in Primary and Metastatic  
908 Breast Carcinoma. *PLoS ONE* 7, e47998.

909 Chooniedass-Kothari, S., Emberley, E., Hamedani, M.K., Troup, S., Wang, X.,  
910 Czosnek, A., Hube, F., Mutawe, M., Watson, P.H., and Leygue, E. (2004).  
911 The steroid receptor RNA activator is the first functional RNA encoding a  
912 protein. *FEBS Lett.* 566, 43–47.

913 Clark, M.B., Amaral, P.P., Schlesinger, F.J., Dinger, M.E., Taft, R.J., Rinn,  
914 J.L., Ponting, C.P., Stadler, P.F., Morris, K.V., Morillon, A., et al. (2011). The  
915 Reality of Pervasive Transcription. *PLoS Biol.* 9, e1000625.

916 Clark, M.B., Johnston, R.L., Inostroza-Ponta, M., Fox, A.H., Fortini, E.,  
917 Moscato, P., Dinger, M.E., and Mattick, J.S. (2012). Genome-wide analysis of  
918 long noncoding RNA stability. *Genome Res.* 22, 885–898.

919 Clark, M.B., Mercer, T.R., Bussotti, G., Leonardi, T., Haynes, K.R., Crawford,  
920 J., Brunck, M.E., Cao, K.-A.L., Thomas, G.P., Chen, W.Y., et al. (2015).  
921 Quantitative gene profiling of long noncoding RNAs with targeted RNA  
922 sequencing. *Nat. Methods* 12, 339–342.

923 Clemson, C.M., Hutchinson, J.N., Sara, S.A., Ensminger, A.W., Fox, A.H.,  
924 Chess, A., and Lawrence, J.B. (2009). An Architectural Role for a Nuclear  
925 Noncoding RNA: NEAT1 RNA Is Essential for the Structure of Paraspeckles.  
926 *Mol. Cell* 33, 717–726.

927 Coccia, E.M., Cicala, C., Charlesworth, A., Ciccarelli, C., Rossi, G.B.,  
928 Philipson, L., and Sorrentino, V. (1992). Regulation and expression of a  
929 growth arrest-specific gene (*gas5*) during growth, differentiation, and  
930 development. *Mol. Cell. Biol.* 12, 3514–3521.

931 Darrow, E.M., and Chadwick, B.P. (2013). Boosting transcription by  
932 transcription: enhancer-associated transcripts. *Chromosome Res.* 21, 713–  
933 724.

934 De Los Angeles, A., Ferrari, F., Xi, R., Fujiwara, Y., Benvenisty, N., Deng, H.,  
935 Hochedlinger, K., Jaenisch, R., Lee, S., Leitch, H.G., et al. (2015). Hallmarks  
936 of pluripotency. *Nature* 525, 469–478.

937 Deng, J., Liang, Y., Liu, C., He, S., and Wang, S. (2015). The up-regulation of  
938 long non-coding RNA AFAP1-AS1 is associated with the poor prognosis of  
939 NSCLC patients. *Biomed. Pharmacother. Biomed. Pharmacothérapie* 75, 8–  
940 11.

941 Denoeud, F., Kapranov, P., Ucla, C., Frankish, A., Castelo, R., Drenkow, J.,  
942 Lagarde, J., Alioto, T., Manzano, C., Chrast, J., et al. (2007). Prominent use of  
943 distal 5' transcription start sites and discovery of a large number of additional  
944 exons in ENCODE regions. *Genome Res.* 17, 746–759.

945 Denzler, R., Agarwal, V., Stefano, J., Bartel, D.P., and Stoffel, M. (2014).  
946 Assessing the ceRNA Hypothesis with Quantitative Measurements of miRNA  
947 and Target Abundance. *Mol. Cell* 54, 766–776.

948 Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H.,  
949 Guernec, G., Martin, D., Merkel, A., Knowles, D.G., et al. (2012). The  
950 GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene  
951 structure, evolution, and expression. *Genome Res.* 22, 1775–1789.

952 De Santa, F., Barozzi, I., Mietton, F., Ghisletti, S., Polletti, S., Tusi, B.K.,  
953 Muller, H., Ragoussis, J., Wei, C.-L., and Natoli, G. (2010). A Large Fraction  
954 of Extragenic RNA Pol II Transcription Sites Overlap Enhancers. *PLoS Biol.* 8,  
955 e1000384.

956 Dey, B.K., Mueller, A.C., and Dutta, A. (2014). Long non-coding RNAs as  
957 emerging regulators of differentiation, development, and disease.  
958 *Transcription* 5, e944014.

959 Dinger, M.E., Pang, K.C., Mercer, T.R., and Mattick, J.S. (2008a).  
960 Differentiating Protein-Coding and Noncoding RNA: Challenges and  
961 Ambiguities. *PLoS Comput. Biol.* 4, e1000176.

962 Dinger, M.E., Amaral, P.P., Mercer, T.R., Pang, K.C., Bruce, S.J., Gardiner,  
963 B.B., Askarian-Amiri, M.E., Ru, K., Solda, G., Simons, C., et al. (2008b). Long  
964 noncoding RNAs in mouse embryonic stem cell pluripotency and  
965 differentiation. *Genome Res.* 18, 1433–1445.

966 Di Ruscio, A., Ebralidze, A.K., Benoukraf, T., Amabile, G., Goff, L.A.,  
967 Terragni, J., Figueroa, M.E., De Figueiredo Pontes, L.L., Alberich-Jorda, M.,  
968 Zhang, P., et al. (2013). DNMT1-interacting RNAs block gene-specific DNA  
969 methylation. *Nature* 503, 371–376.

970 Djebali, S., Kapranov, P., Foissac, S., Lagarde, J., Reymond, A., Ucla, C.,  
971 Wyss, C., Drenkow, J., Dumais, E., Murray, R.R., et al. (2008). Efficient  
972 targeted transcript discovery via array-based normalization of RACE libraries.  
973 *Nat. Methods* 5, 629–635.

974 Du, Y., Kong, G., You, X., Zhang, S., Zhang, T., Gao, Y., Ye, L., and Zhang,  
975 X. (2012). Elevation of highly up-regulated in liver cancer (HULC) by hepatitis  
976 B virus X protein promotes hepatoma cell proliferation via down-regulating  
977 p18. *J. Biol. Chem.* 287, 26302–26311.

978 Ebert, M.S., Neilson, J.R., and Sharp, P.A. (2007). MicroRNA sponges:  
979 competitive inhibitors of small RNAs in mammalian cells. *Nat. Methods* 4,  
980 721–726.

- 981 Ebisuya, M., Yamamoto, T., Nakajima, M., and Nishida, E. (2008). Ripples  
982 from neighbouring transcription. *Nat. Cell Biol.* 10, 1106–1113.
- 983 Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998). Cluster  
984 analysis and display of genome-wide expression patterns. *Proc. Natl. Acad.*  
985 *Sci. U. S. A.* 95, 14863–14868.
- 986 ENCODE Project Consortium, Birney, E., Stamatoyannopoulos, J.A., Dutta,  
987 A., Guigó, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M.,  
988 Dermitzakis, E.T., et al. (2007). Identification and analysis of functional  
989 elements in 1% of the human genome by the ENCODE pilot project. *Nature*  
990 447, 799–816.
- 991 Fanucchi, S., Shibayama, Y., Burd, S., Weinberg, M.S., and Mhlanga, M.M.  
992 (2013). Chromosomal contact permits transcription between coregulated  
993 genes. *Cell* 155, 606–620.
- 994 Fatemi, R.P., Velmeshev, D., and Faghihi, M.A. (2014). De-repressing  
995 LncRNA-Targeted Genes to Upregulate Gene Expression: Focus on Small  
996 Molecule Therapeutics. *Mol. Ther. Nucleic Acids* 3, e196.
- 997 Fatica, A., and Bozzoni, I. (2014). Long non-coding RNAs: new players in cell  
998 differentiation and development. *Nat. Rev. Genet.* 15, 7–21.
- 999 Fragouli, E., Wells, D., Lager, A.E., Kayisli, U.A., and Patrizio, P. (2012).  
1000 Alteration of gene expression in human cumulus cells as a potential indicator  
1001 of oocyte aneuploidy. *Hum. Reprod. Oxf. Engl.* 27, 2559–2568.
- 1002 Franco-Zorrilla, J.M., Valli, A., Todesco, M., Mateos, I., Puga, M.I., Rubio-  
1003 Somoza, I., Leyva, A., Weigel, D., García, J.A., and Paz-Ares, J. (2007).  
1004 Target mimicry provides a new mechanism for regulation of microRNA  
1005 activity. *Nat. Genet.* 39, 1033–1037.
- 1006 Gabory, A., Jammes, H., and Dandolo, L. (2010). The H19 locus: role of an  
1007 imprinted non-coding RNA in growth and development. *BioEssays News Rev.*  
1008 *Mol. Cell. Dev. Biol.* 32, 473–480.
- 1009 Gao, W.-L., Liu, M., Yang, Y., Yang, H., Liao, Q., Bai, Y., Li, Y.-X., Li, D.,  
1010 Peng, C., and Wang, Y.-L. (2012). The imprinted H19 gene regulates human  
1011 placental trophoblast cell proliferation via encoding miR-675 that targets  
1012 Nodal Modulator 1 (NOMO1). *RNA Biol.* 9, 1002–1010.
- 1013 Gartler, S.M., Dyer, K.A., and Goldman, M.A. (1992). Mammalian X  
1014 chromosome inactivation. *Mol. Genet. Med.* 2, 121–160.
- 1015 Gebhardt, K.M., Feil, D.K., Dunning, K.R., Lane, M., and Russell, D.L. (2011).  
1016 Human cumulus cell gene expression as a biomarker of pregnancy outcome  
1017 after single embryo transfer. *Fertil. Steril.* 96, 47–52.e2.
- 1018 Gerstein, M.B., Bruce, C., Rozowsky, J.S., Zheng, D., Du, J., Korbil, J.O.,  
1019 Emanuelsson, O., Zhang, Z.D., Weissman, S., and Snyder, M. (2007). What

- 1020 is a gene, post-ENCODE? History and updated definition. *Genome Res.* 17,  
1021 669–681.
- 1022 Ginger, M.R., Gonzalez-Rimbau, M.F., Gay, J.P., and Rosen, J.M. (2001).  
1023 Persistent changes in gene expression induced by estrogen and progesterone  
1024 in the rat mammary gland. *Mol. Endocrinol. Baltim. Md* 15, 1993–2009.
- 1025 Ginger, M.R., Shore, A.N., Contreras, A., Rijnkels, M., Miller, J., Gonzalez-  
1026 Rimbau, M.F., and Rosen, J.M. (2006). A noncoding RNA is a potential  
1027 marker of cell fate during mammary gland development. *Proc. Natl. Acad. Sci.*  
1028 103, 5781–5786.
- 1029 Gingeras, T.R. (2007). Origin of phenotypes: genes and transcripts. *Genome*  
1030 *Res.* 17, 682–690.
- 1031 Grote, P., Wittler, L., Hendrix, D., Koch, F., Währisch, S., Beisaw, A., Macura,  
1032 K., Bläss, G., Kellis, M., Werber, M., et al. (2013). The Tissue-Specific lncRNA  
1033 Fendrr Is an Essential Regulator of Heart and Body Wall Development in the  
1034 Mouse. *Dev. Cell* 24, 206–214.
- 1035 Gong, C., and Maquat, L.E. (2011). lncRNAs transactivate STAU1-mediated  
1036 mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature* 470, 284–  
1037 288.
- 1038 Guo, J.U., Agarwal, V., Guo, H., and Bartel, D.P. (2014). Expanded  
1039 identification and characterization of mammalian circular RNAs. *Genome Biol*  
1040 15, 409.
- 1041 Guttman, M., and Rinn, J.L. (2012). Modular regulatory principles of large  
1042 non-coding RNAs. *Nature* 482, 339–346.
- 1043 Guttman, M., Amit, I., Garber, M., French, C., Lin, M.F., Feldser, D., Huarte,  
1044 M., Zuk, O., Carey, B.W., Cassady, J.P., et al. (2009). Chromatin signature  
1045 reveals over a thousand highly conserved large non-coding RNAs in  
1046 mammals. *Nature* 458, 223–227.
- 1047 Guttman, M., Garber, M., Levin, J.Z., Donaghey, J., Robinson, J., Adiconis,  
1048 X., Fan, L., Koziol, M.J., Gnirke, A., Nusbaum, C., et al. (2010). Ab initio  
1049 reconstruction of cell type-specific transcriptomes in mouse reveals the  
1050 conserved multi-exonic structure of lincRNAs. *Nat. Biotechnol.* 28, 503–510.
- 1051 Guttman, M., Donaghey, J., Carey, B.W., Garber, M., Grenier, J.K., Munson,  
1052 G., Young, G., Lucas, A.B., Ach, R., Bruhn, L., et al. (2011). lincRNAs act in  
1053 the circuitry controlling pluripotency and differentiation. *Nature* 477, 295–300.
- 1054 Hacısuleyman, E., Goff, L.A., Trapnell, C., Williams, A., Henao-Mejia, J., Sun,  
1055 L., McClanahan, P., Hendrickson, D.G., Sauvageau, M., Kelley, D.R., et al.  
1056 (2014). Topological organization of multichromosomal regions by the long  
1057 intergenic noncoding RNA Firre. *Nat. Struct. Mol. Biol.* 21, 198–206.
- 1058 Hackermüller, J., Reiche, K., Otto, C., Höslér, N., Blumert, C., Brocke-  
1059 Heidrich, K., Böhlig, L., Nitsche, A., Kasack, K., Ahnert, P., et al. (2014). *Cell*

- 1060 cycle, oncogenic and tumor suppressor pathways regulate numerous long  
1061 and macro non-protein coding RNAs. *Genome Biol* 15, R48.
- 1062 Hajjari, M., and Salavaty, A. (2015). HOTAIR: an oncogenic long non-coding  
1063 RNA in different cancers. *Cancer Biol. Med.* 12, 1.
- 1064 Hamazaki, N., Uesaka, M., Nakashima, K., Agata, K., and Imamura, T.  
1065 (2015). Gene activation-associated long noncoding RNAs function in mouse  
1066 preimplantation development. *Development* 142, 910–920.
- 1067 Han, J., Kim, D., and Morris, K.V. (2007). Promoter-associated RNA is  
1068 required for RNA-directed transcriptional gene silencing in human cells. *Proc.*  
1069 *Natl. Acad. Sci.* 104, 12422–12427.
- 1070 Hangauer, M.J., Vaughn, I.W., and McManus, M.T. (2013). Pervasive  
1071 Transcription of the Human Genome Produces Thousands of Previously  
1072 Unidentified Long Intergenic Noncoding RNAs. *PLoS Genet.* 9, e1003569.
- 1073 Hanna, J.H., Saha, K., and Jaenisch, R. (2010). Pluripotency and cellular  
1074 reprogramming: facts, hypotheses, unresolved issues. *Cell* 143, 508–525.
- 1075 Hansen, T.B., Jensen, T.I., Clausen, B.H., Bramsen, J.B., Finsen, B.,  
1076 Damgaard, C.K., and Kjems, J. (2013). Natural RNA circles function as  
1077 efficient microRNA sponges. *Nature* 495, 384–388.
- 1078 Hauptman, N., and Glavač, D. (2013). Long non-coding RNA in cancer. *Int. J.*  
1079 *Mol. Sci.* 14, 4655–4669.
- 1080 He, S., Liu, S., and Zhu, H. (2011). The sequence, structure and evolutionary  
1081 features of HOTAIR in mammals. *BMC Evol. Biol.* 11, 1.
- 1082 He, S., Zhang, H., Liu, H., and Zhu, H. (2015). LongTarget: a tool to predict  
1083 lncRNA DNA-binding motifs and binding sites via Hoogsteen base-pairing  
1084 analysis. *Bioinformatics* 31, 178–186.
- 1085 He, Y., Vogelstein, B., Velculescu, V.E., Papadopoulos, N., and Kinzler, K.W.  
1086 (2008). The antisense transcriptomes of human cells. *Science* 322, 1855–  
1087 1857.
- 1088 Herriges, M.J., Swarr, D.T., Morley, M.P., Rathi, K.S., Peng, T., Stewart, K.M.,  
1089 and Morrisey, E.E. (2014). Long noncoding RNAs are spatially correlated with  
1090 transcription factors and regulate lung development. *Genes Dev.* 28, 1363–  
1091 1379.
- 1092 Higashimoto, K., Soejima, H., Saito, T., Okumura, K., and Mukai, T. (2006).  
1093 Imprinting disruption of the CDKN1C/KCNQ1OT1 domain: the molecular  
1094 mechanisms causing Beckwith-Wiedemann syndrome and cancer. *Cytogenet.*  
1095 *Genome Res.* 113, 306–312.
- 1096 Hillier, S.G. (2008). Research challenge: what is the best non-invasive test of  
1097 oocyte/embryo competence? *Mol. Hum. Reprod.* 14, 665.

1098 Hirose, T., Virnicchi, G., Tanigawa, A., Naganuma, T., Li, R., Kimura, H.,  
1099 Yokoi, T., Nakagawa, S., Bénard, M., Fox, A.H., et al. (2014). NEAT1 long  
1100 noncoding RNA regulates transcription via protein sequestration within  
1101 subnuclear bodies. *Mol. Biol. Cell* 25, 169–183.

1102 Housman, G., and Ulitsky, I. (2016). Methods for distinguishing between  
1103 protein-coding and long noncoding RNAs and the elusive biological purpose  
1104 of translation of long noncoding RNAs. *Biochim. Biophys. Acta BBA - Gene  
1105 Regul. Mech.* 1859, 31–40.

1106 Hu, W., Yuan, B., Flygare, J., and Lodish, H.F. (2011). Long noncoding RNA-  
1107 mediated anti-apoptotic activity in murine erythroid terminal differentiation.  
1108 *Genes Dev.* 25, 2573–2578.

1109 Huang, X., Liu, C., Hao, C., Tang, Q., Liu, R., Lin, S., Zhang, L., and Yan, W.  
1110 (2016a). Identification of altered microRNAs and mRNAs in the cumulus cells  
1111 of PCOS patients. *Reprod. Camb. Engl.*

1112 Huang, X., Hao, C., Bao, H., Wang, M., and Dai, H. (2016b). Aberrant  
1113 expression of long noncoding RNAs in cumulus cells isolated from PCOS  
1114 patients. *J. Assist. Reprod. Genet.* 33, 111–121.

1115 Huarte, M., and Marín-Béjar, O. (2015). Long noncoding RNAs: from  
1116 identification to functions and mechanisms. *Adv. Genomics Genet.* 257.

1117 Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M.J., Kenzelmann-  
1118 Broz, D., Khalil, A.M., Zuk, O., Amit, I., Rabani, M., et al. (2010). A Large  
1119 Intergenic Noncoding RNA Induced by p53 Mediates Global Gene Repression  
1120 in the p53 Response. *Cell* 142, 409–419.

1121 Hung, T., Wang, Y., Lin, M.F., Koegel, A.K., Kotake, Y., Grant, G.D., Horlings,  
1122 H.M., Shah, N., Umbricht, C., Wang, P., et al. (2011). Extensive and  
1123 coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat.  
1124 Genet.* 43, 621–629.

1125 Huo, J.S., and Zambidis, E.T. (2013). Pivots of pluripotency: The roles of non-  
1126 coding RNA in regulating embryonic and induced pluripotent stem cells.  
1127 *Biochim. Biophys. Acta BBA - Gen. Subj.* 1830, 2385–2394.

1128 Jia, H., Osak, M., Bogu, G.K., Stanton, L.W., Johnson, R., and Lipovich, L.  
1129 (2010). Genome-wide computational identification and manual annotation of  
1130 human long noncoding RNA genes. *RNA* 16, 1478–1487.

1131 Jia, W., Chen, W., and Kang, J. (2013). The functions of microRNAs and long  
1132 non-coding RNAs in embryonic and induced pluripotent stem cells. *Genomics  
1133 Proteomics Bioinformatics* 11, 275–283.

1134 Johnson, R. (2012). Long non-coding RNAs in Huntington’s disease  
1135 neurodegeneration. *Neurobiol. Dis.* 46, 245–254.

- 1136 Johnsson, P., Lipovich, L., Grandér, D., and Morris, K.V. (2014). Evolutionary  
1137 conservation of long non-coding RNAs; sequence, structure, function.  
1138 *Biochim. Biophys. Acta BBA - Gen. Subj.* *1840*, 1063–1071.
- 1139 Kaikkonen, M.U., Spann, N.J., Heinz, S., Romanoski, C.E., Allison, K.A.,  
1140 Stender, J.D., Chun, H.B., Tough, D.F., Prinjha, R.K., Benner, C., et al.  
1141 (2013). Remodeling of the Enhancer Landscape during Macrophage  
1142 Activation Is Coupled to Enhancer Transcription. *Mol. Cell* *51*, 310–325.
- 1143 Kallen, A.N., Zhou, X.-B., Xu, J., Qiao, C., Ma, J., Yan, L., Lu, L., Liu, C., Yi,  
1144 J.-S., Zhang, H., et al. (2013). The Imprinted H19 LncRNA Antagonizes Let-7  
1145 MicroRNAs. *Mol. Cell* *52*, 101–112.
- 1146 Kanduri, C. (2011). *Kcnq1ot1*: a chromatin regulatory RNA. *Semin. Cell Dev.*  
1147 *Biol.* *22*, 343–350.
- 1148 Kanduri, C. (2016). Long noncoding RNAs: Lessons from genomic imprinting.  
1149 *Biochim. Biophys. Acta* *1859*, 102–111.
- 1150 Kaneko, S., Son, J., Shen, S.S., Reinberg, D., and Bonasio, R. (2013). PRC2  
1151 binds active promoters and contacts nascent RNAs in embryonic stem cells.  
1152 *Nat. Struct. Mol. Biol.* *20*, 1258–1264.
- 1153 Kaneko, S., Bonasio, R., Saldaña-Meyer, R., Yoshida, T., Son, J., Nishino, K.,  
1154 Umezawa, A., and Reinberg, D. (2014). Interactions between JARID2 and  
1155 noncoding RNAs regulate PRC2 recruitment to chromatin. *Mol. Cell* *53*, 290–  
1156 300.
- 1157 Kapranov, P., Drenkow, J., Cheng, J., Long, J., Helt, G., Dike, S., and  
1158 Gingeras, T.R. (2005). Examples of the complex architecture of the human  
1159 transcriptome revealed by RACE and high-density tiling arrays. *Genome Res.*  
1160 *15*, 987–997.
- 1161 Kapranov, P., Willingham, A.T., and Gingeras, T.R. (2007). Genome-wide  
1162 transcription and the implications for genomic organization. *Nat. Rev. Genet.*  
1163 *8*, 413–423.
- 1164 Karreth, F.A., Tay, Y., Perna, D., Ala, U., Tan, S.M., Rust, A.G., DeNicola, G.,  
1165 Webster, K.A., Weiss, D., Perez-Mancera, P.A., et al. (2011). In Vivo  
1166 Identification of Tumor- Suppressive PTEN ceRNAs in an Oncogenic BRAF-  
1167 Induced Mouse Model of Melanoma. *Cell* *147*, 382–395.
- 1168 Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura,  
1169 M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., et al. (2005). Antisense  
1170 transcription in the mammalian transcriptome. *Science* *309*, 1564–1566.
- 1171 Kelley, D., and Rinn, J. (2012). Transposable elements reveal a stem cell-  
1172 specific class of long noncoding RNAs. *Genome Biol.* *13*, R107.
- 1173 Keniry, A., Oxley, D., Monnier, P., Kyba, M., Dandolo, L., Smits, G., and Reik,  
1174 W. (2012). The H19 lincRNA is a developmental reservoir of miR-675 that  
1175 suppresses growth and *Igf1r*. *Nat. Cell Biol.* *14*, 659–665.



- 1176 Khalil, A.M., Guttman, M., Huarte, M., Garber, M., Raj, A., Morales, D.R.,  
1177 Thomas, K., Presser, A., Bernstein, B.E., van Oudenaarden, A., et al. (2009).  
1178 Many human large intergenic noncoding RNAs associate with chromatin-  
1179 modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci.* *106*,  
1180 11667–11672.
- 1181 Khorkova, O., Hsiao, J., and Wahlestedt, C. (2015). Basic biology and  
1182 therapeutic implications of lncRNA. *Adv. Drug Deliv. Rev.* *87*, 15–24.
- 1183 Khoueiry, R., Khoureiry, R., Ibala-Rhomdane, S., Méry, L., Blachère, T.,  
1184 Guérin, J.-F., Lornage, J., and Lefèvre, A. (2008). Dynamic CpG methylation  
1185 of the KCNQ1OT1 gene during maturation of human oocytes. *J. Med. Genet.*  
1186 *45*, 583–588.
- 1187 Kim, J., Bretz, C.L., and Lee, S. (2015). Epigenetic instability of imprinted  
1188 genes in human cancers. *Nucleic Acids Res.* *43*, 10689–10699.
- 1189 Kim, S.H., Suddath, F.L., Quigley, G.J., McPherson, A., Sussman, J.L., Wang,  
1190 A.H.J., Seeman, N.C., and Rich, A. (1974). Three-dimensional tertiary  
1191 structure of yeast phenylalanine transfer RNA. *Science* *185*, 435–440.
- 1192 Kim, T.-K., Hemberg, M., Gray, J.M., Costa, A.M., Bear, D.M., Wu, J., Harmin,  
1193 D.A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., et al. (2010).  
1194 Widespread transcription at neuronal activity-regulated enhancers. *Nature*  
1195 *465*, 182–187.
- 1196 Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in  
1197 animals. *Nat. Rev. Mol. Cell Biol.* *10*, 126–139.
- 1198 Kino, T., Hurt, D.E., Ichijo, T., Nader, N., and Chrousos, G.P. (2010).  
1199 Noncoding RNA Gas5 is a growth arrest and starvation-associated repressor  
1200 of the glucocorticoid receptor. *Sci. Signal.* *3*, ra8.
- 1201 Klattenhoff, C.A., Scheuermann, J.C., Surface, L.E., Bradley, R.K., Fields,  
1202 P.A., Steinhauser, M.L., Ding, H., Butty, V.L., Torrey, L., Haas, S., et al.  
1203 (2013). Braveheart, a Long Noncoding RNA Required for Cardiovascular  
1204 Lineage Commitment. *Cell* *152*, 570–583.
- 1205 Knauss, J.L., and Sun, T. (2013). Regulatory mechanisms of long noncoding  
1206 RNAs in vertebrate central nervous system development and function.  
1207 *Neuroscience* *235*, 200–214.
- 1208 Kowalczyk, M.S., Higgs, D.R., and Gingeras, T.R. (2012). Molecular biology:  
1209 RNA discrimination. *Nature* *482*, 310–311.
- 1210 Kretz, M., Siprashvili, Z., Chu, C., Webster, D.E., Zehnder, A., Qu, K., Lee,  
1211 C.S., Flockhart, R.J., Groff, A.F., Chow, J., et al. (2012a). Control of somatic  
1212 tissue differentiation by the long non-coding RNA TINCR. *Nature* *493*, 231–  
1213 235.
- 1214 Kretz, M., Webster, D.E., Flockhart, R.J., Lee, C.S., Zehnder, A., Lopez-  
1215 Pajares, V., Qu, K., Zheng, G.X.Y., Chow, J., Kim, G.E., et al. (2012b).

- 1216 Suppression of progenitor differentiation requires the long noncoding RNA  
1217 ANCR. *Genes Dev.* 26, 338–343.
- 1218 Lai, F., Orom, U.A., Cesaroni, M., Beringer, M., Taatjes, D.J., Blobel, G.A.,  
1219 and Shiekhata, R. (2013). Activating RNAs associate with Mediator to  
1220 enhance chromatin architecture and transcription. *Nature* 494, 497–501.
- 1221 Lam, M.T.Y., Cho, H., Lesch, H.P., Gosselin, D., Heinz, S., Tanaka-Oishi, Y.,  
1222 Benner, C., Kaikkonen, M.U., Kim, A.S., Kosaka, M., et al. (2013). Rev-Erbs  
1223 repress macrophage gene expression by inhibiting enhancer-directed  
1224 transcription. *Nature* 498, 511–515.
- 1225 Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J.,  
1226 Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial  
1227 sequencing and analysis of the human genome. *Nature* 409, 860–921.
- 1228 Lanz, R.B., McKenna, N.J., Onate, S.A., Albrecht, U., Wong, J., Tsai, S.Y.,  
1229 Tsai, M.J., and O'Malley, B.W. (1999). A steroid receptor coactivator, SRA,  
1230 functions as an RNA and is present in an SRC-1 complex. *Cell* 97, 17–27.
- 1231 Lazorthes, S., Vallot, C., Briois, S., Aguirrebengoa, M., Thuret, J.-Y., Laurent,  
1232 G.S., Rougeulle, C., Kapranov, P., Mann, C., Trouche, D., et al. (2015). A  
1233 vlincRNA participates in senescence maintenance by relieving H2AZ-  
1234 mediated repression at the INK4 locus. *Nat. Commun.* 6, 5971.
- 1235 Lee, J.T. (2012). Epigenetic Regulation by Long Noncoding RNAs. *Science*  
1236 338, 1435–1439.
- 1237 Léveillé, N., Melo, C.A., Rooijers, K., Díaz-Lagares, A., Melo, S.A., Korkmaz,  
1238 G., Lopes, R., Moqadam, F.A., Maia, A.R., Wijchers, P.J., et al. (2015).  
1239 Genome-wide profiling of p53-regulated enhancer RNAs uncovers a subset of  
1240 enhancers controlled by a lincRNA. *Nat. Commun.* 6, 6520.
- 1241 Leygue, E. (2007). Steroid receptor RNA activator (SRA1): unusual bifaceted  
1242 gene products with suspected relevance to breast cancer. *Nucl. Recept.*  
1243 *Signal.* 5, e006.
- 1244 Li, H., Yu, B., Li, J., Su, L., Yan, M., Zhu, Z., and Liu, B. (2014).  
1245 Overexpression of lincRNA H19 enhances carcinogenesis and metastasis of  
1246 gastric cancer. *Oncotarget* 5, 2318–2329.
- 1247 Li, K., Blum, Y., Verma, A., Liu, Z., Pramanik, K., Leigh, N.R., Chun, C.Z.,  
1248 Samant, G.V., Zhao, B., Garnaas, M.K., et al. (2010). A noncoding antisense  
1249 RNA in tie-1 locus regulates tie-1 function in vivo. *Blood* 115, 133–139.
- 1250 Li, L., Wang, X., Sasidharan, R., Stolc, V., Deng, W., He, H., Korb, J., Chen,  
1251 X., Tongprasit, W., Ronald, P., et al. (2007). Global identification and  
1252 characterization of transcriptionally active regions in the rice genome. *PLoS*  
1253 *One* 2, e294.
- 1254 Lin, N., Chang, K.-Y., Li, Z., Gates, K., Rana, Z.A., Dang, J., Zhang, D., Han,  
1255 T., Yang, C.-S., Cunningham, T.J., et al. (2014). An evolutionarily conserved

- 1256 long noncoding RNA TUNA controls pluripotency and neural lineage  
1257 commitment. *Mol. Cell* 53, 1005–1019.
- 1258 Liu, S., Zhang, X., Shi, C., Lin, J., Chen, G., Wu, B., Wu, L., Shi, H., Yuan, Y.,  
1259 Zhou, W., et al. (2015). Altered microRNAs expression profiling in cumulus  
1260 cells from patients with polycystic ovary syndrome. *J. Transl. Med.* 13, 238.
- 1261 Liu, X., Li, D., Zhang, W., Guo, M., and Zhan, Q. (2012). Long non-coding  
1262 RNA gadd7 interacts with TDP-43 and regulates Cdk6 mRNA decay. *EMBO*  
1263 *J.* 31, 4415–4427.
- 1264 Loewer, S., Cabili, M.N., Guttman, M., Loh, Y.-H., Thomas, K., Park, I.H.,  
1265 Garber, M., Curran, M., Onder, T., Agarwal, S., et al. (2010). Large intergenic  
1266 non-coding RNA-RoR modulates reprogramming of human induced  
1267 pluripotent stem cells. *Nat. Genet.* 42, 1113–1117.
- 1268 Louro, R., Smirnova, A.S., and Verjovski-Almeida, S. (2009). Long intronic  
1269 noncoding RNA transcription: Expression noise or expression choice?  
1270 *Genomics* 93, 291–298.
- 1271 Lu, K., Li, W., Liu, X., Sun, M., Zhang, M., Wu, W., Xie, W., and Hou, Y.  
1272 (2013). Long non-coding RNA MEG3 inhibits NSCLC cells proliferation and  
1273 induces apoptosis by affecting p53 expression. *BMC Cancer* 13, 461.
- 1274 Lundin, K., Bergh, C., and Hardarson, T. (2001). Early embryo cleavage is a  
1275 strong indicator of embryo quality in human IVF. *Hum. Reprod. Oxf. Engl.* 16,  
1276 2652–2657.
- 1277 Magistri, M., Faghihi, M.A., St Laurent, G., and Wahlestedt, C. (2012).  
1278 Regulation of chromatin structure by long noncoding RNAs: focus on natural  
1279 antisense transcripts. *Trends Genet. TIG* 28, 389–396.
- 1280 Mallo, M., and Alonso, C.R. (2013). The regulation of Hox gene expression  
1281 during animal development. *Development* 140, 3951–3963.
- 1282 Marques, A.C., Hughes, J., Graham, B., Kowalczyk, M.S., Higgs, D.R., and  
1283 Ponting, C.P. (2013). Chromatin signatures at transcriptional start sites  
1284 separate two equally populated yet distinct classes of intergenic long  
1285 noncoding RNAs. *Genome Biol.* 14, R131.
- 1286 Martianov, I., Ramadass, A., Serra Barros, A., Chow, N., and Akoulitchev, A.  
1287 (2007). Repression of the human dihydrofolate reductase gene by a non-  
1288 coding interfering transcript. *Nature* 445, 666–670.
- 1289 Mastenbroek, S., van der Veen, F., Aflatoonian, A., Shapiro, B., Bossuyt, P.,  
1290 and Repping, S. (2011). Embryo selection in IVF. *Hum. Reprod. Oxf. Engl.* 26,  
1291 964–966.
- 1292 Matouk, I.J., DeGroot, N., Mezan, S., Ayesh, S., Abu-lail, R., Hochberg, A.,  
1293 and Galun, E. (2007). The H19 non-coding RNA is essential for human tumor  
1294 growth. *PloS One* 2, e845.

- 1295 Mattick, J.S. (2004). RNA regulation: a new genetics? *Nat. Rev. Genet.* *5*,  
1296 316–323.
- 1297 Mattick, J.S. (2006). Non-coding RNA. *Hum. Mol. Genet.* *15*, R17–R29.
- 1298 Mattick, J.S., and Rinn, J.L. (2015). Discovery and annotation of long  
1299 noncoding RNAs. *Nat. Struct. Mol. Biol.* *22*.
- 1300 McClelland, M.L., Mesh, K., Lorenzana, E., Chopra, V.S., Segal, E.,  
1301 Watanabe, C., Haley, B., Mayba, O., Yaylaoglu, M., Gnad, F., et al. (2016).  
1302 CCAT1 is an enhancer-templated RNA that predicts BET sensitivity in  
1303 colorectal cancer. *J. Clin. Invest.* *126*, 639–652.
- 1304 McKenzie, L.J., Pangas, S.A., Carson, S.A., Kovanci, E., Cisneros, P., Buster,  
1305 J.E., Amato, P., and Matzuk, M.M. (2004). Human cumulus granulosa cell  
1306 gene expression: a predictor of fertilization and embryo selection in women  
1307 undergoing IVF. *Hum. Reprod. Oxf. Engl.* *19*, 2869–2874.
- 1308 Melo, C.A., Drost, J., Wijchers, P.J., van de Werken, H., de Wit, E., Vrieling,  
1309 J.A.F.O., Elkon, R., Melo, S.A., Léveillé, N., Kalluri, R., et al. (2013). eRNAs  
1310 Are Required for p53-Dependent Enhancer Activity and Gene Transcription.  
1311 *Mol. Cell* *49*, 524–535.
- 1312 Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier,  
1313 L., Mackowiak, S.D., Gregersen, L.H., Munschauer, M., et al. (2013). Circular  
1314 RNAs are a large class of animal RNAs with regulatory potency. *Nature* *495*,  
1315 333–338.
- 1316 Mercer, T.R., Qureshi, I.A., Gokhan, S., Dinger, M.E., Li, G., Mattick, J.S., and  
1317 Mehler, M.F. (2010). Long noncoding RNAs in neuronal-glia fate specification  
1318 and oligodendrocyte lineage maturation. *BMC Neurosci.* *11*, 14.
- 1319 Milligan, M.J., and Lipovich, L. (2014). Pseudogene-derived lncRNAs:  
1320 emerging regulators of gene expression. *Front. Genet.* *5*, 476.
- 1321 Modarresi, F., Faghihi, M.A., Lopez-Toledano, M.A., Fatemi, R.P., Magistri,  
1322 M., Brothers, S.P., van der Brug, M.P., and Wahlestedt, C. (2012). Inhibition  
1323 of natural antisense transcripts in vivo results in gene-specific transcriptional  
1324 upregulation. *Nat. Biotechnol.* *30*, 453–459.
- 1325 Mondal, T., and Kanduri, C. (2012). Noncoding RNA scaffolds in pluripotency.  
1326 *Circ. Res.* *110*, 1162–1165.
- 1327 Morris, K.V., and Mattick, J.S. (2014). The rise of regulatory RNA. *Nat. Rev.*  
1328 *Genet.* *15*, 423–437.
- 1329 Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008).  
1330 Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat.*  
1331 *Methods* *5*, 621–628.
- 1332 Mourtada-Maarabouni, M., Hedge, V.L., Kirkham, L., Farzaneh, F., and  
1333 Williams, G.T. (2008). Growth arrest in human T-cells is controlled by the non-

- 1334 coding RNA growth-arrest-specific transcript 5 (GAS5). *J. Cell Sci.* *121*, 939–  
1335 946.
- 1336 Mousavi, K., Zare, H., Dell’Orso, S., Grontved, L., Gutierrez-Cruz, G., Derfoul,  
1337 A., Hager, G.L., and Sartorelli, V. (2013). eRNAs Promote Transcription by  
1338 Establishing Chromatin Accessibility at Defined Genomic Loci. *Mol. Cell* *51*,  
1339 606–617.
- 1340 Mullis, K.B. (1990). Target amplification for DNA analysis by the polymerase  
1341 chain reaction. *Ann. Biol. Clin. (Paris)* *48*, 579–582.
- 1342 Murthy, U.M.S., and Rangarajan, P.N. (2010). Identification of protein  
1343 interaction regions of VINC/NEAT1/Men epsilon RNA. *FEBS Lett.* *584*, 1531–  
1344 1535.
- 1345 Nagano, T., Mitchell, J.A., Sanz, L.A., Pauler, F.M., Ferguson-Smith, A.C.,  
1346 Feil, R., and Fraser, P. (2008). The Air noncoding RNA epigenetically silences  
1347 transcription by targeting G9a to chromatin. *Science* *322*, 1717–1720.
- 1348 Nakaya, H.I., Amaral, P.P., Louro, R., Lopes, A., Fachel, A.A., Moreira, Y.B.,  
1349 El-Jundi, T.A., da Silva, A.M., Reis, E.M., and Verjovski-Almeida, S. (2007).  
1350 Genome mapping and expression analyses of human intronic noncoding  
1351 RNAs reveal tissue-specific patterns and enrichment in genes related to  
1352 regulation of transcription. *Genome Biol.* *8*, R43.
- 1353 Necsulea, A., Soumillon, M., Warnefors, M., Liechti, A., Daish, T., Zeller, U.,  
1354 Baker, J.C., Grützner, F., and Kaessmann, H. (2014). The evolution of  
1355 lncRNA repertoires and expression patterns in tetrapods. *Nature* *505*, 635–  
1356 640.
- 1357 Ng, S.-Y., Johnson, R., and Stanton, L.W. (2012). Human long non-coding  
1358 RNAs promote pluripotency and neuronal differentiation by association with  
1359 chromatin modifiers and transcription factors. *EMBO J.* *31*, 522–533.
- 1360 Niakan, K.K., Han, J., Pedersen, R.A., Simon, C., and Pera, R.A.R. (2012).  
1361 Human pre-implantation embryo development. *Dev. Camb. Engl.* *139*, 829–  
1362 841.
- 1363 Nitsche, A., Rose, D., Fasold, M., Reiche, K., and Stadler, P.F. (2015).  
1364 Comparison of splice sites reveals that long noncoding RNAs are  
1365 evolutionarily well conserved. *RNA* *21*, 801–812.
- 1366 Ørom, U.A., and Shiekhattar, R. (2013). Long Noncoding RNAs Usher In a  
1367 New Era in the Biology of Enhancers. *Cell* *154*, 1190–1193.
- 1368 Ørom, U.A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G.,  
1369 Lai, F., Zytnicki, M., Notredame, C., Huang, Q., et al. (2010). Long Noncoding  
1370 RNAs with Enhancer-like Function in Human Cells. *Cell* *143*, 46–58.
- 1371 Otte, A.P., and Kwaks, T.H.J. (2003). Gene repression by Polycomb group  
1372 protein complexes: a distinct complex for every occasion? *Curr. Opin. Genet.*  
1373 *Dev.* *13*, 448–454.

- 1374 Palazzo, A.F., and Gregory, T.R. (2014). The case for junk DNA. *PLoS Genet*  
1375 *10*, e1004351.
- 1376 Pandey, R.R., Mondal, T., Mohammad, F., Enroth, S., Redrup, L.,  
1377 Komorowski, J., Nagano, T., Mancini-DiNardo, D., and Kanduri, C. (2008).  
1378 *Kcnq1ot1* Antisense Noncoding RNA Mediates Lineage-Specific  
1379 Transcriptional Silencing through Chromatin-Level Regulation. *Mol. Cell* *32*,  
1380 232–246.
- 1381 Pang, K.C., Frith, M.C., and Mattick, J.S. (2006). Rapid evolution of  
1382 noncoding RNAs: lack of conservation does not mean lack of function. *Trends*  
1383 *Genet.* *22*, 1–5.
- 1384 Paranjpe, S.S., Jacobi, U.G., van Heeringen, S.J., and Veenstra, G.J. (2013).  
1385 A genome-wide survey of maternal and embryonic transcripts during *Xenopus*  
1386 *tropicalis* development. *BMC Genomics* *14*, 762.
- 1387 Pasmant, E., Laurendeau, I., Héron, D., Vidaud, M., Vidaud, D., and Bièche, I.  
1388 (2007). Characterization of a germ-line deletion, including the entire  
1389 *INK4/ARF* locus, in a melanoma-neural system tumor family: identification of  
1390 *ANRIL*, an antisense noncoding RNA whose expression coclusters with *ARF*.  
1391 *Cancer Res.* *67*, 3963–3969.
- 1392 Pauli, A., Rinn, J.L., and Schier, A.F. (2011). Non-coding RNAs as regulators  
1393 of embryogenesis. *Nat. Rev. Genet.* *12*, 136–149.
- 1394 Pauli, A., Valen, E., Lin, M.F., Garber, M., Vastenhouw, N.L., Levin, J.Z., Fan,  
1395 L., Sandelin, A., Rinn, J.L., Regev, A., et al. (2012). Systematic identification  
1396 of long noncoding RNAs expressed during zebrafish embryogenesis. *Genome*  
1397 *Res.* *22*, 577–591.
- 1398 Pefanis, E., Wang, J., Rothschild, G., Lim, J., Kazadi, D., Sun, J., Federation,  
1399 A., Chao, J., Elliott, O., Liu, Z.-P., et al. (2015). RNA Exosome-Regulated  
1400 Long Non-Coding RNA Transcription Controls Super-Enhancer Activity. *Cell*  
1401 *161*, 774–789.
- 1402 Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S., and Brockdorff, N.  
1403 (1996). Requirement for *Xist* in X chromosome inactivation. *Nature* *379*, 131–  
1404 137.
- 1405 Philippe, N., Salson, M., Commes, T., and Rivals, E. (2013). CRAC: an  
1406 integrated approach to the analysis of RNA-seq reads. *Genome Biol.* *14*, R30.
- 1407 Ponjavic, J., Ponting, C.P., and Lunter, G. (2007). Functionality or  
1408 transcriptional noise? Evidence for selection within long noncoding RNAs.  
1409 *Genome Res.* *17*, 556–565.
- 1410 Ponting, C.P., Oliver, P.L., and Reik, W. (2009). Evolution and Functions of  
1411 Long Noncoding RNAs. *Cell* *136*, 629–641.
- 1412 Prickett, A.R., and Oakey, R.J. (2012). A survey of tissue-specific genomic  
1413 imprinting in mammals. *Mol. Genet. Genomics* *287*, 621–630.

1414 Quinn, J.J., Ilik, I.A., Qu, K., Georgiev, P., Chu, C., Akhtar, A., and Chang,  
1415 H.Y. (2014). Revealing long noncoding RNA architecture and functions using  
1416 domain-specific chromatin isolation by RNA purification. *Nat. Biotechnol.* *32*,  
1417 933–940.

1418 Racowsky, C., Vernon, M., Mayer, J., Ball, G.D., Behr, B., Pomeroy, K.O.,  
1419 Winger, D., Gibbons, W., Conaghan, J., and Stern, J.E. (2010).  
1420 Standardization of grading embryo morphology. *J. Assist. Reprod. Genet.* *27*,  
1421 437–439.

1422 Ramirez, J.-M. (2010). Human pluripotent stem cells: From biology to cell  
1423 therapy. *World J. Stem Cells* *2*, 24.

1424 Ramos, A.D., Andersen, R.E., Liu, S.J., Nowakowski, T.J., Hong, S.J., Gertz,  
1425 C.C., Salinas, R.D., Zarabi, H., Kriegstein, A.R., and Lim, D.A. (2015). The  
1426 long noncoding RNA Pnky regulates neuronal differentiation of embryonic and  
1427 postnatal neural stem cells. *Cell Stem Cell* *16*, 439–447.

1428 Rands, C.M., Meader, S., Ponting, C.P., and Lunter, G. (2014). 8.2% of the  
1429 Human Genome Is Constrained: Variation in Rates of Turnover across  
1430 Functional Element Classes in the Human Lineage. *PLoS Genet.* *10*,  
1431 e1004525.

1432 Rashid, F., Shah, A., and Shan, G. (2016). Long Non-coding RNAs in the  
1433 Cytoplasm. *Genomics Proteomics Bioinformatics* *14*, 73–80.

1434 Ravasi, T. (2005). Experimental validation of the regulated expression of large  
1435 numbers of non-coding RNAs from the mouse genome. *Genome Res.* *16*, 11–  
1436 19.

1437 Rinn, J.L., and Chang, H.Y. (2012). Genome Regulation by Long Noncoding  
1438 RNAs. *Annu. Rev. Biochem.* *81*, 145–166.

1439 Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Brugmann, S.A.,  
1440 Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E., et al. (2007).  
1441 Functional Demarcation of Active and Silent Chromatin Domains in Human  
1442 HOX Loci by Noncoding RNAs. *Cell* *129*, 1311–1323.

1443 Rosa, A., and Brivanlou, A. (2013). Regulatory Non-Coding RNAs in  
1444 Pluripotent Stem Cells. *Int. J. Mol. Sci.* *14*, 14346–14373.

1445 Rouskin, S., Zubradt, M., Washietl, S., Kellis, M., and Weissman, J.S. (2013).  
1446 Genome-wide probing of RNA structure reveals active unfolding of mRNA  
1447 structures in vivo. *Nature* *505*, 701–705.

1448 Ruiz-Orera, J., Messeguer, X., Subirana, J.A., and Alba, M.M. (2014). Long  
1449 non-coding RNAs as a source of new peptides. *eLife* *3*, e03523.

1450 Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P.P. (2011). A  
1451 ceRNA Hypothesis: The Rosetta Stone of a Hidden RNA Language? *Cell* *146*,  
1452 353–358.

- 1453 Sasaki, Y.T.F., Sano, M., Kin, T., Asai, K., and Hirose, T. (2007). Coordinated  
1454 expression of ncRNAs and HOX mRNAs in the human HOXA locus. *Biochem.*  
1455 *Biophys. Res. Commun.* 357, 724–730.
- 1456 Schaukowitch, K., Joo, J.-Y., Liu, X., Watts, J.K., Martinez, C., and Kim, T.-K.  
1457 (2014). Enhancer RNA Facilitates NELF Release from Immediate Early  
1458 Genes. *Mol. Cell* 56, 29–42.
- 1459 Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. (1995). Quantitative  
1460 monitoring of gene expression patterns with a complementary DNA  
1461 microarray. *Science* 270, 467–470.
- 1462 Schmitz, K.-M., Mayer, C., Postepska, A., and Grummt, I. (2010). Interaction  
1463 of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b  
1464 and silencing of rRNA genes. *Genes Dev.* 24, 2264–2269.
- 1465 Schoenhals, M., Kassambara, A., De Vos, J., Hose, D., Moreaux, J., and  
1466 Klein, B. (2009). Embryonic stem cell markers expression in cancers.  
1467 *Biochem. Biophys. Res. Commun.* 383, 157–162.
- 1468 Schorderet, P., and Duboule, D. (2011). Structural and Functional Differences  
1469 in the Long Non-Coding RNA *Hotair* in Mouse and Human. *PLoS Genet.* 7,  
1470 e1002071.
- 1471 Schultes, E.A., Spasic, A., Mohanty, U., and Bartel, D.P. (2005). Compact and  
1472 ordered collapse of randomly generated RNA sequences. *Nat. Struct. 38 Mol.*  
1473 *Biol.* 12, 1130–1136.
- 1474 Schwartz, J.C., Younger, S.T., Nguyen, N.-B., Hardy, D.B., Monia, B.P.,  
1475 Corey, D.R., and Janowski, B.A. (2008). Antisense transcripts are targets for  
1476 activating small RNAs. *Nat. Struct. 38 Mol. Biol.* 15, 842–848.
- 1477 Seto, A.G., Kingston, R.E., and Lau, N.C. (2007). The Coming of Age for Piwi  
1478 Proteins. *Mol. Cell* 26, 603–609.
- 1479 Sheik Mohamed, J., Gaughwin, P.M., Lim, B., Robson, P., and Lipovich, L.  
1480 (2010). Conserved long noncoding RNAs transcriptionally regulated by Oct4  
1481 and Nanog modulate pluripotency in mouse embryonic stem cells. *RNA N. Y.*  
1482 *N* 16, 324–337.
- 1483 Shen, X., Kim, W., Fujiwara, Y., Simon, M.D., Liu, Y., Mysliwiec, M.R., Yuan,  
1484 G.-C., Lee, Y., and Orkin, S.H. (2009). *Jumonji* modulates polycomb activity  
1485 and self-renewal versus differentiation of stem cells. *Cell* 139, 1303–1314.
- 1486 Sleutels, F., Zwart, R., and Barlow, D.P. (2002). The non-coding *Air* RNA is  
1487 required for silencing autosomal imprinted genes. *Nature* 415, 810–813.
- 1488 Smith, M.A., Gesell, T., Stadler, P.F., and Mattick, J.S. (2013). Widespread  
1489 purifying selection on RNA structure in mammals. *Nucleic Acids Res.* 41,  
1490 8220–8236.



1491 Smits, G., Mungall, A.J., Griffiths-Jones, S., Smith, P., Beury, D., Matthews,  
1492 L., Rogers, J., Pask, A.J., Shaw, G., VandeBerg, J.L., et al. (2008).  
1493 Conservation of the H19 noncoding RNA and H19-IGF2 imprinting  
1494 mechanism in therians. *Nat. Genet.* *40*, 971–976.

1495 Souquere, S., Beauclair, G., Harper, F., Fox, A., and Pierron, G. (2010).  
1496 Highly ordered spatial organization of the structural long noncoding NEAT1  
1497 RNAs within paraspeckle nuclear bodies. *Mol. Biol. Cell* *21*, 4020–4027.

1498 Stavropoulos, N., Lu, N., and Lee, J.T. (2001). A functional role for Tsix  
1499 transcription in blocking Xist RNA accumulation but not in X-chromosome  
1500 choice. *Proc. Natl. Acad. Sci. U. S. A.* *98*, 10232–10237.

1501 St Laurent, G., Shtokalo, D., Tackett, M.R., Yang, Z., Eremina, T.,  
1502 Wahlestedt, C., Urcuqui-Inchima, S., Seilheimer, B., McCaffrey, T.A., and  
1503 Kapranov, P. (2012). Intronic RNAs constitute the major fraction of the non-  
1504 coding RNA in mammalian cells. *BMC Genomics* *13*, 1.

1505 St Laurent, G., Shtokalo, D., Dong, B., Tackett, M.R., Fan, X., Lazorthes, S.,  
1506 Nicolas, E., Sang, N., Triche, T.J., McCaffrey, T.A., et al. (2013). VlincRNAs  
1507 controlled by retroviral elements are a hallmark of pluripotency and cancer.  
1508 *Genome Biol* *14*, R73.

1509 St. Laurent, G., Wahlestedt, C., and Kapranov, P. (2015). The Landscape of  
1510 long noncoding RNA classification. *Trends Genet.* *31*, 239–251.

1511 Struhl, K. (2007). Transcriptional noise and the fidelity of initiation by RNA  
1512 polymerase II. *Nat. Struct. Mol. Biol.* *14*, 103–105.

1513 Sunwoo, H., Dinger, M.E., Wilusz, J.E., Amaral, P.P., Mattick, J.S., and  
1514 Spector, D.L. (2008). MEN / nuclear-retained non-coding RNAs are up-  
1515 regulated upon muscle differentiation and are essential components of  
1516 paraspeckles. *Genome Res.* *19*, 347–359.

1517 Stavropoulos, N., Lu, N., and Lee, J.T. (2001). A functional role for Tsix  
1518 transcription in blocking Xist RNA accumulation but not in X-chromosome  
1519 choice. *Proc. Natl. Acad. Sci. U. S. A.* *98*, 10232–10237.

1520 Taft, R.J., Pheasant, M., and Mattick, J.S. (2007). The relationship between  
1521 non-protein-coding DNA and eukaryotic complexity. *BioEssays* *29*, 288–299.

1522 Taft, R.J., Kaplan, C.D., Simons, C., and Mattick, J.S. (2009). Evolution,  
1523 biogenesis and function of promoter-associated RNAs. *Cell Cycle Georget.*  
1524 *Tex* *8*, 2332–2338.

1525 Taft, R.J., Pang, K.C., Mercer, T.R., Dinger, M., and Mattick, J.S. (2010). Non-  
1526 coding RNAs: regulators of disease. *J. Pathol.* *220*, 126–139.

1527 Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K.,  
1528 and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult  
1529 human fibroblasts by defined factors. *Cell* *131*, 861–872.

- 1530 Tay, Y., Rinn, J., and Pandolfi, P.P. (2014). The multilayered complexity of  
1531 ceRNA crosstalk and competition. *Nature* 505, 344–352.
- 1532 Thakur, N., Tiwari, V.K., Thomassin, H., Pandey, R.R., Kanduri, M., Göndör,  
1533 A., Grange, T., Ohlsson, R., and Kanduri, C. (2004). An antisense RNA  
1534 regulates the bidirectional silencing property of the *Kcnq1* imprinting control  
1535 region. *Mol. Cell. Biol.* 24, 7855–7862.
- 1536 Tochitani, S., and Hayashizaki, Y. (2008). *Nkx2.2* antisense RNA  
1537 overexpression enhanced oligodendrocytic differentiation. *Biochem. Biophys.*  
1538 *Res. Commun.* 372, 691–696.
- 1539 Tong, X.-H., Xu, B., Zhang, Y.-W., Liu, Y.-S., and Ma, C.-H. (2014). Research  
1540 resources: comparative microRNA profiles in human *corona radiata* cells and  
1541 cumulus oophorus cells detected by next-generation small RNA sequencing.  
1542 *PLoS One* 9, e106706.
- 1543 Trimarchi, T., Bilal, E., Ntziachristos, P., Fabbri, G., Dalla-Favera, R., Tsirigos,  
1544 A., and Aifantis, I. (2014). Genome-wide Mapping and Characterization of  
1545 Notch-Regulated Long Noncoding RNAs in Acute Leukemia. *Cell* 158, 593–  
1546 606.
- 1547 Tripathi, V., Ellis, J.D., Shen, Z., Song, D.Y., Pan, Q., Watt, A.T., Freier, S.M.,  
1548 Bennett, C.F., Sharma, A., Bubulya, P.A., et al. (2010). The Nuclear-Retained  
1549 Noncoding RNA MALAT1 Regulates Alternative Splicing by Modulating SR  
1550 Splicing Factor Phosphorylation. *Mol. Cell* 39, 925–938.
- 1551 Tripathi, V., Shen, Z., Chakraborty, A., Giri, S., Freier, S.M., Wu, X., Zhang,  
1552 Y., Gorospe, M., Prasanth, S.G., Lal, A., et al. (2013). Long noncoding RNA  
1553 MALAT1 controls cell cycle progression by regulating the expression of  
1554 oncogenic transcription factor B-MYB. *PLoS Genet.* 9, e1003368.
- 1555 Tsai, M.-C., Manor, O., Wan, Y., Mosammamaparast, N., Wang, J.K., Lan, F.,  
1556 Shi, Y., Segal, E., and Chang, H.Y. (2010). Long noncoding RNA as modular  
1557 scaffold of histone modification complexes. *Science* 329, 689–693.
- 1558 Tuck, A.C., and Tollervey, D. (2013). A Transcriptome-wide Atlas of RNP  
1559 Composition Reveals Diverse Classes of mRNAs and lncRNAs. *Cell* 154,  
1560 996–1009.
- 1561 Ulitsky, I., and Bartel, D.P. (2013). lincRNAs: Genomics, Evolution, and  
1562 Mechanisms. *Cell* 154, 26–46.
- 1563 Ulitsky, I., Shkumatava, A., Jan, C.H., Sive, H., and Bartel, D.P. (2011).  
1564 Conserved Function of lincRNAs in Vertebrate Embryonic Development  
1565 despite Rapid Sequence Evolution. *Cell* 147, 1537–1550.
- 1566 Vallot, C., Huret, C., Lesecque, Y., Resch, A., Oudrhiri, N., Bennaceur-  
1567 Griscelli, A., Duret, L., and Rougeulle, C. (2013). XACT, a long noncoding  
1568 transcript coating the active X chromosome in human pluripotent cells. *Nat.*  
1569 *Genet.* 45, 239–241.

1570 Wagner, L.A., Christensen, C.J., Dunn, D.M., Spangrude, G.J., Georgelas, A.,  
1571 Kelley, L., Esplin, M.S., Weiss, R.B., and Gleich, G.J. (2007). EGO, a novel,  
1572 noncoding RNA gene, regulates eosinophil granule protein transcript  
1573 expression. *Blood* 109, 5191–5198.

1574 Wahlestedt, C. (2013). Targeting long non-coding RNA to therapeutically  
1575 upregulate gene expression. *Nat. Rev. Drug Discov.* 12, 433–446.

1576 Wan, Y., Qu, K., Zhang, Q.C., Flynn, R.A., Manor, O., Ouyang, Z., Zhang, J.,  
1577 Spitale, R.C., Snyder, M.P., Segal, E., et al. (2014). Landscape and variation  
1578 of RNA secondary structure across the human transcriptome. *Nature* 505,  
1579 706–709.

1580 Wang, K.C., and Chang, H.Y. (2011). Molecular Mechanisms of Long  
1581 Noncoding RNAs. *Mol. Cell* 43, 904–914.

1582 Wang, K.C., Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen,  
1583 Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A., et al. (2011). A long  
1584 noncoding RNA maintains active chromatin to coordinate homeotic gene  
1585 expression. *Nature* 472, 120–124.

1586 Wang, Y., Xu, Z., Jiang, J., Xu, C., Kang, J., Xiao, L., Wu, M., Xiong, J., Guo,  
1587 X., and Liu, H. (2013). Endogenous miRNA Sponge lincRNA-RoR Regulates  
1588 Oct4, Nanog, and Sox2 in Human Embryonic Stem Cell Self-Renewal. *Dev.*  
1589 *Cell* 25, 69–80.

1590 Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool  
1591 for transcriptomics. *Nat. Rev. Genet.* 10, 57–63.

1592 Warner, J.R., Soeiro, R., Birnboim, H.C., Girard, M., and Darnell, J.E. (1966).  
1593 Rapidly labeled HeLa cell nuclear RNA. I. Identification by zone sedimentation  
1594 of a heterogeneous fraction separate from ribosomal precursor RNA. *J. Mol.*  
1595 *Biol.* 19, 349–361.

1596 West, J.A., Davis, C.P., Sunwoo, H., Simon, M.D., Sadreyev, R.I., Wang, P.I.,  
1597 Tolstorukov, M.Y., and Kingston, R.E. (2014). The Long Noncoding RNAs  
1598 NEAT1 and MALAT1 Bind Active Chromatin Sites. *Mol. Cell* 55, 791–802.

1599 Willingham, A.T., Orth, A.P., Batalov, S., Peters, E.C., Wen, B.G., Aza-Blanc,  
1600 P., Hogenesch, J.B., and Schultz, P.G. (2005). A strategy for probing the  
1601 function of noncoding RNAs finds a repressor of NFAT. *Science* 309, 1570–  
1602 1573.

1603 Wojdacz, T.K., Dobrovic, A., and Algar, E.M. (2008). Rapid detection of  
1604 methylation change at H19 in human imprinting disorders using methylation-  
1605 sensitive high-resolution melting. *Hum. Mutat.* 29, 1255–1260.

1606 Wood, E.J., Chin-Inmanu, K., Jia, H., and Lipovich, L. (2013). Sense-  
1607 antisense gene pairs: sequence, transcription, and structure are not  
1608 conserved between human and mouse. *Front. Genet.* 4.

- 1609 Xie, M., Sun, M., Zhu, Y., Xia, R., Liu, Y., Ding, J., Ma, H., He, X., Zhang, Z.,  
1610 Liu, Z., et al. (2015). Long noncoding RNA HOXA-AS2 promotes gastric  
1611 cancer proliferation by epigenetically silencing P21/PLK3/DDIT3 expression.  
1612 *Oncotarget* 6, 33587–33601.
- 1613 Xu, N., Papagiannakopoulos, T., Pan, G., Thomson, J.A., and Kosik, K.S.  
1614 (2009). MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses  
1615 pluripotency in human embryonic stem cells. *Cell* 137, 647–658.
- 1616 Xu, X.-F., Li, J., Cao, Y.-X., Chen, D.-W., Zhang, Z.-G., He, X.-J., Ji, D.-M.,  
1617 and Chen, B.-L. (2015). Differential Expression of Long Noncoding RNAs in  
1618 Human Cumulus Cells Related to Embryo Developmental Potential: A  
1619 Microarray Analysis. *Reprod. Sci. Thousand Oaks Calif* 22, 672–678.
- 1620 Xue, Z., Huang, K., Cai, C., Cai, L., Jiang, C., Feng, Y., Liu, Z., Zeng, Q.,  
1621 Cheng, L., Sun, Y.E., et al. (2013). Genetic programs in human and mouse  
1622 early embryos revealed by single-cell RNA sequencing. *Nature* 500, 593–597.
- 1623 Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng,  
1624 X., Yan, J., et al. (2013). Single-cell RNA-Seq profiling of human  
1625 preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* 20,  
1626 1131–1139.
- 1627 Yang, F., Zhang, L., Huo, X., Yuan, J., Xu, D., Yuan, S., Zhu, N., Zhou, W.,  
1628 Yang, G., Wang, Y., et al. (2011a). Long noncoding RNA high expression in  
1629 hepatocellular carcinoma facilitates tumor growth through enhancer of zeste  
1630 homolog 2 in humans. *Hepatology* 54, 1679–1689.
- 1631 Yang, L., Lin, C., Liu, W., Zhang, J., Ohgi, K.A., Grinstead, J.D., Dorrestein,  
1632 P.C., and Rosenfeld, M.G. (2011b). ncRNA- and Pc2 methylation-dependent  
1633 gene relocation between nuclear structures mediates gene activation  
1634 programs. *Cell* 147, 773–788.
- 1635 Yang, Y.W., Flynn, R.A., Chen, Y., Qu, K., Wan, B., Wang, K.C., Lei, M., and  
1636 Chang, H.Y. (2014). Essential role of lncRNA binding for WDR5 maintenance  
1637 of active chromatin and embryonic stem cell pluripotency. *Elife* 3, e02046.
- 1638 Yap, K.L., Li, S., Muñoz-Cabello, A.M., Raguz, S., Zeng, L., Mujtaba, S., Gil,  
1639 J., Walsh, M.J., and Zhou, M.-M. (2010). Molecular Interplay of the Noncoding  
1640 RNA ANRIL and Methylated Histone H3 Lysine 27 by Polycomb CBX7 in  
1641 Transcriptional Silencing of INK4a. *Mol. Cell* 38, 662–674.
- 1642 Yerushalmi, G.M., Salmon-Divon, M., Yung, Y., Maman, E., Kedem, A., Ophir,  
1643 L., Elemento, O., Coticchio, G., Dal Canto, M., Mignini Renzinu, M., et al.  
1644 (2014). Characterization of the human cumulus cell transcriptome during final  
1645 follicular maturation and ovulation. *Mol. Hum. Reprod.* 20, 719–735.
- 1646 Yin, Q.-F., Yang, L., Zhang, Y., Xiang, J.-F., Wu, Y.-W., Carmichael, G.G.,  
1647 and Chen, L.-L. (2012). Long noncoding RNAs with snoRNA ends. *Mol. Cell*  
1648 48, 219–230.

- 1649 Yotova, I.Y., Vlatkovic, I.M., Pauler, F.M., Warczok, K.E., Ambros, P.F.,  
1650 Oshimura, M., Theussl, H.-C., Gessler, M., Wagner, E.F., and Barlow, D.P.  
1651 (2008). Identification of the human homolog of the imprinted mouse Air non-  
1652 coding RNA. *Genomics* 92, 464–473.
- 1653 Young, T.L., Matsuda, T., and Cepko, C.L. (2005). The noncoding RNA  
1654 taurine upregulated gene 1 is required for differentiation of the murine retina.  
1655 *Curr. Biol. CB* 15, 501–512.
- 1656 Yu, W., Gius, D., Onyango, P., Muldoon-Jacobs, K., Karp, J., Feinberg, A.P.,  
1657 and Cui, H. (2008). Epigenetic silencing of tumour suppressor gene p15 by its  
1658 antisense RNA. *Nature* 451, 202–206.
- 1659 Zhang, A., Zhou, N., Huang, J., Liu, Q., Fukuda, K., Ma, D., Lu, Z., Bai, C.,  
1660 Watabe, K., and Mo, Y.-Y. (2013). The human long non-coding RNA-RoR is a  
1661 p53 repressor in response to DNA damage. *Cell Res.* 23, 340–350.
- 1662 Zhang, B., Arun, G., Mao, Y.S., Lazar, Z., Hung, G., Bhattacharjee, G., Xiao,  
1663 X., Booth, C.J., Wu, J., Zhang, C., et al. (2012). The lncRNA Malat1 Is  
1664 Dispensable for Mouse Development but Its Transcription Plays a cis-  
1665 Regulatory Role in the Adult. *Cell Rep.* 2, 111–123.
- 1666 Zhang, K., Huang, K., Luo, Y., and Li, S. (2014). Identification and functional  
1667 analysis of long non-coding RNAs in mouse cleavage stage embryonic  
1668 development based on single cell transcriptome data. *BMC Genomics* 15,  
1669 845.
- 1670 Zhang, J.-Y., Weng, M.-Z., Song, F.-B., Xu, Y.-G., Liu, Q., Wu, J.-Y., Qin, J.,  
1671 Jin, T., and Xu, J.-M. (2016). Long noncoding RNA AFAP1-AS1 indicates a  
1672 poor prognosis of hepatocellular carcinoma and promotes cell proliferation  
1673 and invasion via upregulation of the RhoA/Rac2 signaling. *Int. J. Oncol.* 48,  
1674 1590–1598.
- 1675 Zhang, X., Lian, Z., Padden, C., Gerstein, M.B., Rozowsky, J., Snyder, M.,  
1676 Gingeras, T.R., Kapranov, P., Weissman, S.M., and Newburger, P.E. (2009).  
1677 A myelopoiesis-associated regulatory intergenic noncoding RNA transcript  
1678 within the human HOXA cluster. *Blood* 113, 2526–2534.
- 1679 Zhang, X., Weissman, S.M., and Newburger, P.E. (2014). Long intergenic  
1680 non-coding RNA HOTAIRM1 regulates cell cycle progression during myeloid  
1681 maturation in NB4 human promyelocytic leukemia cells. *RNA Biol.* 11, 777–  
1682 787.
- 1683 Zhao, J., Ohsumi, T.K., Kung, J.T., Ogawa, Y., Grau, D.J., Sarma, K., Song,  
1684 J.J., Kingston, R.E., Borowsky, M., and Lee, J.T. (2010). Genome-wide  
1685 identification of polycomb-associated RNAs by RIP-seq. *Mol. Cell* 40, 939–  
1686 953.
- 1687 Zhou, X., Liu, S., Cai, G., Kong, L., Zhang, T., Ren, Y., Wu, Y., Mei, M.,  
1688 Zhang, L., and Wang, X. (2015). Long Non Coding RNA MALAT1 Promotes

1689 Tumor Growth and Metastasis by inducing Epithelial-Mesenchymal Transition  
1690 in Oral Squamous Cell Carcinoma. *Sci. Rep.* 5, 15972.

1691 Zhou, X.-L., Wang, W.-W., Zhu, W.-G., Yu, C.-H., Tao, G.-Z., Wu, Q.-Q.,  
1692 Song, Y.-Q., Pan, P., and Tong, Y.-S. (2016). High expression of long non-  
1693 coding RNA AFAP1-AS1 predicts chemoradioresistance and poor prognosis  
1694 in patients with esophageal squamous cell carcinoma treated with definitive  
1695 chemoradiotherapy. *Mol. Carcinog.*

1696 Zhu, L., Zhang, Y., Zhang, W., Yang, S., Chen, J.-Q., and Tian, D. (2009).  
1697 Patterns of exon-intron architecture variation of genes in eukaryotic genomes.  
1698 *BMC Genomics* 10, 47.

1699 Zhou, Y., Zhong, Y., Wang, Y., Zhang, X., Batista, D.L., Gejman, R., Ansell,  
1700 P.J., Zhao, J., Weng, C., and Klibanski, A. (2007). Activation of p53 by MEG3  
1701 non-coding RNA. *J. Biol. Chem.* 282, 24731–24742.

1702

1703

1704

1705

1706 **FIGURES LEGENDS**

1707 **Figure 1. General classification of lncRNAs and mechanisms of action.**

1708 (A) Sense lncRNAs overlap with a protein-coding gene on the same DNA  
1709 strand and antisense lncRNAs overlap on the opposite strand. Bidirectional  
1710 lncRNAs are transcripts in which the transcription start site is located at less  
1711 than 1kb away from the neighboring protein-coding gene. Intronic lncRNAs  
1712 are located within an intron of a protein-coding gene, while intergenic  
1713 lncRNAs (or lincRNAs) are located more than 5kb away from a protein-coding  
1714 region. Intragenic lncRNAs include sense and antisense overlapping,  
1715 bidirectional and intronic lncRNAs.

1716 (B) lncRNAs can act via different mechanisms of actions:

1717 **Scaffold:** lncRNAs can serve as adaptors by bringing together multiple  
1718 components, such as proteins, RNAs or DNAs, to form ribonucleoprotein  
1719 complexes (examples: HOTAIR, ANRIL, MALAT1, TUG1, NEAT1).

1720 **Guide:** via standard base pairing, lncRNAs can guide ribonucleoprotein  
1721 complexes to a specific genomic location (examples: HOTTIP, XIST, AIR,  
1722 KCNQ1ot1, lincRNA-p21, HOTTIP).

1723 **Molecular decoy:** lncRNAs can bind to protein complexes and prevent their  
1724 interaction with their natural targets (examples: GAS5, PANDA, NEAT1).

1725 **miRNA sponge:** ceRNAs can bind to and sequester miRNA, leading to the  
1726 active transcription of their mRNA targets (examples: MD1, H19, ROR).

1727 **Enhancer:** eRNAs are derived from enhancer regions and regulate  
1728 neighboring gene expression in cis and trans, being often implicated in  
1729 creating chromatin loops for stabilization and activation of the promoter region  
1730 of their target gene (examples: lncRNA-CSR, HOTTIP).

1731 **Figure 2: LncRNA expression in early human embryos, embryonic stem**  
1732 **cells (ESC), foetal and adult tissues and cancer samples.**

1733 (A) Pie charts representing the proportion of lncRNAs and protein-coding  
1734 genes at each stage of human early embryonic development. Embryonic  
1735 genome activation (EGA) takes place at the 4 to 8-cell stage and is  
1736 accompanied by a shift in the proportion of these three RNA categories.  
1737 During and after EGA, lncRNAs become predominant over protein-coding  
1738 genes. (B) Coverage plots of RNA-seq reads of selected lncRNA genes that  
1739 are differentially expressed in early embryos and foetal/adult tissues, and in  
1740 early embryos and cancer tissues. Emb: Embryo; Foe: foetal; Ad: adult.

1741 BCAR4 is detected specifically during early development in pre-implantation  
1742 embryos, PTENP1 in early embryos and also in foetal/adult tissues, while  
1743 COLCA1 is expressed only in foetal/adult tissues. AFAP1-AS1, HOTAIR,  
1744 AACSP1, LUCAT1 are all expressed during embryo development and in  
1745 cancer samples, whereas they are absent in foetal/adult tissues.

1746

## 1747 **TABLES**

1748 **Table 1: Examples of lncRNAs to illustrate their variety of mechanisms**  
1749 **of action.**

1750 **Table 2: Examples of lncRNAs with demonstrated functionality.**

1751

## 1752 **Supplemental Tables**

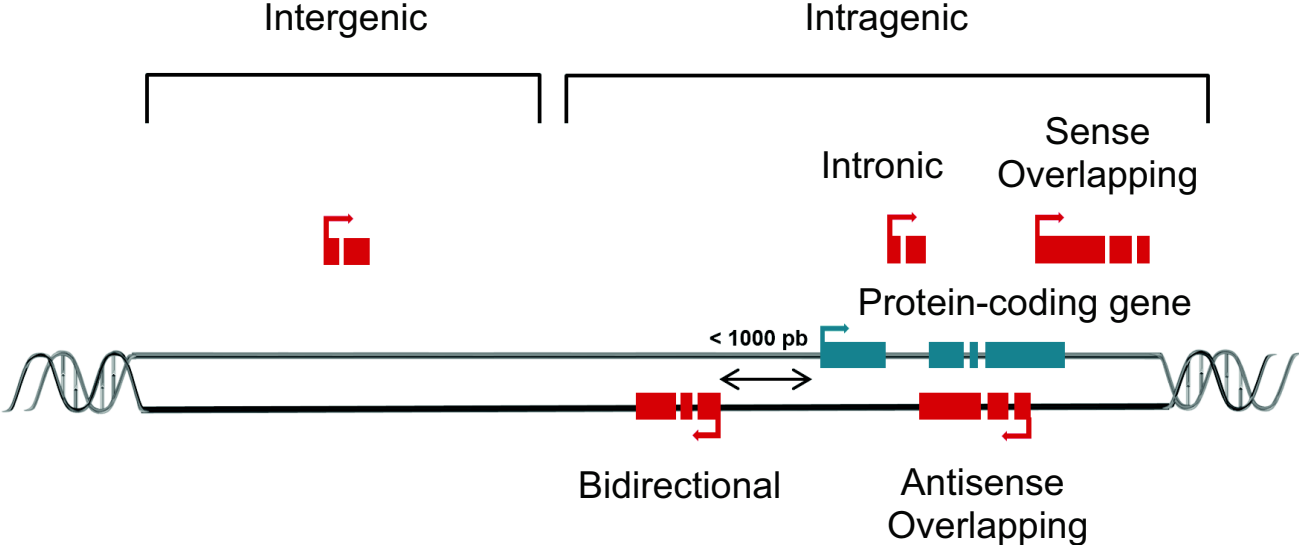
1753 **Table S1: Samples included in the mRNAseq data meta-analysis**

1754

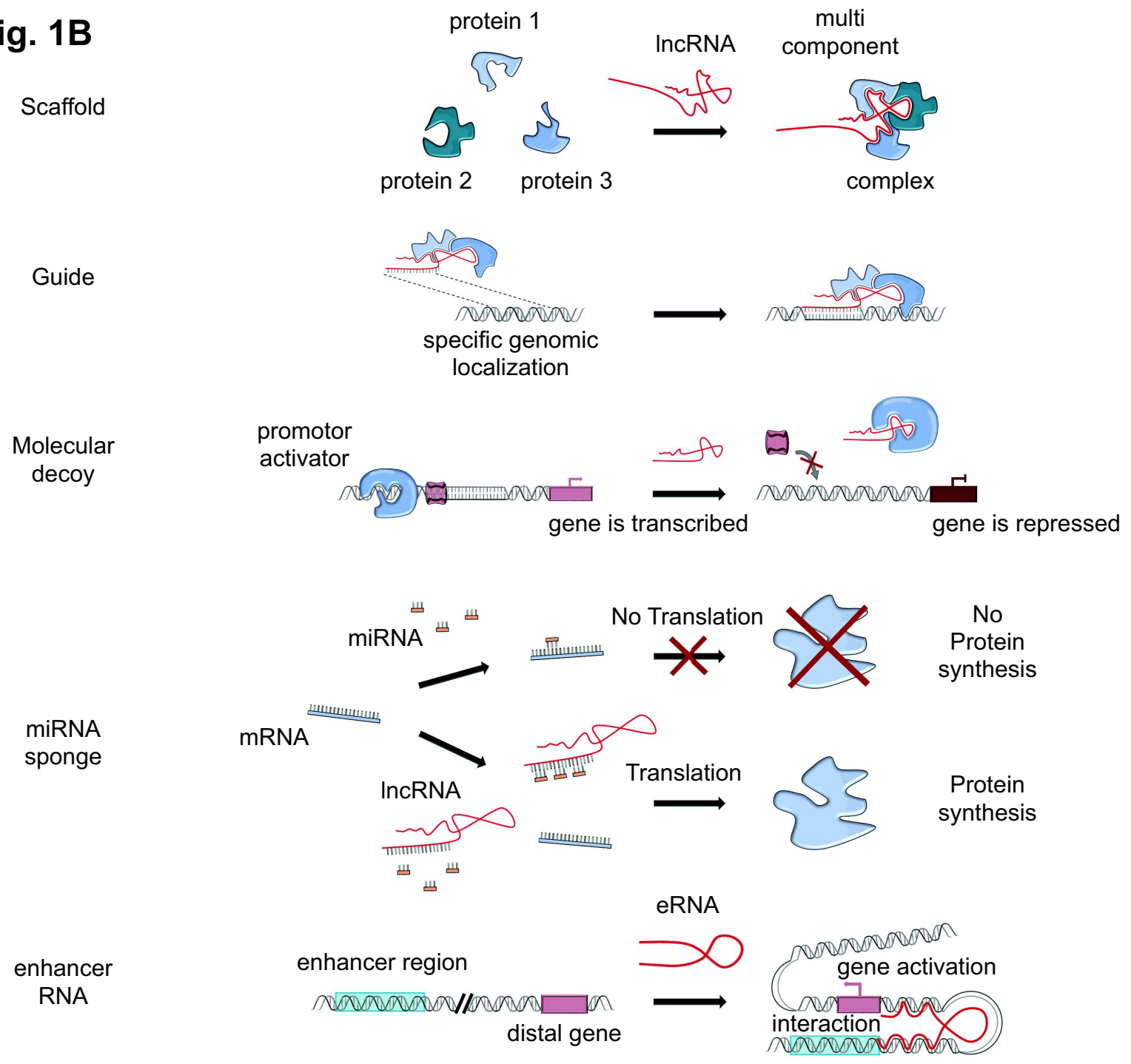
1755



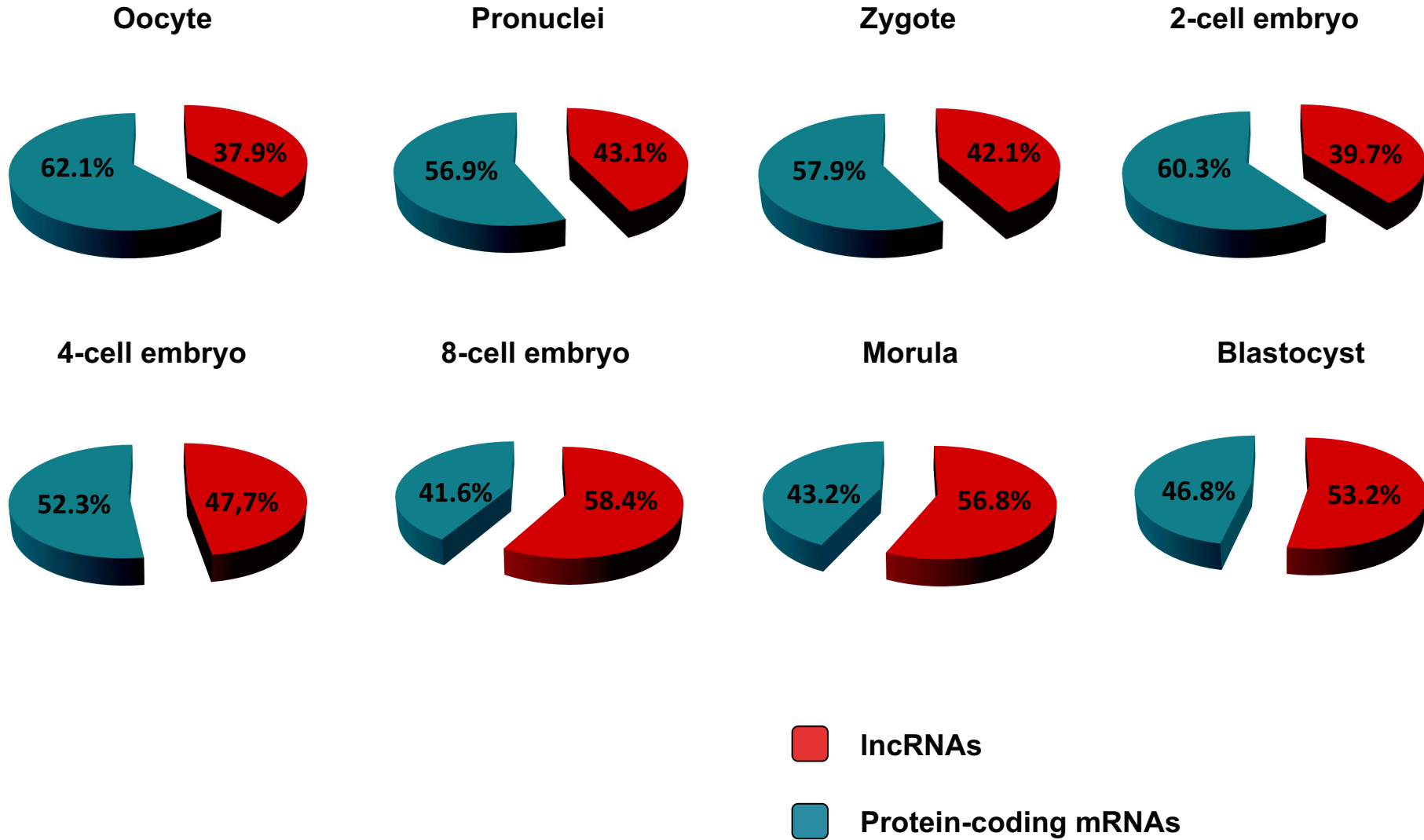
Fig. 1A



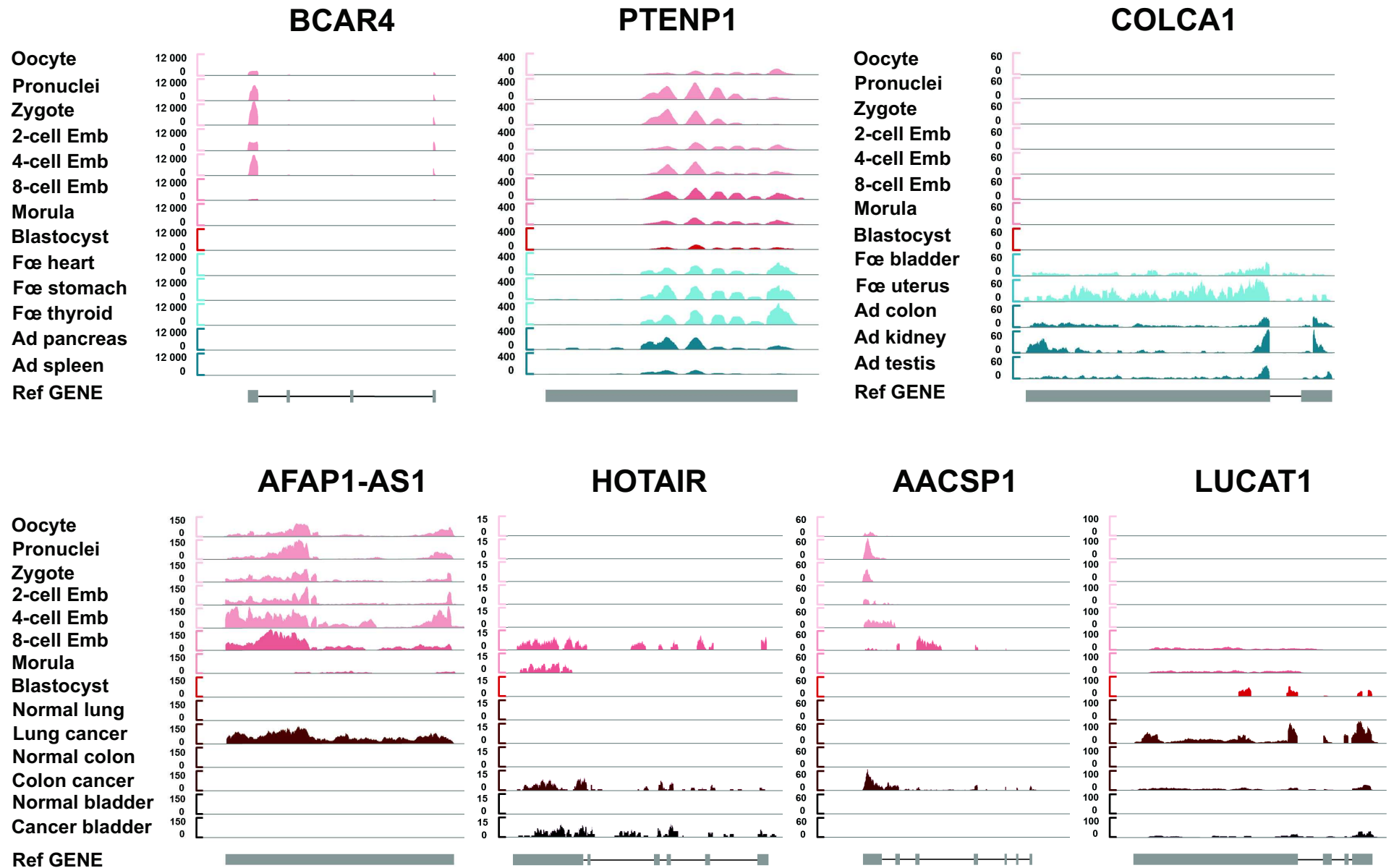
**Fig. 1B**



**Fig. 2A**



**Fig. 2B**



# **PARTIE 3 : La Différenciation**

Les CSP sont capables de se différencier en tout type cellulaire. Cette propriété fondamentale des CSP peut s'observer aussi bien *in vivo* qu'*in vitro*. De plus, il existe deux stratégies pour obtenir la différenciation des CSP : laisser cette différenciation se réaliser au hasard, selon un patron chaotique (« différenciation non dirigée »), ou bien intervenir à l'aide de facteurs de signalisation en combinaisons et doses variables au cours du temps (« différenciation dirigée »). Ces différentes méthodes de différenciation seront reprises dans les chapitres suivants et sont résumés dans le tableau ci-dessous.

<b>DIFFERENCIATION</b>	<b>IN VIVO</b>	<b>IN VITRO</b>
<b>DIRIGEE</b>	(Embryon*)	Différenciation dirigée <i>in vitro</i>
<b>NON DIRIGEE</b>	Tératome	Corps Embryoïdes

**Tableau 6 : Récapitulatif de différents grands types de différenciation.**

\* : le développement embryonnaire est par définition une différenciation *in vivo* d'un groupe de cellules souches pluripotentes (la masse cellulaire interne) sous l'influence de facteurs de croissances précis, à doses définies et dans une fenêtre temporelle donnée : il s'agit bien d'une différenciation dirigée.

## 1. Différenciation non dirigée

La différenciation non dirigée est une différenciation anarchique. Cela implique que les tissus qui en découlent ne sont pas organisés correctement par opposition au développement embryonnaire normal.

## 1.1 Le corps embryoïde

Les corps embryoïdes (embryoid bodies ou EB) sont un modèle de différenciation non dirigée *in vitro*. L'avantage majeur de ce type de différenciation est qu'elle est en trois dimensions et qu'elle permet d'appréhender les interactions des tissus entre eux.

Il existe plusieurs protocoles pour former des EB. La plupart de ces protocoles utilisent une différenciation spontanée des cellules souches pluripotentes en l'absence de cytokines. Hiroshi Kurosawa les a classé en cinq catégories (Kurosawa, 2007) : (1) la suspension de la culture cellulaire en un puits de grade bactériologique (c'est à dire sans adhérence), (2) la culture en méthylcellulose, (3) la culture en goutte suspendue, (4) culture dans un puits à faible adhérence et (5) les bioréacteurs (Tableau 7).

Suspension cellulaire de CSP	Puits de grade bactériologique	<ul style="list-style-type: none"> <li>- méthode basique</li> <li>- peu d'homogénéité dans la morphologie et dans la différenciation</li> </ul>
	Méthylcellulose	<ul style="list-style-type: none"> <li>- généralement utilisée pour la différenciation hématopoïétique</li> <li>- pas facile d'utilisation</li> </ul>
	Goutte suspendue	<ul style="list-style-type: none"> <li>- EB de taille contrôlée</li> <li>- volume de la goutte &lt; 50µl</li> </ul>
	Puits à faible adhérence	<ul style="list-style-type: none"> <li>- EB de taille contrôlée</li> <li>- possibilité de changer le</li> </ul>

		milieu  - culture comparable à la méthode « goutte suspendue »
	Bioréacteur	- possibilité de production d'EB à grande échelle  - possibilité de contrôler la stratégie de production  - possibilité de former directement des EB

Tableau 7 : Les protocoles de formation d'EB (adapté de (Kurosawa, 2007)).

Cette différenciation permet d'obtenir des tissus issus des trois feuillets embryonnaires mais de façon désorganisée. L'absence de vascularisation des EB bride le processus de différenciation en limitant la taille des structures obtenues.

## 1.2 Le tératome

L'étymologie du mot tératome signifie « tumeur monstrueuse » (issu du grec *teratos* et du suffixe *-ome*). L'académie de médecine définit le tératome comme suit : « Formation tumorale constituée d'une variété de cellules ou de tissus par définition étrangers à la région qui l'abrite mais provenant du développement, à des stades variés, y compris jusqu'au stade adulte (tératome adulte ou mature), d'un ou de plusieurs des trois feuillets



embryonnaires (tératome uni- ou pluritissulaire), ce qui explique son polymorphisme histologique; mais il existe presque toujours une composante neuro-épithéliale ».

### **1.2.1 Tératomes et tératocarcinomes en clinique**

Les tératomes sont dérivés des cellules germinales primordiales (CGP). Ce sont des tumeurs non séminomateuses. Elles peuvent être gonadiques (testiculaires ou ovariens) dans 80% des cas, ou prendre naissance dans d'autres territoires le long de l'axe médian du corps (médiastin, rétro-péritoine, hypophyse, sacro-coccygiens). Les tératomes sacro-coccygiens peuvent être endo- ou exo-pelvien, en particulier au niveau du médiastin. Ils sont plus fréquents chez les nouveau-nés. Ils s'observent toutes les 20 000 à 70 000 naissances avec une proportion trois à quatre fois plus importante chez les filles. Autrefois, on parlait de jumeau absorbé à l'état fœtal, ou d'embryon égaré. Les tératomes sont des tumeurs correspondant à un patchwork de tissus histologiques différenciés issus des trois feuilletts (ectoderme, mésoderme, endoderme) allant du stade embryonnaire jusqu'au stade adulte. Selon le degré de différenciation des tissus qui composent le tératome, on distingue les tératomes matures, les tératomes immatures et les tératocarcinomes (Damjanov and Andrews, 2007). Ils peuvent être kystiques ou solides.

Les tératomes matures sont les formes bénignes ou totalement différenciées. Ils sont constitués exclusivement de tissus bien différenciés de type adulte. Cependant, la composante ectodermique est généralement prépondérante, c'est à dire la peau, le tissu sébacé et le tissu neuro-épithélial (Lewis et al., 1983). Ils sont très souvent asymptomatiques, leur proportion semble donc être largement sous-estimée. Ils sont le plus fréquemment localisés au niveau du médiastin (95% antérieur, 5% postérieur) lorsqu'ils sont extragonadiques (Moeller et al., 1997). Le traitement est exclusivement chirurgical par l'exérèse de la tumeur.

Les tératomes immatures ou malins sont parfois appelés tératocarcinomes. En plus d'être constitués de tissus différenciés, ils sont aussi composés de cellules souches

indifférenciées. Leur degré de malignité est généralement en fonction inverse du degré de différenciation tissulaire. Ils sont donc composés de tissus issus des trois feuillets parfaitement différenciés ainsi que des tissus embryonnaires et peuvent être capables de produire de l'alpha-fœto protéine.

Les cellules à l'origine des tératocarcinomes sont les cellules de carcinome embryonnaire (EC). Il est établi que les EC sont le pendant malin des ES. En plusieurs points, les EC ressemblent aux ES notamment par la présence de marqueurs de pluripotence, tel que SSEA3, la conservation de la propriété « cellule souche » au cours de la culture *in vitro* et la possibilité de former des corps embryoïdes *in vitro* et des tératomes *in vivo* chez l'animal immunosupprimé (Cooke et al., 2006; Yang et al., 2012). Cependant les EC et les ES diffèrent. Non seulement, les EC présentent des marqueurs membranaires de différenciation tels que brachyury et VE-cadhérine, mais également le profil d'expression des gènes diffèrent (Yang et al., 2012).

L'étude des tératomes et tératocarcinomes spontanés chez l'homme, et la correspondance entre les EC et les ES amènent à utiliser les tératomes dans l'animal immunodéficient comme un outil dans la sécurité et la qualité du développement des thérapies cellulaires issues des CSP (Yang et al., 2012).

### **1.2.2 Tératomes dans l'animal immunosupprimé**

Le tératome est l'outil de référence afin de vérifier la pluripotence des cellules pluripotentes, iPS ou ES. Pour cela, les CSP sont introduites dans un animal fortement immunodéprimé, ce qui empêche le système immunitaire de l'animal de détruire les cellules pluripotentes. Après quelques semaines, des tératomes apparaissent chez l'animal. Il s'agit de la différenciation anarchique, mais *in vivo*, des cellules pluripotentes en un patchwork de tissus issus des trois feuillets embryonnaires.

Il existe plusieurs types d'animaux immunodéprimés, des chevaux, des chiens et des souris. Les tératomes réalisés pour évaluer la pluripotence se font dans des souris. Au cours des dernières décennies, les souris immunodéprimées disponibles ont évoluées considérablement au niveau de leur patrimoine génétique (Shultz et al., 2007). Il s'agit d'organismes génétiquement modifiés, et ceux utilisés avec le plus de succès pour former des tératomes sont des souris NOD – SCID – Gamma (NOD.Cg-Prkdc<sup>scid</sup> Il2rg<sup>tm1Wjl</sup>/SzJ, NSG). Ce qui signifie que ces souris sont diabétiques non obèses (NOD) - entraînant une altération de l'immunité innée pour les cellules NK (natural killer), l'immunodéficiences sévère combinée (SCID) - caractérisée par une déficience du système immunitaire adaptatif via la mutation du gène PRKCD conduisant à une réponse immunitaire inadaptée des lymphocytes T et B, et enfin la mutation du gène IL2RG entraîne l'absence totale du récepteur gamma de l'interleukine 2 et de ce fait favorise la greffe de cellules primaires humaines (Shultz et al., 2005).

L'utilisation d'animaux à des fins scientifiques répond à une législation très stricte mise à jour en 2013 avec la publication d'une directive européenne (DIRECTIVE 2010/63/UE DU PARLEMENT EUROPÉEN ET DU CONSEIL du 22 septembre 2010) relative à leur protection. Le principe des « trois R » y fait son apparition. Cela regroupe les principes du Remplacement, de la Réduction et du Raffinement. Remplacer les animaux à chaque fois que cela est possible, des outils informatiques performants peuvent substituer à l'expérimentation animale ou d'autres méthodes ne nécessitant pas d'animaux. Réduire au nombre minimal d'animaux nécessaires à une relevance scientifique. Le raffinement est dans l'hébergement, l'élevage et la qualité des soins qui sont apportés aux animaux.

Plusieurs localisations éthiquement acceptables sont possibles pour la greffe de cellules pluripotentes pour réaliser des tératomes dans la souris: la capsule du testicule, la capsule du rein, la cavité intra péritonéale, le genou, le tissu sous cutanés, etc. De nombreuses études s'accordent sur le fait que le site d'implantation des cellules souches pluripotentes influence la formation du tératome (Cooke et al., 2006; Prokhorova et al., 2009). Outre la localisation de la greffe, d'autres facteurs influencent la formation de tératomes tel que la lignée de CSP injectée (Gutierrez-Aranda et al., 2010; Narsinh et al.,

2011), ou le nombre de cellules injectées, en effet le seuil critique de cellules nécessaire à la formation du tératome est différent selon le site d'injection (Cao et al., 2007; Lee et al., 2009a). De plus, lorsque les cellules pluripotentes sont injectées en sous-cutané, la présence de Matrigel® permet d'augmenter considérablement la réussite de formation de tératomes (Prokhorova et al., 2009). En effet, cette combinaison permet 94% de chance de formation de tératomes contre seulement 33% avec une injection sous-cutané de cellules sans Matrigel®, 12,5% pour une injection en intramusculaire.

### **1.2.3 Tissus obtenus dans les tératomes expérimentaux**

Les cellules ES humaines injectées dans des souris immunodéprimées se différencient en tissus des trois feuillets embryonnaires (ectoderme, le mésoderme et l'endoderme) parfaitement identifiables. Contrairement à l'embryon, le tératome ressemble plus à un patchwork de tissus différenciés. En effet, c'est au cours de la gastrulation que le disque embryonnaire donne trois feuillets distincts qui se différencieront en tissus puis organes tout au long de l'embryogénèse puis de l'organogénèse. L'architecture de la gastrulation n'est pas reproduite dans le tératome.

#### **1.2.3.1 Ectoderme ou ectoblaste**

L'ectoderme est tiré du grec « ektos » qui signifie « en dehors » et de « derma » pour « la peau » ou « blastos » pour « bourgeon ». Il représente le feuillet embryonnaire externe de l'embryon tridermique.

Dans le développement embryonnaire normal, il se compose d'une couche cellulaire issues de l'épiblaste et qui divise en deux : l'ectoderme a proprement parler et la plaque neurale. Cette dernière s'épaissira, s'invaginera et migrera, c'est neurulation, pour former dès la fin du premier mois le tube neural. Le développement ultérieur de ce tube donnera le système nerveux central : l'encéphale, le cervelet, les neurones, les cellules gliales et la

moelle épinière. Quant à l'ectoderme, il donnera les placodes sensibles (la placode olfactive qui donnera l'épithélium olfactif, la placode auditive qui donnera l'oreille interne et la placode oculaire qui donnera le cristallin et la cornée), ainsi que la cavité buccale, le canal anal, l'épiderme, les poils, les glandes sébacées et les glandes sudoripares.

Dans le tératome expérimental, comme pour le tératome pathologique, on ne retrouve jamais de tissu cardiaque. Bien que le développement ne soit pas assez ordonné ni long dans le temps (maximum deux mois pour un tératome dans la souris immunodéprimée) pour permettre d'obtenir des organes sensoriels complets, on trouve néanmoins des tissus de la plupart des organes.

### **1.2.3.2 Mésoderme ou mésoblaste**

Le mésoderme est tiré du grec « méso » qui signifie « milieu » et de « derma » pour « la peau » ou « blastos » pour « bourgeon ». Il représente le feuillet embryonnaire intermédiaire de l'embryon tridermique qui se forme au cours de la gastrulation pendant la troisième semaine de développement.

Il se forme entre l'épiblaste et l'entoblaste lors de la formation de la notochorde, prolifère à côté de la ligne médiane et forme trois structures : le mésoblaste para-axial, le mésoblaste intermédiaire et le mésoblaste latéral. Le premier va subir une métamérisation pour former plusieurs amas ou somites. C'est ce qui va permettre de segmenter l'embryon en métamères avec 35 à 37 paires de somites. Il fournira également le matériel cellulaire du sclérotome pour le squelette axial, du myotome pour la musculature striée du tronc, cou et extrémités et du dermatome pour les tissus sous-cutanés. Le mésoblaste intermédiaire forme, en regard de chaque somite, des néphrotomes qui constitueront l'appareil urinaire. Enfin le mésoblaste latéral se divise en deux lames. L'une ventrale qui constituera la splanchnopleure et l'autre dorsale pour former la somatopleure. Très précocement, au niveau de la partie céphalique de l'embryon, un regroupement de cellules mésoblastiques forme l'aire cardiogénique qui évoluera pour donner le cœur.

Au niveau du tératome expérimental, les structures d'origine mésodermiques les plus souvent retrouvées sont le cartilage et les cellules musculaires.

### **1.2.3.3 Endoderme ou endoblaste**

L'endoderme est tiré du grec « endo » qui signifie « dedans » et de « derma » pour « la peau » ou « blastos » pour « bourgeon ». Il représente le feuillet embryonnaire interne de l'embryon tridermique.

Il est à l'origine des épithélia du tube digestif, à l'exception de la bouche et du rectum, des glandes digestives annexes (le foie, les glandes salivaires, la vésicule biliaire et une partie du pancréas), ainsi que des voies respiratoires.

Au niveau du tératome expérimental, on retrouve de nombreuses formations d'origine endodermique. Après marquage spécifique, il apparaît qu'une même structure épithéliale peut avoir à la fois une origine digestive et respiratoire. Nous approfondirons ce résultat surprenant au cours du chapitre Résultats.

## **2. Différenciation dirigée**

### **2.1 Modéliser des tissus *in vitro***

La différenciation non dirigée ne permet pas d'obtenir un seul type de tissu. En effet, que ce soit dans le tératome ou dans le corps embryoïde, c'est un mélange de tissus qui est obtenu. Au contraire, la différenciation dirigée permet d'obtenir un tissu bien spécifique. Cette homogénéité est recherchée à la fois pour les applications de modélisation *in vitro* et pour les applications médicales.

La différenciation dirigée s'appuie sur des protocoles définis qui visent à mimer au plus près, du point de vue de la signalisation moléculaire, les événements qui se déroulent au cours du développement embryonnaire du tissu d'intérêt. Il s'agit d'activer ou d'inhiber des voies de signalisation par l'ajout de cocktails de cytokines et/ou de molécules chimiques à des concentrations bien établies, selon un calendrier précis. Ces protocoles permettent d'obtenir des tissus des trois feuillets embryonnaires (Tableau 8).

<b>Feuillet embryonnaire</b>	<b>Tissu obtenu à l'issue de la différenciation</b>	<b>Références</b>
<b>Ectoderme</b>	Neurones	(Chambers et al., 2009)
	Epithélium de la rétine pigmentaire	(Kanemura et al., 2013)
	Peau	(Nissan et al., 2011; Shalom-Feuerstein et al., 2013)
<b>Mésoderme</b>	Myocarde	(Moretti et al., 2010)
	Cartilage	(Umeda et al., 2015)
	Tissu hématopoïétique	(Kobari et al., 2012; Smith et al., 2015)
	Endothélium	(Orlova et al., 2014)
	Tissu adipeux	(Mohsen-Kanson et al., 2014)

	Muscle squelettique	(Chal et al., 2015)
<b>Endoderme</b>	Tissu hépatique	(Hannan et al., 2013)
	Cellules $\beta$ du Pancréas	(Rezania et al., 2014)
	Thyroïde	(Kurmann et al., 2015)
	Epithélium bronchique	(Wong et al., 2015)

**Tableau 8 : Liste non exhaustive de protocoles de différenciation dirigée (adapté de (De Vos et al., 2016)).**

La différenciation dirigée est une technique clé pour la modélisation des tissus *in vitro* par les CSP et a fait l'objet d'un article de revue. Il est non seulement possible de modéliser des tissus sains à partir de CSP normales, mais également des tissus malades issus de CSP porteuses d'anomalies génétiques, comme nous le traiterons dans la Partie 4, Chapitre 2.1).

**Article 3 :** « Les cellules souches pluripotentes induites : un nouveau paradigme pour l'étude des tissus humains »

Publié dans *Biol Aujourd'hui* en 2016.



## Les cellules souches pluripotentes induites : un nouveau paradigme pour l'étude des tissus humains

Caroline Sansac<sup>1,2,3</sup>, Saïd Assou<sup>1,2,4</sup>, Julien Bouckenheimer<sup>1,2,3</sup>, Jean-Marc Lemaître<sup>1,2,5</sup>  
et John De Vos<sup>1,2,4,5,6,7</sup>

<sup>1</sup> CHU Montpellier, Institut de Recherche de Médecine Régénératrice et de Biothérapies, Hôpital Saint-Éloi, 80 avenue Augustin Fliche, 34295 Montpellier Cedex 5, France

<sup>2</sup> INSERM, U1183, 34000 Montpellier, France

<sup>3</sup> Université de Montpellier, UFR de Pharmacie, 34000 Montpellier, France

<sup>4</sup> Université de Montpellier, UFR de Médecine, 34000 Montpellier, France

<sup>5</sup> CHU Montpellier, Plateforme iPSCs SAFE-IPS, Institut de Recherche de Médecine Régénératrice et de Biothérapies, 34000 Montpellier, France

<sup>6</sup> Institut de Biologie Computationnelle, 34000 Montpellier, France

<sup>7</sup> CHU Montpellier, Département d'Ingénierie Cellulaire et Tissulaire, Hôpital Saint-Éloi, 34000 Montpellier, France  
Auteur correspondant : John De Vos, [john.devos@inserm.fr](mailto:john.devos@inserm.fr)

Reçu le 3 avril 2016

**Résumé** – La reprogrammation cellulaire permet d'obtenir des cellules souches pluripotentes induites (iPSC) par l'expression forcée, et limitée dans le temps, de quatre facteurs de transcription embryonnaires. La découverte de cette technologie permettant de convertir une cellule différenciée en une cellule pluripotente a conduit à un profond changement de paradigme dans la notion d'identité cellulaire, puisqu'il est dorénavant possible de produire *in vitro* n'importe quel type cellulaire à partir d'un prélèvement d'un donneur sain ou d'un patient. Le champ d'applications des iPSC est très large, et comporte notamment la modélisation *in vitro* de tissus normaux ou pathologiques, y compris pour le criblage à haut débit de petites molécules pour le développement de nouveaux médicaments. La technologie des iPSC ouvre également de nouvelles voies thérapeutiques dans le domaine de la médecine régénératrice.

**Mots clés** : Cellules souches / cellules souches pluripotentes induites / pluripotence / médecine régénératrice / modélisation de tissus

**Abstract** – Induced pluripotent stem cells: a new paradigm to study human tissues.

Induced pluripotent stem cells (iPSCs) are obtained by reprogramming differentiated cells through forced expression of four embryonic transcription factors. The discovery of this technology, able to transform a differentiated cell into a pluripotent cell, has profoundly shifted the paradigm of the concept of cell identity, since it is now possible to obtain *in vitro* any cell type from an initial sample of skin or blood cells from a healthy volunteer or patient. Applications of iPSCs are exceedingly large, and comprise the *in vitro* modeling of normal or pathological tissues, including for massive drug screening. They also open new therapeutic avenues in the field of regenerative medicine.

**Key words**: Stem cells / induced pluripotent stem cells / cell reprogramming / pluripotency / regenerative medicine

---

### Abréviations

BMP *Bone Morphogenetic Protein*  
CSE cellules souches embryonnaires

CSP cellules souches pluripotentes  
EB *embryoid bodies*,  
FdT facteurs de transcription  
iPSC cellules souches pluripotentes induites

## Les cellules souches pluripotentes

Les cellules souches pluripotentes (CSP) sont des cellules souches capables de se différencier en tout type cellulaire de l'organisme. La découverte des cellules souches pluripotentes découle de l'étude dans les années 70 des tératomes, ces tumeurs bénignes comportant de nombreux tissus matures, y compris, de manière spectaculaire, des phanères tels que des dents et des cheveux. Des CSP ont pu être identifiées et isolées à partir de l'embryon au début de son développement – le stade blastocyste – d'abord dans la souris puis chez l'homme (Evans & Kaufman, 1981; Thomson *et al.*, 1998). Les CSP isolées à partir de l'embryon prolifèrent *in vitro* et constituent des lignées de cellules souches embryonnaires (CSE). Les applications des CSE sont nombreuses, permettant de produire une très large diversité de tissus, ouvrant ainsi la porte à l'étude de leur développement *in vitro*, de leur physiologie ainsi qu'à la production en quantité théoriquement illimitée de cellules et tissus pour une médecine régénératrice (Ramirez *et al.*, 2010). Il convient de souligner que l'utilisation des CSE de souris a permis la création des souris *knock-out* (KO) et transgéniques qui ont apporté une contribution majeure à la biologie contemporaine (Vogel, 2007). Cependant, les CSE posent notamment le problème de la difficulté de choisir le génotype de la lignée ES. Cette difficulté peut être résolue par le clonage thérapeutique, qui a été proposé comme stratégie d'avenir pour la médecine régénératrice (Wakayama, 2004). Cependant, cette technique est très difficile à mettre en œuvre sur le plan technologique, et n'a finalement été appliquée à des cellules humaines qu'en 2013, après près de 10 ans de tentatives infructueuses (Tachibana *et al.*, 2013). De plus, elle requiert un grand nombre d'ovules ce qui pose des problèmes éthiques spécifiques. Elle est interdite en France car elle comporte une étape de création d'embryons pour la recherche. Et enfin, elle n'évite pas la destruction d'un embryon humain et les questions éthiques spécifiques liées à cette recherche.

## Les cellules souches pluripotentes induites

Partant du constat qu'une cellule spécialisée est capable d'être reprogrammée par transplantation nucléaire dans un ovule, le japonais Shinya Yamanaka a proposé de forcer l'expression d'un nombre limité de facteurs de transcription (FdT) associés à la pluripotence dans une cellule différenciée, pour la dédifférencier vers la pluripotence. Par élimination, il a pu montrer que l'expression du quatuor minimum de FdT Oct4, cMyc, Klf4 et Sox2 pendant environ deux

semaines peut entraîner la reprogrammation complète d'une cellule différenciée en cellule pluripotente. Les cellules obtenues sont appelées *induced Pluripotent Stem Cells* (iPSCs) (Takahashi & Yamanaka, 2006; Yamanaka, 2009). L'épigénome, le transcriptome et le protéome de la cellule spécialisée de départ ont été intégralement transformés d'un état spécialisé vers un état pluripotent. La physiologie et les propriétés fonctionnelles des iPSC sont typiquement celles des cellules souches pluripotentes : expression des FdT caractéristiques tels que Oct4 et Nanog, expression des marqueurs de surface tels que SSEA3, SSEA4, TRA-1-60 et TRA-1-81, pousse en colonies à bord net, dépendance au bFGF, prolifération intense, et capacité à se différencier en tout type cellulaire. La pluripotence des iPSC a été démontrée aussi bien *in vitro* par des tests de différenciation dirigée qu'*in vivo* dans la souris immunodéprimée NOD.Cg-Prkdc<sup>scid</sup> Il2rg<sup>tm1Wjl</sup>/SzJ (NSG) par la formation de tératomes comportant des tissus issus des trois feuillets germinaux (*cf. infra*) (Ramirez *et al.*, 2013). Certaines variations dans le nombre et la nature de ces facteurs ont permis de généraliser cette technologie en la rendant applicable à partir de nombreux types cellulaires y compris ceux issus de personnes très âgées et même centenaires (Lapasset *et al.*, 2011). La démonstration qu'il était possible, par une technique relativement facile à mettre en œuvre, d'obtenir des CSP à partir d'un prélèvement d'un individu adulte, a été une découverte majeure. La technologie des iPSC permet donc de produire en laboratoire n'importe quel type cellulaire à partir de cellules de n'importe quel individu, y compris des patients porteurs d'anomalies génétiques, en quantité théoriquement illimitée.

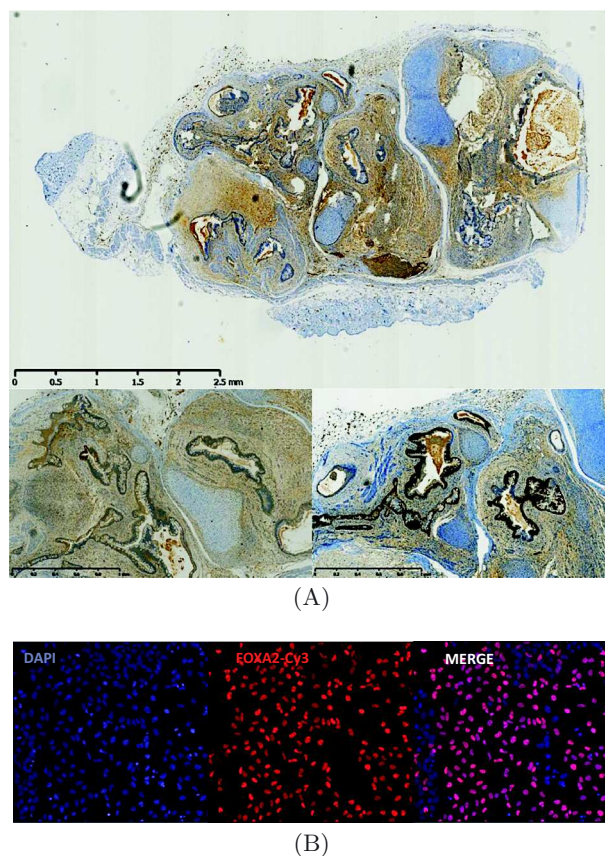
La simplicité et la robustesse de la technologie des iPSC, ainsi que les perspectives scientifiques et médicales des iPSC, ont valu à S. Yamanaka de recevoir le prix Nobel de Physiologie et de Médecine en 2012, conjointement avec John Gurdon (*Science*, 2012).

## Différenciation des iPSC

Les CSP se caractérisent par leur capacité à se différencier en tout type cellulaire. Aussi, dès leur découverte, un enjeu majeur fut d'identifier des protocoles pour guider les CSP à se différencier dans la direction recherchée. On peut schématiquement classer les méthodes pour promouvoir la différenciation des CSP en deux catégories : celles basées sur une différenciation spontanée et désordonnée et celles dites « dirigées ». La première catégorie repose sur l'observation que lorsque des CSP forment des agrégats cellulaires, elles se différencient spontanément. Cela

peut être obtenu *in vitro* par culture en suspension et en l'absence de facteurs de croissance. On parle alors de corps embryoides (*embryoid bodies*, EB), qui sont constitués de divers tissus mélangés (tissu nerveux, tissu cardiaque, cellules hématopoïétiques, différents épithéliums, *etc.*) (Martin & Evans, 1975). Une autre manière de promouvoir une différenciation spontanée est d'injecter des CSP chez une souris immunodéprimée NSG, ce qui induit la croissance d'un tératome en quelques semaines. Le tératome ainsi obtenu est composé d'un mélange de tissus issus des trois feuillets germinaux, notamment du cartilage, des structures cylindriques tapissées d'épithélium digestif et pulmonaire et des structures neurales, ainsi que de nombreux autres tissus (Figure 1A) (Prokhorova *et al.*, 2009). La formation de tératomes reste la technique de référence pour démontrer formellement la pluripotence d'une lignée de CSP.

La différenciation spontanée est donc avantageuse pour montrer les propriétés de pluripotence, mais cette technique présente l'inconvénient de ne pouvoir conduire à l'obtention majoritaire que d'un seul type cellulaire. La production *in vitro* d'un tissu à partir de CSP et notamment des iPSC tente en effet de récapituler le développement humain normal. C'est pourquoi se sont développées des approches de différenciation dirigée, dont l'ambition est de reproduire *in vitro* les étapes successives du développement d'un tissu en mimant les voies de signalisation auxquelles ces tissus sont physiologiquement soumis au cours de leur développement. Les protocoles initiaux faisaient appel à une première étape de formation d'EB, à des co-cultures avec des cellules fœtales (Roy *et al.*, 2006) ou à des surexpressions de gènes (Kim *et al.*, 2002). Pour simplifier les protocoles et éviter les contaminations par des cellules indésirables ou des modifications génétiques des cellules produites, les protocoles actuels privilégient l'usage exclusif de combinaisons successives de facteurs de croissance, d'inhibiteurs chimiques de certaines voies de transduction et de composants de la matrice extracellulaire. Il est par exemple relativement simple d'obtenir une population quasi pure de cellules de l'endoderme définitif exprimant FoxA2 (Figure 1B) par l'ajout d'activine A et de WNT3A. Un autre exemple est apporté par le système nerveux central, pour lequel une amélioration majeure a été apportée par l'équipe de Lorenz Studer qui a montré que l'inhibition duale de la signalisation SMAD – en bloquant la voie du TGF $\beta$  et des BMP – conduisait à une différenciation massive et rapide vers des cellules neurales (Chambers *et al.*, 2009). Des protocoles pour la production de très nombreux tissus, y compris des gamètes, ont été publiés. Une synthèse de ces protocoles a été compilée en 2012 par Kevin Eggan (Williams *et al.*, 2012).



**Fig. 1.** Tératome obtenu à partir d'une lignée d'iPSC. (A) Tératome. Analyse immuno-histochimique d'une section de tératome généré par l'injection de la lignée iPSC humaine M4C7 dans une souris immunodéprimée NOD.Cg-Prkdc<sup>scid</sup> Il2rg<sup>tm1Wjl</sup>/SzJ (NSG). Marquage par GFAP, marqueur astrocytaire (en haut), par TTF1/NKX2.1, marqueur de progéniteurs bronchiques et neuraux (en bas à gauche) et par CDX2, marqueur notamment de l'épithélium intestinal sous-diaphragmatique (en bas à droite). (B) Marquage FOXA2 par immunofluorescence de CSP différenciées en endoderme définitif par traitement pendant 5 jours par activine A et Wnt3a. Coloration de l'ADN par DAPI.

## Les iPSC, un nouveau système modèle pour l'étude des tissus humains

L'immense majorité des tissus primaires ne peut être cultivée *in vitro*. En effet, la mise en culture d'un explant tissulaire se traduit le plus souvent par sa dégénérescence rapide ou au mieux par une croissance limitée dans le temps. Pour ces raisons, l'étude du développement des tissus humains, de leur physiologie et de leurs fonctions au sein des organes dans le cadre de pathologies génétiques provenait essentiellement d'analyses statiques issues de biopsies humaines prélevées à différents stades du développement ou chez

**Tableau 1.** Applications potentielles des iPSC et comparaison avec celles des CSE.

	iPSC	CSE
Étude de la modification du destin cellulaire	X	
Étude <i>in vitro</i> du développement humain normal	X	X
Étude <i>in vitro</i> du développement humain dans le contexte d'une pathologie génétique ou d'un génotype particulier (exemple : individus acétyleurs lents ou rapides)	X	
Étude <i>in vitro</i> des tissus humains normaux	X	X
Étude <i>in vitro</i> des pathologies génétiques	X	
Criblage pharmacologique à grande échelle sur du tissu humain	X	X
Criblage pharmacologique à grande échelle sur des tissus humains du génotype souhaité	X	
Production de cellules et de tissus pour de la thérapie cellulaire	X	X
Production de cellules et de tissus <i>autologues</i> pour de la thérapie cellulaire	X	
Rajeunissement cellulaire	X	

l'adulte, complétées par l'étude du développement des tissus *in vivo* chez l'animal, notamment sur les modèles génétiques des souris *knock-out* et des souris transgéniques. Les cellules ES puis les iPSC ont bouleversé ce paysage en apportant une solution technique au développement des tissus humains *in vitro*. Les perspectives apportées par ce bouleversement peuvent-être classées en trois grandes catégories : (i) étude *in vitro* du développement et du fonctionnement des tissus normaux et pathologiques humains; (ii) production de cellules et de tissus, potentiellement autologues, pour une médecine régénérative; (iii) un rajeunissement extrême de la cellule âgée et/ou sénescence. Les applications potentielles des iPSC sont précisées et comparées avec celles apportées par les ES dans le Tableau 1. Plusieurs revues détaillent ce potentiel, notamment dans le domaine de la pharmacologie pour déterminer l'efficacité ou la toxicité des molécules chimiques candidates à un usage clinique (Kolaja, 2014), ou l'utilisation des iPSC pour la médecine régénératrice (Fox *et al.*, 2014). La technologie CRISPR/Cas9 a ajouté un potentiel supplémentaire en permettant de produire, à partir des iPSC issues d'un donneur porteur d'une mutation génétique, la lignée iPSC témoin isogénique dans laquelle la mutation a été réparée (Chang *et al.*, 2015).

L'étude des tissus *via* les iPSC se heurte cependant à deux limitations : (i) le degré de maturité des tissus obtenus et (ii) la possibilité de produire un tissu complexe (point abordé au chapitre suivant). La production *in vitro* d'un tissu à partir de CSP tente en effet de récapituler, comme nous l'avons pointé plus haut, le développement humain normal. Pour la majorité des tissus, les protocoles de différenciation *in vitro* actuels parviennent à un stade foetal mais pas à un stade adulte. Il est ainsi notoire que les CSP se différencient en cellules hépatocytaires ou plutôt *hepatocyte-like* qui

expriment un niveau d'albumine inférieur à celui des hépatocytes adultes et un profil de cytochrome p450 de type foetal (Funakoshi *et al.*, 2011). Un autre exemple illustrant ce point est l'impossibilité, en l'état actuel des protocoles de différenciation, d'obtenir *in vitro* le basculement d'une hémoglobine foetale vers une hémoglobine adulte au cours de la production de globules rouges à partir d'iPSC alors que l'injection de progéniteurs issus d'iPSC dans la « niche hématopoïétique » adulte murine permet cette conversion (Kobari *et al.*, 2012).

## Production de tissus en 3D *in vitro*

La majorité des protocoles actuels de différenciation est basée sur une culture des tissus en deux dimensions (2D), comportant une ou un nombre limité de couches cellulaires. Or les tissus sont des ensembles cellulaires plus complexes, se développant et fonctionnant en trois dimensions (3D). Même les épithéliums les plus simples, monocouche cellulaire reposant sur une lame basale, développent une relation étroite avec le tissu sous-épithélial et leur fonctionnement normal est lié à ce tissu sous-jacent. Les protocoles de différenciation actuels des CSP aboutissant à des tissus en 2D ne reproduisent donc qu'imparfaitement les tissus humains. Les EB sont une exception, mais leur différenciation diversifiée et chaotique limite leur utilisation pour produire un tissu précis.

Pour répondre à cette insuffisance, des protocoles visant à produire une réplique de petite dimension d'organes complexes ont été développés. Ces structures 3D flottantes dans le milieu de culture sont appelées « organoïdes » et sont devenues une technique majeure pour mimer *in vitro* le développement normal et pathologique (Takasato *et al.*, 2015). Parmi les exemples les plus spectaculaires, on citera le

développement *in vitro* de la rétine humaine, reproduisant celui de la cupule optique puis de la vésicule optique, et aboutissant à une rétine comportant l'épithélium pigmenté de la rétine, une couche de photorécepteurs comportant cônes et bâtonnets et une couche neuronale incluant cellules ganglionnaires et interneurons (Nakano *et al.*, 2012). Ces structures rétinienne complexes et organisées peuvent être congelées et décongelées ultérieurement pour étude physiologique. Un autre exemple est le développement *in vitro* d'une structure mimant certaines parties du cerveau. Ces structures, appelées « *mini-brain* », reproduisent plusieurs régions du cerveau humain, notamment un cortex comportant différents sous-types de neurones corticaux, et une zone sous-jacente contenant des progéniteurs sous forme de cellules gliales radiales (Lancaster *et al.*, 2013). Le mimétisme du développement cérébral est attesté par la reproduction *in vitro*, grâce à cette approche, de la microcéphalie, une pathologie humaine se traduisant par un défaut prononcé du développement cérébral. En effet, la génération de *mini-brain* à partir d'iPSC dérivées d'un patient atteint d'une forme sévère de microcéphalie par mutations dans le gène *CDK5RAP2* résulte en des *mini-brains* réduits en taille en raison d'une maturation prématurée des progéniteurs neuronaux, et donc un déficit final en neurones matures. De plus, ce phénotype a été reproduit en réalisant un *knock-down* du gène *CDK5RAP2* par shRNA à partir d'iPSC issues d'un donneur sain. D'autres exemples, reflétant le développement d'autres organes, ont été rapportés dans la littérature récente, notamment la production de sphéroïdes pulmonaires comportant un épithélium alvéolaire et mimant les structures distales du poumon et comprenant des cellules alvéolaires de type II porteuses de corps lamellaires dans lesquels est stocké le surfactant (Gotoh *et al.*, 2014). Un autre exemple est la génération d'organoïdes mimant le développement et la physiologie de la muqueuse gastrique. Ces organoïdes gastriques peuvent être infectés par *Helicobacter pylori* et reproduire ainsi *in vitro* une pathologie gastrique d'importance clinique majeure (McCracken *et al.*, 2014). Des organoïdes dérivés des CSP reproduisant des parties d'organes ont également été décrits pour le rein (Takasato *et al.*, 2015), l'intestin grêle (Watson *et al.*, 2014), ou le foie (Takebe *et al.*, 2013) et le pancréas exocrine (Huang *et al.*, 2015).

## Production de tissus humains *in vivo* dans l'animal

Cependant, malgré l'intérêt considérable que suscitent les organoïdes dérivés des CSP, certaines différences persistent par rapport aux tissus *in vivo*. Le défaut le plus notable est l'absence de microvascularisation. En effet, tout tissu, à l'exception du cartilage et des

tissus transparents de l'œil, est parcouru d'un réseau plus ou moins dense de capillaires. Ces vaisseaux apportent oxygène et nutriments au cœur des tissus et permettent l'évacuation du gaz carbonique et des déchets métaboliques, mais pourraient également directement contribuer au développement de certains tissus (Armulik *et al.*, 2011). L'absence de microvascularisation est donc un frein au développement d'organoïdes de grandes tailles, en raison de la nécrose qui survient en leur centre. Pour pallier cette insuffisance, certains auteurs ont proposé d'utiliser l'impression cellulaire 3D, afin de déposer de manière précise des cellules de différents sous-types, y compris des cellules endothéliales (Collins, 2014). La faisabilité à court terme de cette approche, qui mérite d'être explorée, reste cependant incertaine pour plusieurs raisons : (i) les connections intercellulaires, qui se mettent en place au cours du développement, pourraient avoir du mal à se mettre en place *a posteriori*; (ii) les différents sous-types cellulaires et les différents degrés de maturité qui sont impliqués dans un tissu donné sont complexes et seront difficiles à reproduire; et (iii) il sera difficile de rendre fonctionnelle la microcirculation. Pour ce dernier point, il convient de souligner que pour produire une microcirculation fonctionnelle, il faudrait non seulement construire les capillaires, structures cylindriques étanches – ce qui paraît déjà un obstacle difficile à franchir par impression cellulaire 3D – mais il faudrait, en outre, positionner des globules rouges (GR) à l'intérieur de ces vaisseaux, les oxygéner et les faire circuler sous pression. Répondre à ce quadruple cahier des charges (créer le vaisseau, le remplir, oxygéner le contenu et le faire circuler sous pression) est un défi qui sera très difficile à relever *in vitro*. En revanche, c'est exactement ce qui se produit au cours de tout développement normal d'un organe chez le vivant. Aussi, l'environnement le plus adapté à une différenciation des CSP humaines – mis à part l'être humain lui-même – est l'animal. Comme décrit précédemment, la différenciation de CSP dans l'animal se fait classiquement par l'induction de tératomes en les injectant à des souris profondément immunodéprimées NSG. Mais des approches plus sophistiquées permettent d'orienter la différenciation dans une voie spécifique, et par exemple d'obtenir des organoïdes hépatiques humains vascularisés et connectés à la circulation murine (Takebe *et al.*, 2013). Les perspectives dans le traitement des insuffisances hépatiques sont notables. Récemment, une autre stratégie a été proposée, qui vise, *via* la création d'organismes chimères animal/humain, à produire des organes humains structurés dans des animaux porteurs. La preuve de concept a été apportée chez le petit animal à faible divergence évolutive par l'équipe de Hiromitsu Nakauchi en 2010 en injectant des iPSC de rat à des embryons de souris au stade

de blastocyste (Kobayashi *et al.*, 2010). Les souris obtenues par cette approche comportent un contingent significatif de cellules de rats. Ce qui est particulièrement remarquable est que ces iPSC de rat sont capables de permettre la survie des souris *pdx1*<sup>-/-</sup> dont le gène *pdx1*, nécessaire au développement du pancréas a été invalidé, en complétant la fonction défaillante. Les souriceaux vivants produisent alors de l'insuline et leur pancréas est formé de cellules de rat. Ce travail montre qu'il est possible de créer un animal porteur d'un organe d'une autre espèce animale. Si on extrapole cette stratégie au couple porc/humain, il serait concevable de créer des porcs porteurs d'organes humains tels qu'un pancréas, un cœur ou un foie (Rashid *et al.*, 2014). Les applications médicales d'une telle approche sont nombreuses, en particulier la production d'organes humains pour la transplantation d'organes. Bien sûr, de nombreux obstacles restent à lever, notamment s'assurer que les voies de signalisation développementales soient suffisamment conservées entre le porc et l'humain, développer des stratégies pour prévenir le risque de transmission de zoonoses à l'homme, en particulier des rétrovirus porcins (Yang *et al.*, 2015), et bien sûr résoudre les objections éthiques que cette approche soulève, notamment prévenir les risques d'humanisation de certaines parties du porc (cerveau, face, gonades, *etc.*) (Giquel *et al.*, 2015).

### Critiques de la technologie iPSC

Après l'euphorie qui a suivi la découverte de la technologie des iPSC (Cibelli, 2007), des critiques se sont élevées, accusant les iPSC de biais épigénétiques pouvant modifier leurs propriétés (notamment leurs capacités de différenciation) mais également d'altérations de leur patrimoine génétique (Kim *et al.*, 2010; Pera, 2011; Bai *et al.*, 2013). Avec le recul, il apparaît que les variations épigénétiques ou les anomalies génomiques ne sont pas systématiquement retrouvées, et pourraient dépendre de la qualité des protocoles de reprogrammation (Carey *et al.*, 2011; Yamanaka, 2012; Bai *et al.*, 2013). S'il existe une grande variabilité dans les capacités de différenciation des CSP, cette variabilité n'est pas spécifique aux iPSC mais partagée avec les CSE (Osafune *et al.*, 2008) et semble relever du polymorphisme génétique (Rouhani *et al.*, 2014; Choi *et al.*, 2015). Ainsi, les iPSC apparaissent aujourd'hui comme l'équivalent biologique des CSE, ne différant que par leur origine, sous réserve d'utiliser des protocoles optimisés de reprogrammation (Vallier, 2015).

### Conclusion

Les iPSC représentent une technologie révolutionnaire de la biologie. Sur le plan fondamental, elles ont

démonstré que le destin cellulaire imposé par la différenciation ou le vieillissement sont entièrement réversibles par la simple surexpression transitoire de gènes de la pluripotence. Sur le plan des biotechnologies, les iPSC apportent des outils profondément innovants, pour la modélisation *in vitro* des tissus normaux ou pathologiques porteurs de maladies génétiques, ou encore pour la production de cellules voire de tissus ou d'organes pour la médecine régénératrice. Certaines applications telles que la production de tissus *in vitro* à des fins de recherche sont déjà bien établies, tandis que d'autres, telles que le criblage de nouveaux médicaments ou la production de cellules et de tissus, pour la thérapie cellulaire, sont encore au stade du développement.

### Références

- Armulik, A., Genové, G., and Betsholtz, C. (2011). Pericytes : developmental, physiological, and pathological perspectives, problems, and promises. *Dev Cell*, 21, 193-215.
- Bai, Q., Desprat, R., Klein, B., Lemaitre, J.-M., and De Vos, J. (2013). Embryonic stem cells or induced pluripotent stem cells? A DNA integrity perspective. *Curr Gene Ther*, 13, 93-98.
- Carey, B.W., Markoulaki, S., Hanna, J.H., Faddah, D.A., Buganim, Y., Kim, J., Ganz, K., Steine, E.J., Cassady, J.P., Creighton, M.P., Welstead G.G., Gao Q., and Jaenisch R. (2011). Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. *Cell Stem Cell*, 9, 588-598.
- Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sadelain, M., and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol*, 27, 275-280.
- Chang, C.-W., Lai, Y.-S., Westin, E., Khodadadi-Jamayran, A., Pawlik, K.M., Lamb, L.S., Goldman, F.D., and Townes, T.M. (2015). Modeling Human Severe Combined Immunodeficiency and Correction by CRISPR/Cas9-Enhanced Gene Targeting. *Cell Rep*, 12, 1668-1677.
- Choi, J., Lee, S., Mallard, W., Clement, K., Tagliazucchi, G.M., Lim, H., Choi, I.Y., Ferrari, F., Tsankov, A.M., Pop, R., Lee G., Rinn J.L., Meissner A., Park P.J., and Hochedlinger K. (2015). A comparison of genetically matched cell lines reveals the equivalence of human iPSCs and ESCs. *Nat Biotechnol*, 33, 1173-1181.
- Cibelli, J.B. (2007). Development. Is therapeutic cloning dead? *Science*, 318, 1879-1880.
- Collins, S.F. (2014). Bioprinting is changing regenerative medicine forever. *Stem Cells Dev*, 23, 79-82.
- Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature*, 292, 154-156.

- Fox, I.J., Daley, G.Q., Goldman, S.A., Huard, J., Kamp, T.J., and Trucco, M. (2014). Stem cell therapy. Use of differentiated pluripotent stem cells as replacement therapy for treating disease. *Science*, 345, 1247391-1247391.
- Funakoshi, N., Duret, C., Pascucci, J.-M., Blanc, P., Maurel, P., Daujat-Chavanieu, M., and Gerbal-Chaloin, S. (2011). Comparison of hepatic-like cell production from human embryonic stem cells and adult liver progenitor cells: CAR transduction activates a battery of detoxification genes. *Stem Cell Rev Rep*, 7, 518-531.
- Giquel, C., De Vos, J., and Bourret, R. (2015). *Creation of chimeric animals bearing human organs*. Médecine et Droit. Elsevier Masson France, pp. 1-11.
- Gotoh, S., Ito, I., Nagasaki, T., Yamamoto, Y., Konishi, S., Korogi, Y., Matsumoto, H., Muro, S., Hirai, T., Funato, M., Mae S., Toyoda T., Sato-Otsubo A., Ogawa S., Osafune K., and Mishima M. (2014). Generation of Alveolar Epithelial Spheroids via Isolated Progenitor Cells from Human Pluripotent Stem Cells. *Stem Cell Reports*, 3, 394-403.
- Huang, L., Holtzinger, A., Jagan, I., BeGora, M., Lohse, I., Ngai, N., Nostro, C., Wang, R., Muthuswamy, L.B., Crawford, H.C., Arrowsmith C., Kalloger S.E., Renouf D.J., Connor A.A., Cleary S., Schaeffer D.F., Roehrl M., Tsao M.S., Gallinger S., Keller G., and Muthuswamy S.K. (2015). Ductal pancreatic cancer modeling and drug screening using human pluripotent stem cell- and patient-derived tumor organoids. *Nat Med*, 21, 1364-1371.
- Kim, J.-H., Auerbach, J.M., Rodríguez-Gómez, J.A., Velasco, I., Gavin, D., Lumelsky, N., Lee, S.-H., Nguyen, J., Sánchez-Pernaute, R., Bankiewicz, K., and McKay, R. (2002). Dopamine neurons derived from embryonic stem cells function in an animal model of Parkinson's disease. *Nature*, 418, 50-56.
- Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M.J., Ji, H., Ehrlich, L.I.R., Yabuuchi A., Takeuchi, A., Cunniff, K.C., Hongguang H., McKinney-Freeman, S., Naveiras, O., Yoon, T.J., Irizarry, R.A., Jung, N., Seita, J., Hanna, J., Murakami, P., Jaenisch, R., Weissleder, R., Orkin, S.H., Weissman, I.L., Feinberg, A.P., and Daley, G.Q. (2010). Epigenetic memory in induced pluripotent stem cells. *Nature*, 467, 285-290.
- Kobari, L., Yates, F., Oudrhiri, N., Francina, A., Kiger, L., Mazurier, C., Rouzbeh, S., Nemer, El, W., Hebert, N., Giarratana, M.-C.C., François S., Chapel A., Lapillonne H., Luton D., Bennaceur-Griscelli A., and Douay L. (2012). Human induced pluripotent stem cells can reach complete terminal maturation : in vivo and in vitro evidence in the erythropoietic differentiation model. *Haematologica*, 97, 1795-1803.
- Kobayashi, T., Yamaguchi, T., Hamanaka, S., Kato-Itoh, M., Yamazaki, Y., Iyata, M., Sato, H., Lee, Y.-S., Usui, J.-I., Knisely, A.S., Hirabayashi M., and Nakauchi H. (2010). Generation of rat pancreas in mouse by interspecific blastocyst injection of pluripotent stem cells. *Cell*, 142, 787-799.
- Kolaja, K. (2014). Stem cells and stem cell-derived tissues and their use in safety assessment. *J Biol Chem*, 289, 4555-4561.
- Lancaster, M.A., Renner, M., Martin, C.-A., Wenzel, D., Bicknell, L.S., Hurles, M.E., Homfray, T., Penninger, J.M., Jackson, A.P., and Knoblich, J.A. (2013). Cerebral organoids model human brain development and microcephaly. *Nature*, 501, 373-379.
- Lapasset, L., Milhavel, O., Prieur, A., Besnard, E., Babled, A., Ait-Hamou, N., Leschik, J., Pellestor, F., Ramirez, J.M., De Vos, J., Lehmann, S., and Lemaitre, J.M. (2011). Rejuvenating senescent and centenarian human cells by reprogramming through the pluripotent state. *Genes Dev*, 25, 2248-2253.
- Martin, G.R., and Evans, M.J. (1975). Differentiation of clonal lines of teratocarcinoma cells : formation of embryoid bodies in vitro. *Proc Natl Acad Sci USA*, 72, 1441-1445.
- McCracken, K.W., Catá, E.M., Crawford, C.M., Sinagoga, K.L., Schumacher, M., Rockich, B.E., Tsai, Y.-H., Mayhew, C.N., Spence, J.R., Zavros, Y., and Wells J.M. (2014). Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. *Nature*, 516, 400-404.
- Nakano, T., Ando, S., Takata, N., Kawada, M., Muguruma, K., Sekiguchi, K., Saito, K., Yonemura, S., Eiraku, M., and Sasai, Y. (2012). Self-Formation of Optic Cups and Storable Stratified Neural Retina from Human ESCs. *Cell Stem Cell*, 10, 771-785.
- Osafune, K., Caron, L., Borowiak, M., Martinez, R.J., Fitz-Gerald, C.S., Sato, Y., Cowan, C.A., Chien, K.R., and Melton, D.A. (2008). Marked differences in differentiation propensity among human embryonic stem cell lines. *Nat Biotechnol*, 26, 313-315.
- Pera, M.F. (2011). Stem cells : The dark side of induced pluripotency. *Nature*, 471, 46-47.
- Prokhorova, T.A., Harkness, L.M., Frandsen, U., Ditzel, N., Schröder, H.D., Burns, J.S., and Kassem, M. (2009). Teratoma formation by human embryonic stem cells is site dependent and enhanced by the presence of Matrigel. *Stem Cells Dev*, 18, 47-54.
- Ramirez, J.M., Bai, Q., Dijon-Grinand, M., Assou, S., Gerbal-Chaloin, S., Hamamah, S., and De Vos, J. (2010). Human pluripotent stem cells : From biology to cell therapy. *World J Stem Cells*, 2, 24-33.
- Ramirez, J.M., Bai, Q., Pequignot, M., Becker, F., Kassambara, A., Bouin, A., Kalatzis, V., Dijon-Grinand, M., and De Vos, J. (2013). Side scatter intensity is highly heterogeneous in undifferentiated pluripotent stem cells and predicts clonogenic self-renewal. *Stem Cells Dev*, 22, 1851-1860.
- Rashid, T., Kobayashi, T., and Nakauchi, H. (2014). Revisiting the Flight of Icarus : Making Human Organs from PSCs with Large Animal Chimeras. *Cell Stem Cell*, 15, 406-409.
- Rouhani, F., Kumasaka, N., de Brito, M.C., Bradley, A., Vallier, L., and Gaffney, D. (2014). Genetic background drives transcriptional variation in human induced pluripotent stem cells. *PLoS Genet*, 10, e1004432.

- Roy, N.S., Cleren, C., Singh, S.K., Yang, L., Beal, M.F., and Goldman, S.A. (2006). Functional engraftment of human ES cell-derived dopaminergic neurons enriched by coculture with telomerase-immortalized midbrain astrocytes. *Nat Med*, 12, 1259-1268.
- Science, A.A.F.T.A.O. (2012). Nobel Prize in physiology or medicine. Reprogrammed cells earn biologists top honor. *Science*, 338, 178-179.
- Tachibana, M., Amato, P., Sparman, M., Gutierrez, N.M., Tippner-Hedges, R., Ma, H., Kang, E., Fulati, A., Lee, H.-S., Sritanaudomchai, H., Masterson K., Larson J., Eaton D., Sadler-Fredd K., Battaglia D., Lee D., Wu D., Jensen J., Patton P., Gokhale S., Stouffer R.L., Wolf D., and Mitalipov S. (2013). Human embryonic stem cells derived by somatic cell nuclear transfer. *Cell*, 153, 1228-1238.
- Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, 126, 663-676.
- Takasato, M., Er, P.X., Chiu, H.S., Maier, B., Baillie, G.J., Ferguson, C., Parton, R.G., Wolvetang, E.J., Roost, M.S., Chuva de Sousa Lopes, S.M., and Little M.H. (2015). Kidney organoids from human iPSC cells contain multiple lineages and model human nephrogenesis. *Nature*, 526, 564-568.
- Takebe, T., Sekine, K., Enomura, M., Koike, H., Kimura, M., Ogaeri, T., Zhang, R.-R., Ueno, Y., Zheng, Y.-W., Koike, N., Aoyama S., Adachi Y., and Taniguchi H. (2013). Vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature*, 499, 481-484.
- Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science*, 282, 1145-1147.
- Vallier, L. (2015). Putting induced pluripotent stem cells to the test. *Nat Biotechnol*, 33, 1145-1146.
- Vogel, G. (2007). Nobel Prizes. A knockout award in medicine. *Science*, 318, 178-179.
- Wakayama, T. (2004). On the road to therapeutic cloning. *Nat Biotechnol*, 22, 399-400.
- Watson, C.L., Mahe, M.M., Múnera, J., Howell, J.C., Sundaram, N., Poling, H.M., Schweitzer, J.I., Vallance, J.E., Mayhew, C.N., Sun, Y., Grabowski G., Finkbeiner S.R., Spence J.R., Shroyer N.F., Wells J.M., and Helmrath M.A. (2014). An in vivo model of human small intestine using pluripotent stem cells. *Nat Med*, 20, 1310-1314.
- Williams, L.A., Davis-Dusenbery, B.N., and Eggan, K.C. (2012). SnapShot : Directed Differentiation of Pluripotent Stem Cells. *Cell*, 149, 1174-1174.e1.
- Yamanaka, S. (2009). Ekiden to iPSC Cells. *Nat Med*, 15, 1145-1148.
- Yamanaka, S. (2012). Induced pluripotent stem cells : past, present, and future. *Cell Stem Cell*, 10, 678-684.
- Yang, L., Guell, M., Niu, D., George, H., Lesho, E., Grishin, D., Aach, J., Shrock, E., Xu, W., Poci, J., Cortazio R, Wilkinson RA, Fishman JA, and Church G. (2015). Genome-wide inactivation of porcine endogenous retroviruses (PERVs). *Science*, 350, 1101-1104.



Nous nous attarderons sur la différenciation dirigée dans la voie pulmonaire dans le chapitre suivant.

## 2.2 Modéliser l'épithélium bronchique

La modélisation de l'épithélium bronchique humain par la différenciation de CSP s'appuie sur une bonne connaissance différentes étapes du développement embryonnaire et notamment du système respiratoire. L'objectif est de délivrer les bonnes cytokines pour activer ou inhiber les bonnes voies de signalisation au bon moment, en mimant les processus embryonnaires et fœtaux physiologiques (voir la Partie 1, Chapitre 1). Le bon déroulement de la différenciation est suivi à chaque étape grâce aux marqueurs nucléaires et membranaires.

La première étape consiste donc à obtenir de l'endoderme définitif ou definitive endoderm en anglais (DE). Kévin D'Amour et collaborateurs sont les premiers à avoir mis au point un protocole robuste de production *in vitro* de DE humain (D'Amour et al., 2005). Ce protocole fait intervenir l'Activine A à forte concentration et peu de sérum. En effet, l'Activine A à faible concentration promeut la voie mésodermique (Green and Smith, 1990), tandis qu'à forte concentration mais en présence d'une forte activité PI3K – par exemple par une forte concentration de sérum – l'Activin A maintient la pluripotence. L'obtention de DE est d'une étape commune avec la différenciation vers le foie, le tube digestif et le pancréas car ce sont aussi des organes d'origine endodermique. Les marqueurs nucléaires de la pluripotence disparaissent et les marqueurs nucléaires, SOX17 et FOXA2, et membranaires, CXCR4, apparaissent.

L'étape suivante est l'obtention de l'intestin primitif antérieur, ou anterior foregut endoderm (AFE). Différents protocoles de différenciation du DE en AFE ont été décrits. Une double inhibition de BMP, via l'ajout de NOGGIN et d'un inhibiteur chimique de la voie de signalisation TGF- $\beta$ , permet une production efficace d'AFE (Dye et al., 2015; Gotoh et al.,

2014; Green et al., 2011). Une autre approche permet d'obtenir l'AFE par l'ajout des cytokines Sonic Hedgehog (SHH) et FGF2 à forte concentration comme le délivrerait le bourgeon cardiaque *in vivo* (Gotoh et al., 2014). Une faible concentration de FGF2 dirige en effet vers la différenciation hépatique (Wong et al., 2012). L'AFE se caractérise par la réinduction de l'expression de SOX2 et l'absence de CDX2, marqueur de l'intestin moyen et postérieur.

L'étape suivante vise à « ventraliser » l'AFE. C'est en effet au niveau de la partie ventrale de l'AFE, au niveau du sillon laryngotrachéal, que l'ébauche pulmonaire va émerger. Cette zone développementale se caractérise par l'apparition du facteur de transcription Thyroid Transcription Factor 1 (TTF1), aussi appelé NK2 homeobox 1 (NKX2.1). Ce facteur de transcription est également important pour le développement de la thyroïde, ainsi qu'une partie du cerveau antérieur (Lazzaro et al., 1991). La ventralisation de l'AFE peut-être obtenue par l'ajout de WNT3a, KGF, FGF10, BMP4 et EGF (" WKFBE") (Green et al., 2011) ou par l'ajout de BMP4, ATRA et CHIR99021 – un agoniste de la voie WNT (Gotoh et al., 2014), ou bien par l'ajout de l'agoniste de Smoothened, SAG (Dye et al., 2015).

Les progéniteurs bronchiques ou lung progenitor (PL) sont ensuite obtenus à partir de l'AFE ventralisé. Différentes associations de cytokines ont été utilisées pour cette étape. Parfois, cette étape et la précédente sont regroupées en une seule. Un facteur de croissance essentiel à ce stade est le FGF 10, agoniste du récepteur FGFR2 et qui est nécessaire au développement du bourgeon pulmonaire et des bifurcations/branchements qui s'ensuivent (Bellusci et al., 1997; Min et al., 1998). L'activation des voies BMP et WNT est également importante. Il est à noter que l'induction de la voie WNT plus précocement induit plus la différenciation vers le hindgut (Sherwood et al., 2011). Ce qui laisse à penser que la différenciation est un jeu très complexe d'activation et de répression des voies de signalisation à des instants précis. De plus, d'autres voies de signalisation, également impliquées dans la formation des bourgeons pulmonaires, sont représentées dans les différents protocoles publiés comme SHH (Peng and Morrissey, 2013), l'acide rétinoïque (Chen et al., 2010) et TGF- $\beta$  (Que et al., 2006).

Enfin, la dernière étape consiste à diriger la différenciation des PL vers un type d'épithélium de l'arbre respiratoire plus ou moins mature et plus ou moins distal. C'est donc tout naturellement à ce stade que les protocoles varient le plus en fonction du type cellulaire final souhaité.

La différenciation dirigée vers les cellules de l'arbre respiratoire peut ainsi être schématisée en quelques grandes étapes et suivies par l'apparition successive de différents marqueurs (Figure 22).

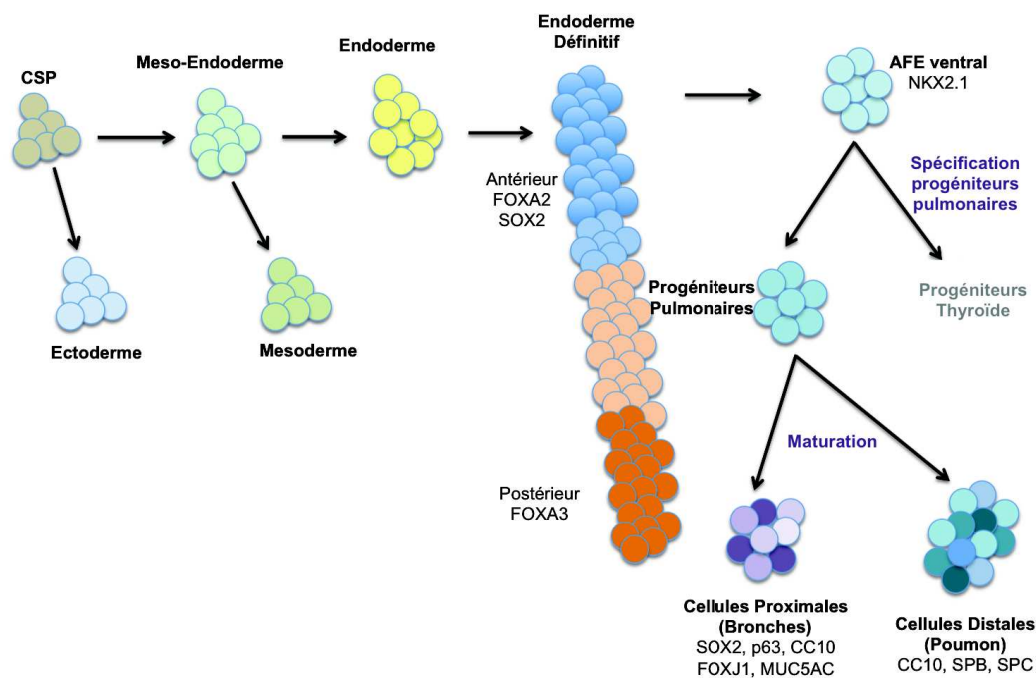


Figure 24 : Schématisation des grandes étapes de la différenciation dirigée de l'arbre respiratoire (d'après (Kadzik and Morrisey, 2012)).

Les protocoles princeps utilisés dans la modélisation de l'épithélium sont résumés dans la Figure 23.

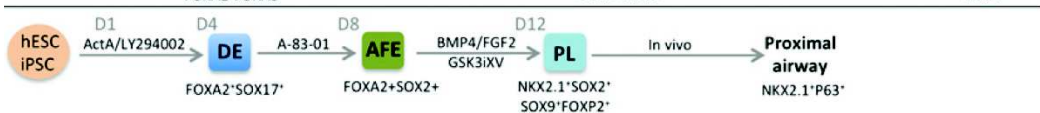
(Green et al., 2011)



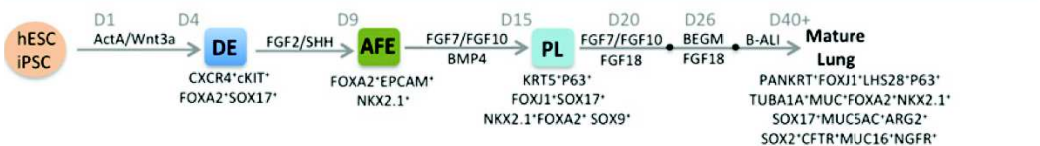
(Longmire et al., 2012)



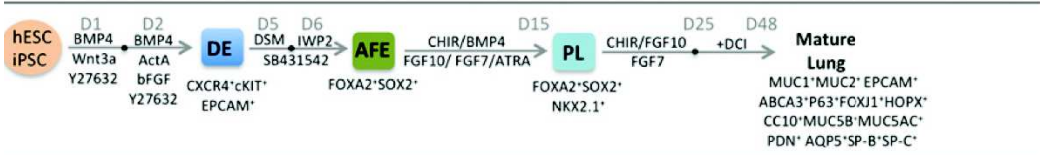
(Mou et al., 2012)



(Wong et al., 2012)



(Huang et al., 2015)



(Firth et al., 2014)



Figure 25 : Protocoles de différenciation en épithélium de l'appareil respiratoire (d'après (Gomperts, 2014)). Dans ce schéma, l'étape de ventralisation de l'AFE n'est pas individualisée.

Une autre approche est de générer directement *in vitro* des mini-poumons à partir d'iPS (Dye et al., 2015). En effet, une alternative à la différenciation des CSP en 2D est leur différenciation en structures 3D. La technique de culture des organoïdes sera traitée ultérieurement dans la Partie 4 chapitre 2.2.

# **PARTIE 4 : Les Applications Médicales**

Les applications des cellules souches pluripotentes sont nombreuses. Nous nous arrêterons sur les utilisations des iPS chez l'homme. En effet, outre le fait que ces cellules permettent de s'affranchir des problèmes éthiques liés à la destruction d'embryons humains, elles offrent la possibilité de choisir le patrimoine génétique du patient à l'origine des iPS. Les deux principaux usages dans le milieu médical de ces cellules sont la médecine régénératrice et la modélisation de tissus d'intérêt pour les industries pharmaceutiques (Figure 24).

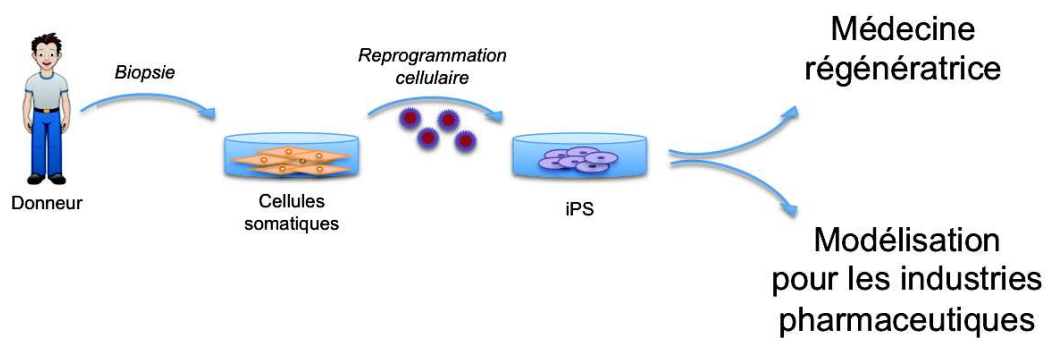


Figure 26 : Les deux grands types d'applications des iPS pour le développement de médicaments.

## 1. Médecine régénératrice

Les CSP ont une place de choix dans la médecine régénératrice. En effet, une des difficultés majeures de ce domaine naissant de la médecine est la source des cellules. Pour de nombreux tissus, il est difficile d'obtenir des cellules thérapeutiques à partir de cellules souches adultes. Les CSP sont une solution car elles peuvent produire n'importe quel type cellulaire et de différents degrés de maturité. Les applications potentielles sont nombreuses, et pourraient passer par l'injection de cellules-médicaments à la production d'un tissu voire d'un organe pour une greffe. Les dérivés différenciés des CSP entrent dans le grand groupe des médicaments de thérapie innovante.

## 1.1 Les médicaments de thérapie innovante

Les thérapies issues des hiPS feront partie de la famille des médicaments de thérapie innovante ou MTI. L'agence européenne du médicament (EMA) a créé un comité, le comité des thérapies avancées (Committee for Advanced Therapies), responsable d'évaluer la qualité, la sécurité et l'efficacité des MTI.

Les textes européens définissent, directive 2001/83/CE, complétée par 1394/2007/CE quatre types de médicaments de thérapie innovante ou « advanced therapy medicinal products » en anglais (ATMP)<sup>2</sup>.

- Les médicaments de thérapie génique :

a) il contient une substance active qui contient ou constitue un acide nucléique recombinant administré à des personnes en vue de réguler, de réparer, de remplacer, d'ajouter ou de supprimer une séquence génétique; et

b) son effet thérapeutique, prophylactique ou diagnostique dépend directement de la séquence d'acide nucléique recombinant qu'il contient ou du produit de l'expression génétique de cette séquence.

- Les médicaments de thérapie cellulaire somatique:

a) contient ou consiste en des cellules ou des tissus qui ont fait l'objet d'une manipulation substantielle de façon à modifier leurs caractéristiques biologiques, leurs fonctions physiologiques ou leurs propriétés structurelles, ou des cellules ou tissus qui ne sont pas destinés à être utilisés pour la ou les mêmes fonctions essentielles chez le receveur et le donneur; et

b) est présenté comme possédant des propriétés permettant de traiter, de prévenir ou de diagnostiquer une maladie à travers l'action métabolique, immunologique ou pharmacologique de ses cellules ou tissus, ou est utilisé chez une personne ou administré à

---

<sup>2</sup> <http://ansm.sante.fr/L-ANSM2/Medicaments-de-therapie-innovante-et-preparations-cellulaires-a-finalite-therapeutique/Les-medicaments-de-therapie-innovante-prepare-ponctuellement-MTI-PP-Exemption-hospitaliere>

une personne dans une telle perspective.

- Les médicaments issus de l'ingénierie tissulaire ou cellulaire :

a) les cellules ou tissus ont été soumis à une manipulation substantielle, de façon à obtenir des caractéristiques biologiques, des fonctions physiologiques ou des propriétés structurales utiles à la régénération, à la réparation ou au remplacement recherchés ; ou  
b) les cellules ou les tissus ne sont pas destinés à être utilisés pour la (les) même(s) fonction(s) essentielle(s) chez le receveur et chez le donneur

- Les médicaments combinés de thérapie innovante :

Ce sont des médicaments de thérapie cellulaire, de thérapie génique ou d'ingénierie tissulaire qui répondent aux conditions suivantes :

a) intègrent dans leur composition un ou plusieurs dispositifs médicaux au sens de l'article 1er, paragraphe 2, point a), de la directive 93/42/CEE, ou bien un ou plusieurs dispositifs médicaux implantables actifs au sens de l'article 1er, paragraphe 2, point c), de la directive 90/385/CEE ;

b) et : leur partie cellulaire ou tissulaire doit contenir des cellules ou des tissus viables, ou, leur partie cellulaire ou tissulaire contenant des cellules ou des tissus non viables doit être susceptible d'avoir sur le corps humain une action qui peut être considérée comme essentielle par rapport à celle des dispositifs précités.

La fabrication de ces médicaments doit répondre aux exigences de bonnes pratiques de fabrication applicables aux médicaments à usage humain (directive 2003/94/CE) et aux guidelines européennes (EudraLEX volume 4). Le développement ainsi que la mise sur le marché devront satisfaire également à ces exigences de qualité, sécurité et efficacité. Pour faciliter et consolider cela, l'ANSM et l'EMA proposent l'aide de différents comités : comités scientifiques, comités pour thérapies avancées, comités pour produits médicaux humains.



## 1.2 Des MTI à base d'iPS

Les iPS ne peuvent pas être utilisées telles quelles dans un but de thérapie régénérative. Le risque majeur serait la formation de tumeurs ou tératomes en raison de leur propriété de pluripotence. Leur utilisation comme MTI doit donc se faire après différenciation en un type cellulaire d'intérêt. Il est donc nécessaire d'optimiser les protocoles de différenciation déjà existants en les adaptant aux exigences relatives aux MTI. A ce jour, il est également envisageable de former des mini-organes ou organoïdes, tel que le foie (Huch and Koo, 2015). Si ces organoïdes ne pourront se substituer à un organe complet pour des besoins de transplantation, ils pourraient constituer une source très intéressante de cellules ou de tissus pour une médecine régénératrice, voire constituer un organe d'appoint *in vivo* (Takebe et al., 2013).

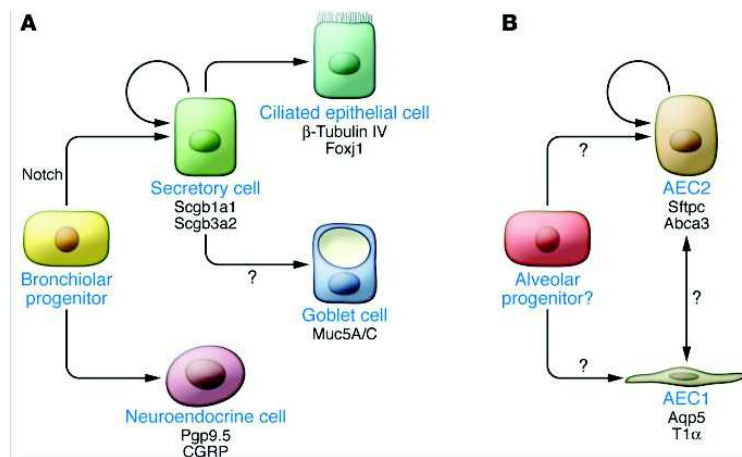
Nous avons vu qu'il est possible d'obtenir le phénotype d'une pathologie dans une boîte de Pétri. Deux modes d'action sont possible pour rendre ce phénotype normal et donc améliorer les traitements spécifiques : soit reverser le phénotype par l'utilisation de médicaments (Lee et al., 2009b) – et dans ce cas les iPS ont servi à identifier un nouveau médicament qui pourra être administré à un patient pour le soigner, soit corriger le gène muté (Raya et al., 2009). Dans ce dernier cas, il s'agit d'une médecine personnalisée : les cellules du patients sont prélevées, génétiquement corrigées puis réinjectées : elles servent de médicament. Cette approche permet d'éviter les problèmes immunologiques rencontrés dans le cas d'utilisation de cellules allogéniques.

De nombreuses pistes de développement de médicaments issus d'iPS sont donc possibles dans les applications cliniques d'ingénierie cellulaire et tissulaire.

## 1.3 Médecine régénératrice pour le poumon

Après une agression de l'appareil respiratoire, qu'elle soit intrinsèque (âge, immunité, génétique) ou extrinsèque (cigarette, radiation, infection, drogue, particules,

etc.), il est connu que le poumon, tout comme le foie et le pancréas, peut - dans une certaine mesure - se réparer (Figure 25) (Beers and Morrisey, 2011). Ce phénomène permet d'envisager de des interventions thérapeutiques pour régénérer un poumon endommagé. Des nombreux travaux pour étudier la régénération du poumon chez la souris ont déjà été amorcés.



**Figure 27 : Régénération connue d'un poumon adulte (A) dans les voies aériennes inférieures et (B) dans les voies respiratoires (d'après (Beers and Morrisey, 2011)).**

Une des approches de médecine régénératrice utilisant la technologie des cellules iPS pourrait recourir à la stratégie de la greffe des cellules souches hématopoïétiques. La greffe hématopoïétique se base sur le principe de niche et celui de cellules progénitrices. Lors d'une greffe hématopoïétique, la moelle osseuse, qui est la niche des cellules souches hématopoïétiques, est nettoyée de toutes ses cellules souches mais sa structure de niche reste intacte et permet la nidation des cellules greffées ainsi que leur maturation en cellules sanguines. Ainsi, l'équipe de Yair Reisner (Rosen et al., 2015) vient de démontrer qu'il est possible de greffer dans un poumon des cellules progénitrices du stade canaliculaire après avoir préparé le poumon en le vidant d'une partie de ses cellules progénitrices tout en maintenant la niche de celles-ci. Les iPS, par la modélisation dirigée en épithélium bronchique, sont donc des candidats de choix comme source cellulaire pour une approche par thérapie cellulaire de la réparation pulmonaire (Figure 26).

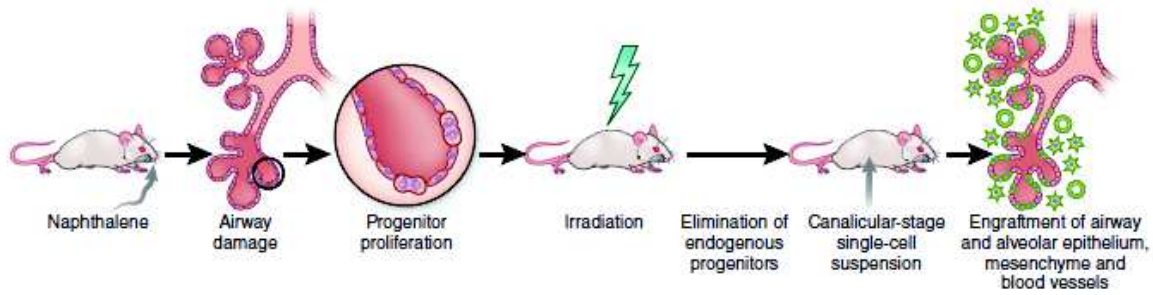


Figure 28 : Principe de la greffe de cellules progénitrices chez la souris (d’après (Snoeck, 2015))

## 1.4 Le premier essai clinique utilisant des iPS

Un premier essai clinique utilisant des cellules dérivées d’iPS a débuté chez l’homme au Japon en 2014 (Reardon and Cyranoski, 2014). Masayo Takahashi, une ophtalmologue du centre RIKEN à Kobe, a mis au point un traitement de la dégénérescence maculaire liée à l’âge par un feuillet d’épithélium pigmentaire de rétine (RPE) formé à partir d’iPS de patient (Kamao et al., 2014). Cependant, après l’inclusion d’une première patiente qui a reçu le traitement, des modifications génétiques ont été observées dans la lignée iPS préparée pour le deuxième patient. L’essai clinique a été arrêté par principe de précaution fin 2015 bien que la tumorigénicité de ces mutations ne soit pas prouvée (Garber, 2015). L’essai clinique devrait reprendre, mais sur un mode allogénique. Les cellules de RPE seront issues d’une lignée iPS fournies par le détenteur du prix NOBEL, Shinya Yamanaka, qui n’est pas dérivée à partir des cellules des patients à traiter.

La capacité de fournir en quantité illimitée des tissus d’intérêt peut également être utilisée d’un point de vue industriel.

## 2. Modélisation des tissus humains pour l'industrie pharmaceutique

Les CSP ont créé une avancée technologique majeure dans le milieu de la recherche biomédicale, et bientôt dans la recherche pharmaceutique.

Le développement de médicaments est un procédé long et très onéreux. Le taux d'échec reste important tout au long du processus du développement d'un médicament, et un échec est d'autant plus coûteux qu'il intervient plus tardivement. Une cause d'arrêt particulièrement tardif est une toxicité ou un manque d'efficacité lors des essais cliniques. Cela se produit assez fréquemment car les modèles qui étaient utilisés jusqu'à présent – lignées immortalisées, modèles animaux – ne reflètent que très imparfaitement le tissu humain ciblé par le médicament ou touché par effet collatéral du fait d'une toxicité. Les CSP auront une place de choix en offrant la possibilité de modéliser des pathologies simples, comme la mucoviscidose, et d'autres plus complexes, telle que la bronchopathie chronique obstructive (BPCO), pour tester l'efficacité des médicaments candidats. Ce type d'approche a déjà été utilisée avec succès dans le domaine neurologique par exemple (Marchetto et al., 2010; Naryshkin et al., 2014). De plus, les tissus normaux permettront de tester la toxicité médicamenteuse *in vitro*, en particulier pour des organes sensibles tels que le foie ou le cœur (Burridge et al., 2016; Kolaja, 2014).

### 2.1 Modéliser des pathologies

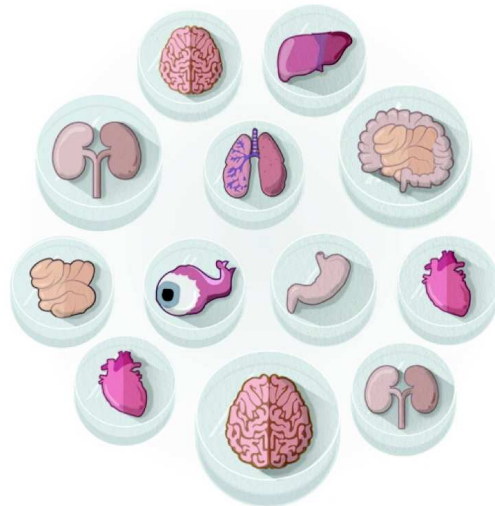
Les CSP représentent également un intérêt majeur au niveau des industries pharmaceutiques. Il est en effet possible de modéliser le développement embryonnaire normal, ou d'un tissu spécifique, en suivant les protocoles de différenciation dirigée à partir d'ES ou iPS exempts de maladie génétique.

De la même manière, comme nous l'avons évoqué dans la Partie 3 Chapitre 2.1, si les cellules pluripotentes de départ sont porteuses d'une maladie génétique, il sera possible

d'étudier le développement anormal du tissu touché par la mutation. Ceci permettra de produire en quantité quasi illimitée le tissu malade *in vitro* et ainsi modéliser les pathologies résultant de ces anomalies génétiques. Il a ainsi été possible de reproduire *in vitro* plusieurs pathologies : des iPS ont été reprogrammées à partir d'un patient porteur d'un syndrome QT long puis différenciées en cardiomyocytes avec une activité pace-maker reproduisant la pathologie du patient donneur (Moretti et al., 2010). D'autres pathologies mono et multigéniques ont ainsi pu être reproduites *in vitro*, telles que des cellules musculaires du syndrome de Hutchinson-Gilford Progeria (Zhang et al., 2011), des hépatocytes avec un déficit en alpha1-antitrypsine modélisant d'hypercholestérolémie de type Ia (Rashid et al., 2010), des cardiomyocytes hypertrophiés du syndrome LEOPARD (Carvajal-Vergara et al., 2010), des progéniteurs neuraux du Syndrome de Rett (Marchetto et al., 2010; Muotri et al., 2010), jusqu'à des cellules hématopoïétiques de l'anémie de Fanconi (Raya et al., 2009). Cette approche est appelée par les anglo-saxons « disease modelling in a dish », modéliser la pathologie dans un puits.

## 2.2 Produire des organes

Cependant, les pathologies ainsi modélisées dans les puits sont en deux dimensions. La structure 2D ne reflète pas parfaitement la physiopathologie des organes qui sont structurés en 3D. Nous vivons dans un monde à trois dimensions, et plusieurs tissus composent un même organe. L'organisation, les interactions et les comportements d'un tissu sont influencés par les tissus qui l'entourent au sein d'un même organe, et même des organes voisins au sein du même organisme. Une étape a été franchie récemment en ajoutant un niveau de complexité pour atteindre la 3D. C'est ainsi que sont produits des mini-organes fonctionnels dans les laboratoires (Marx, 2015) (Figure 27).



**Figure 29 : illustration de la production d'organes dans un puits (d'après (Willyard, 2015)).**

La technique de culture des organoïdes est en plein développement et a été appliquée avec succès à d'autres organes. La liste des organes recréés *in vitro* à partir de CSP s'allonge régulièrement : un mini-cortex de cerveau (Eiraku et al., 2008), un mini-œil (Eiraku et al., 2011), une mini glande pituitaire (Suga et al., 2011), un mini-estomac (McCracken et al., 2014), un mini-intestin (Spence et al., 2011; Watson et al., 2014), un mini-rein (Takasato et al., 2014), un mini-foie (Si-Tayeb et al., 2010; Takebe et al., 2013) et un mini-pancréas (Hrvatin et al., 2014). Le principe général est de placer des CSP en cours de différenciation dans une matrice 3D mimant la lame basale tel que le Matrigel® (voir Partie 2, Chapitre 4.2), une matrice riche en laminine et collagène IV extraite à partir de la tumeur murine Engelbreth-Holm-Swarm. Appliqué au domaine pulmonaire, cela permet de produire des organoïdes humains de poumon (Human Lung Organoid, HLO). Les HLO sont des mini-poumons immatures, dont la composition cellulaire est similaire à celle des poumons fœtaux (Dye et al., 2015; Konishi et al., 2016). Cette immaturité est retrouvée *in vitro* dans les organoïdes mimant les autres organes et présente à ce jour une des limites de cette technologie. Les organoïdes n'en sont pas moins très prometteurs car leur organisation est plus proche du tissu natif que les cultures 2D. Récemment, des organoïdes produits par des cellules souches adultes intestinales ont été utilisés avec succès pour le

criblage de molécules pour la réparation fonctionnelle de certaines mutations du gène responsable de la mucoviscidose CFTR (Dekkers et al., 2013).

Il est possible d'imaginer de produire dans le futur des tissus sains à partir de ces mini-organes et proposer ainsi une stratégie thérapeutique à mi-chemin entre la thérapie cellulaire et la transplantation d'organe. Ces organoïdes peuvent également être utilisés pour optimiser et raccourcir le long procédé de mise sur le marché de nouveaux médicaments.

## **2.3 Découvrir de nouveaux médicaments**

Ainsi après avoir sélectionné des patients porteurs de pathologies d'intérêt, reprogrammé leurs cellules en iPS qualifiées, il est possible avec un protocole robuste de différenciation dirigée d'obtenir un tissu d'intérêt en quantité quasi-illimitée. La genèse d'une banque de lignées iPS chacune porteuse d'une anomalie connue et caractérisée sera un atout pour ce type d'applications. Cela permettra également de tester la robustesse des protocoles de différenciation, ainsi que de comparer les phénotypes des tissus pathologiques aux phénotypes des tissus sains issus d'iPS non porteuses d'anomalie génétique.

Avoir un tissu sain et pathologique en quantité quasi-illimitée permet d'accélérer le procédé de développement des médicaments en offrant la possibilité de cribler des molécules candidat-médicament directement sur le tissu (Grskovic et al., 2011; Matsa et al., 2016; Mordwinkin et al., 2013).

Alors que dans le déroulement traditionnel de découverte de médicaments, l'humain n'intervient que tardivement, pendant les essais cliniques, les CSP permettent de travailler dès le départ sur un modèle humain porteur de la pathologie ciblée et sur le tissu ciblé par la future thérapie (Figure 28). Il sera donc possible de réaliser un criblage des millions de molécules présentes dans les bibliothèques des industriels sur des tissus humains issus des

CSP (Ebert et al., 2009; McNeish et al., 2010), tout en écartant plus rapidement les molécules toxiques (Liang et al., 2013). En effet, et contrairement au procédé classique de développement d'un médicament, l'efficacité et la sécurité pourront être testées directement sur du tissu humain sain ou pathologique pour étudier la toxicologie organe-spécifique (cardio-toxicologie sur du tissu cardiaque, l'hépatotoxicologie sur du tissu hépatique, etc.) et même apprécier les effets de la polymédication sur ces tissus *in vitro*. Enfin, une fois les candidats-médicaments retenus, les iPS seront également un outil idéal pour identifier les mécanismes et les cibles d'action (Grskovic et al., 2011).

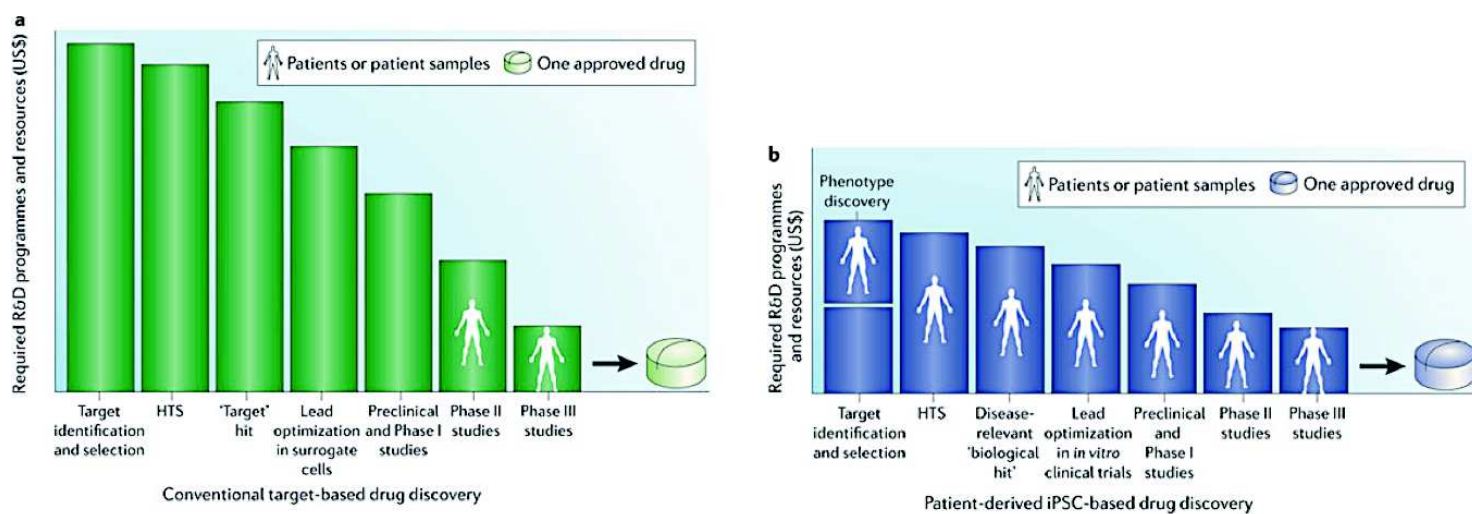


Figure 30 : Comparaison du processus conventionnel de découverte d'un médicament (A), avec le processus utilisant les cellules iPS de patient (B) (d'après (Grskovic et al., 2011)).

## 2.4 Développer la pharmacogénomique

D'un point de vue pharmaceutique, ces criblages permettront également d'approfondir nos connaissances concernant la pharmacogénomique. En effet, le polymorphisme de certains gènes peut être impliqué dans des phénomènes de toxicité ou d'inefficacité vis à vis des médicaments (Abubakar and Gan, 2016; Cacabelos et al., 2016; Chasman et al., 2004). Sont surtout impliqués les gènes ayant un rôle dans le métabolisme des principes actifs. Prenons comme exemple le cytochrome P450 2D6. Le polymorphisme de celui-ci explique qu'il existe des sujets métaboliseurs rapides, intermédiaires ou lents vis



à vis, entre autres, des antalgiques codéinés (Eichelbaum and Evert, 1996), entraînant une inefficacité du médicament ou a contrario une toxicité accrue. La pharmacogénomique tient donc une part importante dans notre nouveau mode de thérapie ciblée, permettant d'optimiser les traitements en fonction du patient receveur. Les iPS sont un atout pour étudier *in vitro* la pharmacogénomique par la production illimitée de tissus issus de patients porteurs de polymorphismes connus.

# RESULTATS

Au cours de mon travail de thèse je me suis attachée à montrer comment la technologie des pluripotentes induites (iPS) pouvait être utilisée pour améliorer notre connaissance de l'appareil respiratoire normal et pathologique, et pour le développement de nouveaux traitements. Je me suis d'abord focalisé sur les tératomes issus des iPS pour déterminer la part de différenciation dans la voie pulmonaire dans ces structures. Puis j'ai mis au point les conditions de différenciation des iPS humaines en épithélium bronchique. Afin de démontrer la preuve de concept de l'intérêt de l'utilisation d'iPS pour modéliser une pathologie pulmonaire, j'ai reprogrammé des fibroblastes d'un patient atteint de dyskinésie ciliaire et j'ai caractérisé cette lignée iPS.

J'ai mis en place la thématique de la différenciation des iPS en épithélium pulmonaire dans le laboratoire. A cet effet, j'ai séjourné un mois dans le laboratoire de Janet Rossant à Toronto pour importer les techniques qui y sont développées. Le travail que j'ai effectué pendant ma thèse s'inscrit dans un projet global du laboratoire de se spécialiser dans cette nouvelle thématique.

# 1. Les tératomes

## 1.1 Résumé

Les cellules hiPS sont obtenues par la reprogrammation de cellules différenciées en cellules souches pluripotentes. Elles sont alors capables de se différencier *in vitro* dans tous les types cellulaires. Les hiPS sont une source prometteuse de cellules afin de modéliser les tissus ou organes sains ou pathologiques selon le patient dont sont issues les hiPS. Cependant, *in vitro*, la différenciation est limitée non seulement par la culture en deux dimensions, soit monocouche, mais aussi par l'absence de vascularisation. Par contraste, et comme nous l'avons vu dans la Partie 3, chapitre 1.2, l'injection d'hiPS dans une souris fortement immunodéprimée produit un tératome. C'est une structure en trois dimensions fortement vascularisée et qui donne des tissus issus des trois feuillets embryonnaires. La formation de tissu pulmonaire dans les tératomes humains est très peu documentée (Gertow et al., 2004; Ozolek and Castro, 2011). J'ai donc procédé à une analyse de l'expression de marqueurs pulmonaires dans des tératomes issus d'hiPS que nous avons obtenus dans le laboratoire.

Les tératomes, que nous avons produits à partir d'hiPS, recèlent de nombreuses formations d'origine endodermique. Ces tubes épithéliaux résultent d'un patchwork de tissus bronchiques et digestifs. En effet, nous avons démontré qu'un même épithélium peut exprimer à la fois des marqueurs bronchiques TTF1/NKX2.1 et CK7, et des marqueurs digestifs CDX2 et CK20 de manière remarquablement interposée.

Ces résultats montrent que les tératomes issus d'hiPS sont un outil prometteur pour étudier le développement des voies aériennes humaines. Cette technique ouvre également la voie pour modéliser les pathologies pulmonaires génétiques d'une manière unique.

## **1.2 Article**

**Article 4** : « Modeling 3D human airway development using induced pluripotent stem cell-derived teratomas »

Article en cours de préparation

1 **Modeling 3D human airway development using induced pluripotent stem**  
2 **cell-derived teratomas**

3

4 Caroline Sansac<sup>1,2,3,\*</sup>, Anass Belbachir<sup>4,\*</sup>, Engi Ahmed <sup>1,2,8</sup>, Jean-Marie  
5 Ramirez<sup>1,2</sup>, Aurélie Fort<sup>5</sup>, Marie Péquignot<sup>6</sup>, Isabelle Serre<sup>7</sup>, Frédérique  
6 Penault-Llorca<sup>8</sup>, Arnaud Bourdin<sup>3,5,9</sup>, John De Vos <sup>1,2,3,10</sup>

7

8 <sup>1</sup> CHRU Montpellier, Institute for Regenerative Medicine and Biotherapy,  
9 Hôpital Saint-Eloi, Montpellier, F34000 France; <sup>2</sup> INSERM, U1183,  
10 Montpellier, F34000 France; <sup>3</sup> Université de Montpellier, UFR de Médecine,  
11 Montpellier, F34000 France; <sup>4</sup> Faculty of Medicine and Pharmacy, Cadi Ayyad  
12 University, 2390 Marrakech, Morocco; <sup>5</sup> INSERM, U1046, Montpellier, F34000  
13 France; <sup>6</sup> Institute for Neurosciences of Montpellier, Montpellier, F34000  
14 France; <sup>7</sup> CHRU Montpellier, Department of Pathology, Montpellier, F34000  
15 France; <sup>8</sup> Department of Pathology, Centre Jean Perrin, EA 4677, Université  
16 d'Auvergne Clermont 1, F63000 Clermont-Ferrand, France; <sup>9</sup> Department of  
17 Respiratory Diseases, Hôpital Arnaud de Villeneuve, Montpellier, F34000  
18 France; <sup>10</sup> CHU Montpellier, Department of Cell and Tissue Engineering,  
19 Hospital Saint-Eloi, Montpellier, F 34000 France;

20 \* These authors contributed equally to this work.

21

22 **Key words:** pluripotent stem cells, pluripotency, endoderm, teratoma, airway  
23 epithelium.

24

25

1

2

3

4 **Corresponding authors:**

5 John De Vos, Department of Cell and Tissue Engineering, Hôpital Saint-Eloi,

6 80 Avenue Augustin Fliche, 34295 Montpellier Cedex 5, France. E-mail:

7 john.devos@inserm.fr

8

9 Arnaud Bourdin, Department of Respiratory Diseases, 371 Av Doyen Giraud,

10 34295 Montpellier Cedex 5, France. E-mail: a-bourdin@chu-montpellier.fr

11

12 Running Title: airway tissues in teratomas

13

1 **Abstract**

2 Human induced pluripotent stem cells (hiPSC) are obtained by  
3 reprogramming differentiated cells into a pluripotent state. As they can  
4 differentiate in vitro into numerous (if not all) human cell types, hiPSC are a  
5 promising source of cells for modeling normal and diseased organs/tissues.  
6 However, in vitro, hiPSC differentiation is limited by the fact that only  
7 monolayers or oligo-layered structures are produced, mainly because of the  
8 absence of vasculature. By contrast, injection of hiPSC in immunosuppressed  
9 mice induces teratomas, which are 3D structures that include tissues from the  
10 three germinal layers, including endoderm. Here, we report that within  
11 teratomas, hiPSC formed typical endodermal tubes that comprised patches of  
12 airway epithelium. These epithelia showed strong expression of CK7,  
13 TTF1/NKX2.1 and CC10 that was mutually exclusive with that of CD20 and  
14 CDX2. These results show that hiPSC-derived teratomas are a promising tool  
15 to study airway development and to model hereditary lung diseases.

16



## 1 **Introduction**

2 For obvious ethical issues, human development can be studied *in utero* only  
3 with non-invasive techniques, such as echography, or *in vitro* in cultured  
4 embryos but only until the blastocyst stage. Human pluripotent stem cells  
5 (hPSC) have the capacity to differentiate into all the body cell types (12). They  
6 are derived from the blastocyst inner cell mass (human embryonic stem cells,  
7 hESC), or by reprogramming differentiated cells into human induced  
8 pluripotent stem cells (hiPSC) through forced expression of reprogramming  
9 factors, such as OCT4, SOX2, c-MYC and KLF4 (3, 15). However,  
10 differentiation of cultured hPSC produces only monolayers of cells or oligo-  
11 layered structures, in part because of the absence of vasculature. By contrast,  
12 Teratoma formation by xenografting hPSC into immunosuppressed mice  
13 produces 3D systems that can help studying human development. Indeed,  
14 teratomas benefit from mouse's vasculature to form relatively large structures.  
15 This conveys two main benefits: (1) allow a 3D development mimicking the  
16 physiological development, including complex interactions between different  
17 cell types, (2) promote a differentiation during a long time period, i.e. at least  
18 two month. Indeed, teratomas comprise numerous tissues, from all three  
19 germ layer, with often the microscopic aspects of equivalent adult fetal or  
20 human somatic tissues. However, the contribution to teratomas of lung tissue,  
21 though a major organ of endodermal origin, has only rarely been mentioned  
22 (5, 6, 10). Here we investigated into more details the extent of airway  
23 development in teratomas derived from hiPSC.

## 24 **Materials and method**

## 1 **Cell culture**

2 Human foreskin fibroblasts (HFF) and adult fibroblasts (FMO2) were cultured  
3 with DMEM supplemented with 10% KO-SR (Gibco, Life Technologies®).  
4 M4C7 was cultured under standard feeder conditions on irradiated HFF.  
5 iPCD02-30 was cultured under feeder-free conditions on Matrigel (BD  
6 Biosciences) in mTESR1 (StemCell Technologies) or on Geltrex in E8 (Gibco,  
7 Life Technologies®).

## 8 **hiPSC generation**

9 We used the two different hiPSC lines: M4C7 and iPCD02-30. M4C7 was  
10 obtained by reprogramming human foreskin fibroblasts (hFFs) using lentiviral  
11 vectors that encode for the open reading frames (ORFs) of human POU5F1,  
12 SOX2, NANOG and LIN28 (14). We generated and characterized ten hiPSC  
13 clones, including M4C7. iPCD02-30 was obtained by reprogramming human  
14 adult fibroblasts FMO2 from a patient with PCD (primary ciliary dyskinesia) -  
15 with a mutation in the CCDC39 gene, as described in (7). We generated thirty  
16 hiPSC clones, including iPCD02-30.

17 FACS and QRTPCR analysis demonstrated that both M4C7 and iPCD02-30  
18 had the typical hiPSC expression profile characterized by expression of cell  
19 surface pluripotency markers (TRA-1-60, TRA-1-81, SSEA3, SSEA4) and key  
20 transcription factors (for instance, OCT4 and NANOG). After injection in  
21 immunocompromised mice, they produced teratomas that included tissues  
22 from all three germ layers (13).

## 23 **Teratoma formation**

24 Undifferentiated cells were dissociated (Trypsin or Versen), suspensions  
25 ( $1 \times 10^7$ ) were mixed with Matrigel (BFD Biosciences) and injected  
26 subcutaneously into NOD – SCID – Gamma (NOD.Cg-Prkdc<sup>scid</sup> Il2rg<sup>tm1Wjl</sup>/SzJ,

1 NSG) mice under anesthesia (11). Mice were sacrificed after seven to ten  
2 weeks after the injection. Teratoma were collected, fixed in 4%  
3 paraformaldehyde and embedded in paraffin.

#### 4 ***Immunohistochemistry***

5 Paraffin-embedded teratoma sections were deparaffinised with xylene and  
6 rehydrated using an alcohol gradient (100% ethanol, 95% ethanol, 70%  
7 ethanol, 50% ethanol). The IHC slides were subjected to heat-induced antigen  
8 retrieval in Tris-EDTA buffer (Tris 100 mM, EDTA 1mM and Tween20, pH =  
9 9.0), blocked and incubated 1hr with primary antibodies (NKX2.1/TTF1, Dako  
10 M3575; CC10, Biovendor RD181022220; CK20, Abcam ab15205).  
11 Immunohistochemical staining was visualized using the avidin-biotin system  
12 (Dako, LSAB2 System-HRP) in combination with 3-3'-diaminobenzidine  
13 (DAB). Ki67 xxxx AML xxxx

#### 14 **Results**

##### 15 **Teratomas exhibit numerous tubular structures surrounded by** 16 **mesenchymal tissue**

17 All hiPSC-derived teratomas displayed tubular structures (Figure 1A)  
18 composed of cuboidal or pseudostratified columnar epithelium as indicated by  
19 the tightly packed cells that formed continuous sheets. The simple or  
20 pseudostratified columnar epithelium organization and the frequent presence  
21 of goblet cells indicated that these structures were of endodermal origin. The  
22 number of tubes varied from 3 to 59 per section, depending on the teratoma  
23 and the section position within the teratoma. These tubes were apparently  
24 randomly associated with other tissues types such as cartilage (Figure 1C).

1 Noteworthy, cartilage is an important component of the upper respiratory tract  
2 The endodermal tubes were mostly positive for the Ki67 markers, indicating  
3 an intense proliferation, witness of immature, developing structure (Figure  
4 1D). In line with this immaturity, we noted that neural tissue appeared  
5 frequently and was immature, consisting of neural rosette (Figure 1B). These  
6 epithelia developed on an underlying, vimentin-positive connective tissue  
7 layer (Figure 2). The interface between the epithelium and the subjacent  
8 mesenchymal tissue was very well delimited, demonstrating the highly  
9 organized developing tissue (Figure 2B). Moreover, a smooth muscle actin  
10 labeling showed that smooth muscle cells surrounded the endodermal tubes,  
11 as typically observed in airway and digestive tracts (Figure 2D).

12

### 13 **Opposing expression patterns of airway and digestive markers**

14 As endoderm gives rise to the epithelia that line the digestive and respiratory  
15 tracts, we studied the expression of cytokeratin-7 (CK7), a cytoplasmic marker  
16 expressed in various epithelia, such as simple ductal epithelium, mesothelium  
17 and pseudostratified epithelium, but not in intestinal epithelium (8),  
18 cytokeratin-20 (CK20), a cytoplasmic marker of the gastrointestinal tract (1),  
19 thyroid transcription factor-1 (TTF1/NKX2.1) involved in lung, thyroid and  
20 forebrain specification and CDX2, a nuclear homeobox transcription factor  
21 expressed from the proximal duodenum to the distal rectum (midgut and  
22 hindgut) from week 8 of gestation and in postnatal life (17). CDX2 and CK20  
23 are considered good markers of intestinal origin of metastatic  
24 adenocarcinoma, and CK7 of pulmonary origin in TTF1/NKX2.1-  
25 positive/CK20-negative tumor samples (16, 17). In a first analysis of four

1 M4C7-derived teratomas, most tubes expressed at least one or both  
2 cytokeratins (74.0%) as well as TTF1/NKX2.1 or/and CDX2 (75.6%) (Table 1).  
3 Specifically, 41.8% were CK7-positive and 22% were TTF1/NKX2.1-positive,  
4 while 47.2% were CK20-positive and 70.7% CDX2-positive. CK7 often  
5 overlapped with TTF1/NKX2.1 and CK20 with CDX2, indicating the bronchial  
6 nature of the CK7/TTF1-positive epithelia and the digestive tract nature of  
7 CK20/CDX2-positive epithelia (Figure 3 and 4). Strikingly, we repeatedly  
8 observed the simultaneous, but not overlapping expression in the same tube  
9 of CDX2 and TTF1/NKX2.1 (Figure 4A) or of CK20 and CK7 (Figure 3).  
10 However, CK7 or TTF1/NKX2.1 never overlapped with CK20 or CDX2.  
11 Furthermore, CK7 and CK20 expression were always mutually exclusive  
12 (Figure 3). The proportion of CK7-positive, CK20-positive or CK20/CK7-  
13 positive tubes varied greatly within each and among teratomas, suggesting a  
14 chaotic determinism of the dominant differentiation direction in the endoderm  
15 tubes.

16

### 17 **Expression of mature airway markers in teratomas**

18 We next tested the expression of SCGB1A1/CC10, a mature bronchial marker  
19 of club cells. SCGB1A1/CC10 specifically stained epithelial cells in  
20 TTF1/NKX2.1-positive, but never in TTF1/NKX2.1-negative areas (Figure 5A).  
21 Finally, we observed the expression of the mucin family member epithelial  
22 membrane antigen (EMA)/MUC1 in CK7 positive tubes (Figure 5B). Thus,  
23 teratomas included endoderm-derived epithelia simultaneously engaged in  
24 digestive tract and also in airway differentiation with high frequency.

## 1 **Discussion**

2 Thus, hiPSC-derived teratomas can be used for 3D modelling of human  
3 bronchial epithelium differentiation as they include highly organized airway  
4 epithelial structures surrounded by vascularized mesenchyme and cartilage.  
5 As the hiPSC technology allows choosing the genetic background of the  
6 starting cell donor, or using the CRISP/cas9 technology to knock down  
7 developmental genes, it will be possible to mimic human embryonic airway  
8 development defects or the early steps of genetic diseases. Modeling normal  
9 or diseased human tissues is one of the nine applications of hiPSC (3), and  
10 teratoma offers some advantages over common 2D differentiation protocols.  
11 The 3D architecture obtained in teratoma is notably closer to that of human  
12 tissues than that obtained by the 2D models. Its microscopic architecture  
13 resembles that of their normal tissue counterpart. The complex and organized  
14 interconnection between the endodermal epithelium and its underlying  
15 mesenchymal vimentin positive tissue is a good illustration of this feature of  
16 teratomas. While organoids offer interesting 3D, in vitro differentiated tissues,  
17 including examples of airway cells mixed with mesenchymal cells (4), the  
18 volume of tissue offered by teratomas is notably bigger, suggesting a better  
19 differential potential.

20 These observations suggest that teratomas recapitulate the dynamic of the  
21 early human development, such as gastrulation and the formation of  
22 endodermal tubes. Indeed, the intimate development of vimentin-positive  
23 mesenchymal tissue with the endodermal tubes recalls the normal airway  
24 structure development in which the mesenchyme plays a crucial role (18).  
25 Hence, hiPSC-derived teratomas will be a precious tool to study these

1 essential early steps of the airway development, for instance by using hiPSC  
2 for which specific growth factors or transcription factors have been genetically  
3 deleted.

4 Though we observed numerous places of lung differentiation initiation, and  
5 though we observed differentiated cell types such as club cells and goblets  
6 cells, we did not observe ciliated cells. These observations are in line with the  
7 fact that most often, the tissues observed in PSC-derived teratomas are  
8 immature (2). This is illustrated, for instance, by the occurrence of neural  
9 tubes. Therefore, we were not able to evaluate whether the PCD hiPSC line  
10 iPCD02-30 recapitulated the cilia defect of the patient from whom the line was  
11 generated. Nevertheless, though such a terminal differentiation into ciliated  
12 cells is rare, it has nonetheless already been reported by Lensch et al. (6).

13 However, as pointed by Damjanov and Ahrlund-Richter, bona fide  
14 organogenesis is not taking place in teratomas (2). The different tissues are  
15 scattered with no global organization. We observed the immediate  
16 cohabitation of immature neural tubes with endodermal tube. One teratoma  
17 often contains many different endodermal tubes while only one tube is  
18 generated during normal development – though it cannot be excluded that  
19 these tubes pertain to a unique tube that meanders in the tumor and is thus  
20 observed several time on the same section. In line with the absence of the  
21 global organization that govern embryogenesis, we observed the expression  
22 of NKX2.1, a transcription factor of the foregut and that of CDX2, a  
23 transcription factor expressed in the mid- and hindgut in the close proximity in  
24 the same tube. Therefore, unless it is possible to manipulate the global  
25 organization of the differentiation processes, it will not be possible to use

1 hiPSC-derived teratomas to model complex developmental stem such lung  
2 branching.

3 Evidently, teratoma formation is not easily amenable to experimental  
4 intervention to guide the differentiation process, in contrast to the 2D or the  
5 organoids models. Nonetheless, it is conceivable to guide teratoma formation  
6 toward the lung pathway by at least two means. Firstly, instead of injecting  
7 undifferentiated PSC to immunodeficient mice, prior differentiation into a  
8 specific organ direction can be carried out. This strategy was used by Mou et  
9 al. (9). They differentiated hiPSC (mice or human) into NKX2.1+ cells that  
10 were injected subcutaneously in immunodeficient mice and formed  
11 “tracheospheres” containing terminally differentiated CC10+ club cells,  
12 FOXJ1+ ciliated cells or MUC5A1 goblet cells (for mice cells) or P63+  
13 progenitor, basal cells (for human cells). Secondly, it is may be possible to  
14 use soaked microbeads as a device to deliver over time specific growth  
15 factors or chemical inhibitors. Such signal may reproduce some early events  
16 taking place during embryonic life and direct organogenesis towards a desired  
17 tissue type.

18 In conclusion, we have showed that human hiPSC-derived teratomas  
19 recapitulate early steps of lung development. The histology of these airways  
20 display a marked analogy with their counterpart found during normal  
21 development. Teratomas are therefore tools to study the mechanism of  
22 normal and altered early airway development.

23

24

25 **Acknowledgement**



1 Supported by grants from the University Hospital of Montpellier (Appel d'offre  
2 interne, projet 9174).

3

#### 4 DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST

5

6 The authors indicate no potential conflicts of interest.

7

#### 8 **References**

9

- 10 1. **Botta MC, Ambu R, Liguori C, Van Eyken P, Pisanu A, Cabras A,**  
11 **Hofler H, Werner M, Faa G.** [CK20 expression in the gastrointestinal  
12 tract of the embryo and fetus]. *Pathologica* 93: 640–644, 2001.
- 13 2. **Damjanov I, Ahrlund-Richter L.** Characterization of Human Embryonic  
14 Stem Cell-derived Teratomas. In:., edited by Loring JF, Wesselschmidt  
15 RL, Schwartz PH. *Human Stem Cell Manuel*, 2013, p. 171–182.
- 16 3. **De Vos J, Bouckenheimer J, Sansac C, Lemaître J-M, Assou S.**  
17 Human induced pluripotent stem cells: A disruptive innovation. *Curr Res*  
18 *Transl Med* 64: 91–96, 2016.
- 19 4. **Dye BR, Hill DR, Ferguson MAH, Tsai Y-H, Nagy MS, Dyal R, Wells**  
20 **JM, Mayhew CN, Nattiv R, Klein OD, White ES, Deutsch GH,**  
21 **Spence JR.** In vitro generation of human pluripotent stem cell derived  
22 lung organoids. *Elife* 4, 2015.
- 23 5. **Gertow K, Wolbank S, Rozell B, Sugars R, Andäng M, Parish CL,**  
24 **Imreh MP, Wendel M, Ahrlund Richter L.** Organized development  
25 from human embryonic stem cells after injection into immunodeficient  
26 mice. *Stem Cells Dev* 13: 421–435, 2004.
- 27 6. **Lensch MW, Schlaeger TM, Zon LI, Daley GQ.** Teratoma formation  
28 assays with human embryonic stem cells: a rationale for one type of  
29 human-animal chimera. *Cell Stem Cell* 1: 253–258, 2007.
- 30 7. **Macarthur CC, Fontes A, Ravinder N, Kuninger D, Kaur J, Bailey M,**  
31 **Taliana A, Vemuri MC, Lieu PT.** Generation of human-induced  
32 pluripotent stem cells by a nonintegrating RNA Sendai virus vector in

- 1 feeder-free or xeno-free conditions. *Stem Cells Int* 2012: 564612, 2012.
- 2 8. **Moll R, Divo M, Langbein L.** The human keratins: biology and  
3 pathology. *Histochem Cell Biol* 129: 705–733, 2008.
- 4 9. **Mou H, Zhao R, Sherwood R, Ahfeldt T, Lapey A, Wain J, Sicilian L,**  
5 **Izvolosky K, Musunuru K, Cowan CA, Rajagopal J.** Generation of  
6 Multipotent Lung and Airway Progenitors from Mouse ESCs and  
7 Patient-Specific Cystic Fibrosis iPSCs. *Cell Stem Cell* 10: 385–397,  
8 2012.
- 9 10. **Ozolek JA, Castro CA.** Teratomas Derived from Embryonic Stem Cells  
10 as Models for Embryonic Development, Disease, and Tumorigenesis.  
11 Kallos, Michael S. InTech, 2011.
- 12 11. **Prokhorova TA, Harkness LM, Frandsen U, Ditzel N, Schröder HD,**  
13 **Burns JS, Kassem M.** Teratoma formation by human embryonic stem  
14 cells is site dependent and enhanced by the presence of Matrigel. *Stem*  
15 *Cells Dev* 18: 47–54, 2009.
- 16 12. **Ramirez JM, Bai Q, Dijon-Grinand M, Assou S, Gerbal-Chaloin S,**  
17 **Hamamah S, De Vos J.** Human pluripotent stem cells: From biology to  
18 cell therapy. *World J Stem Cells* 2: 24–33, 2010.
- 19 13. **Ramirez JM, Bai Q, Pequignot M, Becker F, Kassambara A, Bouin**  
20 **A, Kalatzis V, Dijon-Grinand M, De Vos J.** Side Scatter Intensity Is  
21 Highly Heterogeneous in Undifferentiated Pluripotent Stem Cells and  
22 Predicts Clonogenic Self-Renewal. *Stem Cells Dev* (March 5, 2013).  
23 doi: 10.1089/scd.2012.0658.
- 24 14. **Ramirez JM, Gerbal-Chaloin S, Milhavet O, Bai Q, Becker F, Assou**  
25 **S, Lemaitre J-M, Hamamah S, De Vos J.** Brief report: benchmarking  
26 human pluripotent stem cell markers during differentiation into the three  
27 germ layers unveils a striking heterogeneity: all markers are not equal.  
28 *Stem Cells* 29: 1469–1474, 2011.
- 29 15. **Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K,**  
30 **Yamanaka S.** Induction of pluripotent stem cells from adult human  
31 fibroblasts by defined factors. *Cell* 131: 861–872, 2007.
- 32 16. **Tot T.** Cytokeratins 20 and 7 as biomarkers: usefulness in  
33 discriminating primary from metastatic adenocarcinoma. *Eur J Cancer*  
34 38: 758–763, 2002.
- 35 17. **Werling RW, Yaziji H, Bacchi CE, Gown AM.** CDX2, a highly sensitive  
36 and specific marker of adenocarcinomas of intestinal origin: an  
37 immunohistochemical survey of 476 primary and metastatic  
38 carcinomas. *Am J Surg Pathol* 27: 303–310, 2003.
- 39 18. **Zorn AM, Wells JM.** Vertebrate Endoderm Development and Organ  
40 Formation. *Annu Rev Cell Dev Biol* 25: 221–251, 2009.

1

## 2 **Figure Legend**

### 3 Figure 1

4 General organization of endodermal tubes in hiPSC-derived teratomas.

5 All sections were counterstained with hematoxylin, eosin and safran  
6 (HES). When indicated, antibody binding was visualized using the avidin-  
7 biotin system (Dako, LSAB2 System-HRP) in combination with 3-3'-  
8 diaminobenzidine.

9 A. Whole section of teratoma 13F0358 stained with an anti-CDX2 antibody.

10 B. Neural fetal tubes. Teratoma 104.

11 C. Proximity of cartilage with an endodermal tube. Teratoma 13F0358

12 D. An endodermal tube stained with an anti-Ki67 antibody showing intense  
13 cell proliferation. Teratoma 13F0358

14

### 15 Figure 2

16 Interplay between endodermal tubes epithelium and their underlying  
17 mesenchymal tissue.

18 A. B. C. Sections from teratoma 13F0358, stained with an anti-vimentin  
19 antibody. B is an enlargement of the tube displayed in A.

20 D. Section from teratoma 13F0517 stained with an anti-actin smooth muscle  
21 (AML) antibody to highlight smooth muscle.

22

23

### 24 Figure 3

1 Opposing patterns of expression of airway and digestive cytokeratins  
2 A. B. Consecutive sections of teratoma 13F0358 were stained, respectively,  
3 with anti-CK20 (Dako M7019) and -CK7 (Dako M7018) antibodies. They  
4 show the mutually exclusive expression patterns of CK7 (as marker of  
5 airway tissue differentiation) with CK20 (as marker of digestive tissue  
6 differentiation).

7

8 Figure 4

9 CDX2 and NKX2.1 expression in endodermal tubes.

10 A –C: Consecutive sections stained with anti-CDX2 and -TTF1/NKX2.1  
11 antibodies. They show the non-overlapping expression of TTF1/NKX2.1 (as  
12 marker of airway tissue differentiation) and CDX2 (as marker of digestive  
13 tissue differentiation) in the same endodermal tube or exclusive expression in  
14 other tubes. Teratoma 103: A; Teratoma 104 : B and C.

15

16 Figure 5

17 Late airway marker expression

18 A: Two consecutive sections of teratoma 13F0517 stained with anti-  
19 TTF1/NKX2.1 and anti-CC10 (a mature bronchial marker of club cells)  
20 antibodies. CC10 staining is evidenced in the TTF1/NKX2.1 positive zone.

21 B: Two consecutive sections of teratoma 13F0351 stained with an anti-EMA  
22 antibody, showing the expression of mucin in a CK7 positive endodermal  
23 tube.

24

25

26

1 **Table**

2 Table 1

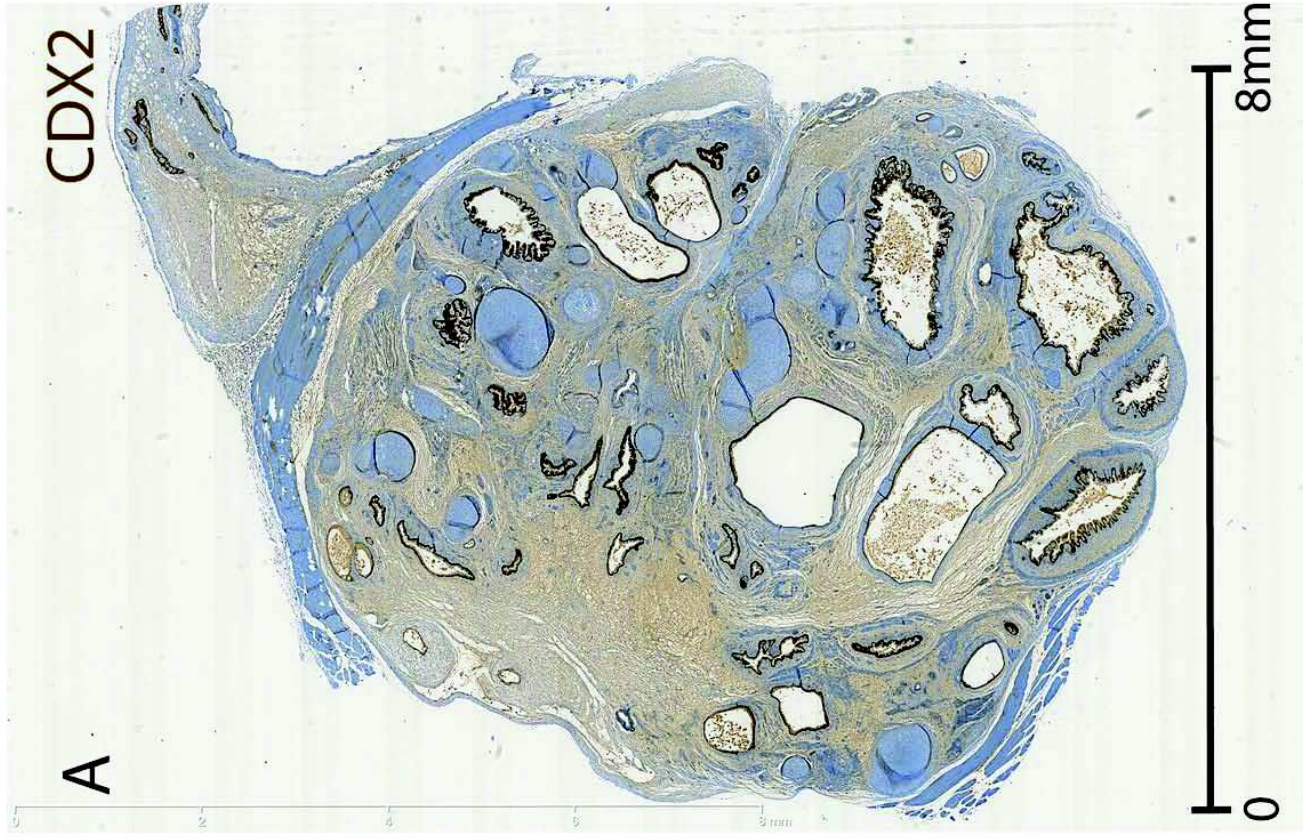
3 Number (and percentage) of tubular structures that express CK7 or/and CK20  
 4 (in all four M4C7-derived teratomas), or TTF1/NKX2.1 and/or CDX2 (only in  
 5 teratomas 13F0351 and 13F0353). Tubular structures were counted in one  
 6 representative section for each teratoma. The percentages indicate the  
 7 fraction of tubes that express a specific protein relative to all the tubes in that  
 8 representative section

<b>Staining</b>	<b>13F0351</b>	<b>13F0353</b>	<b>13F0355</b>	<b>13F0358</b>	<b>TOTAL</b>
<b>CK7+/CK20 -</b>	1	15	7	11	34
	(33.3%)	(75.0%)	(15.6%)	(18.6%)	(26.8%)
<b>CK7-/CK20+</b>	1	1	20	19	41
	(33.3%)	(5.0%)	(44.4%)	(32.2%)	(32.2%)
<b>CK7+/CK20+</b>	0	3	6	10	19
	(0.0%)	(15.0%)	(13.3%)	(16.9%)	(15.0%)
<b>CK7-/CK20-</b>	1	1	12	19	33
	(33.3%)	(5.0%)	(26.7%)	(32.2%)	(26.0%)
<b>Total tubes</b>	<b>3</b>	<b>20</b>	<b>45</b>	<b>59</b>	<b>127</b>
<b>TTF1+/CDX2</b>	-	-	1	1	2
-			(9.1%)	(3.3%)	(4.9%)
<b>TTF1-/CDX2+</b>	-	-	6	16	22
			(54.5%)	(53.3%)	(53.6%)
<b>TTF1+/CDX2+</b>	-	-	1	6	7
			(9.1%)	(20.0%)	(17.1%)

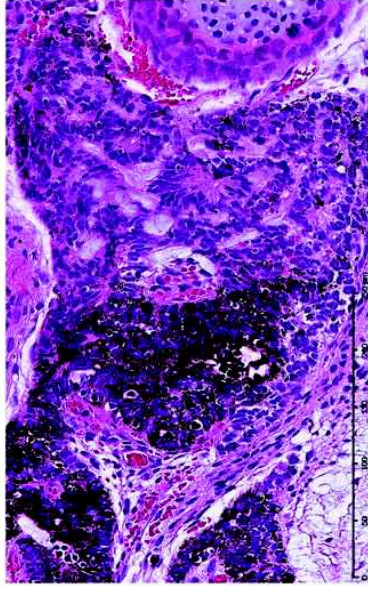
<b>TTF1-/CDX2-</b>	-	-	3 (27.3%)	7 (23.3%)	10 (24.4%)
<b>Total tubes</b>	-	-	11	30	41

1

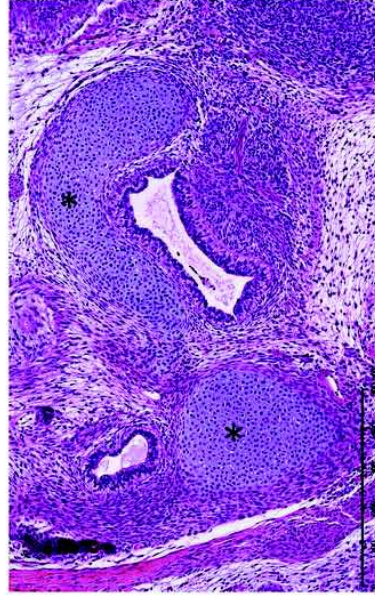
Figure 1



**B**



**C**



**D**

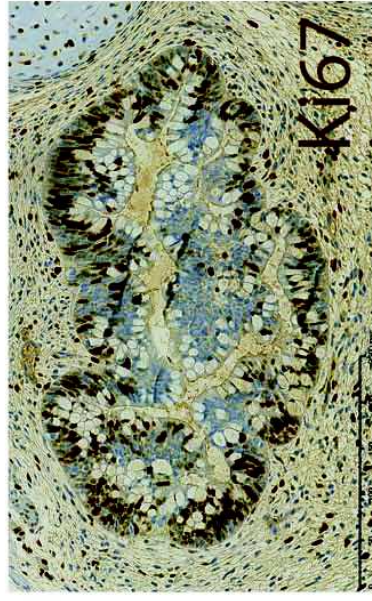


Figure 2

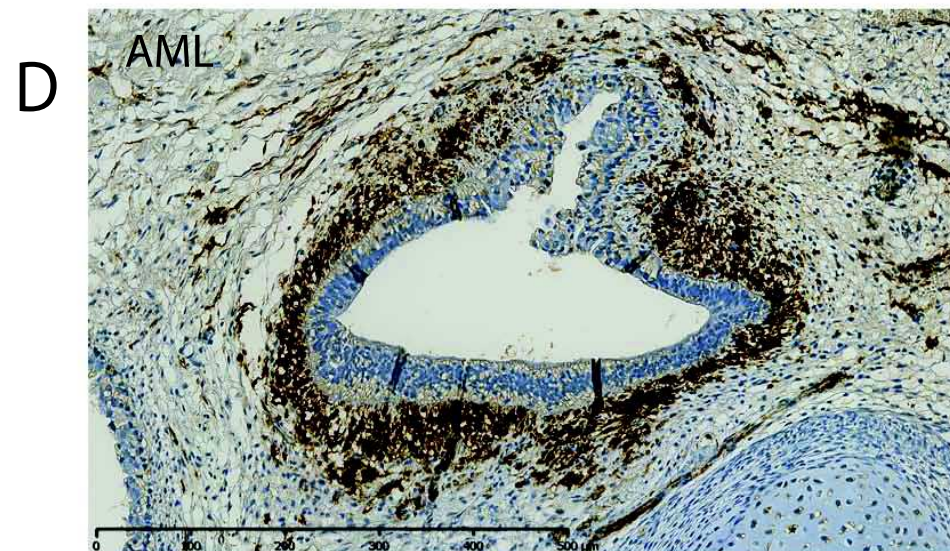
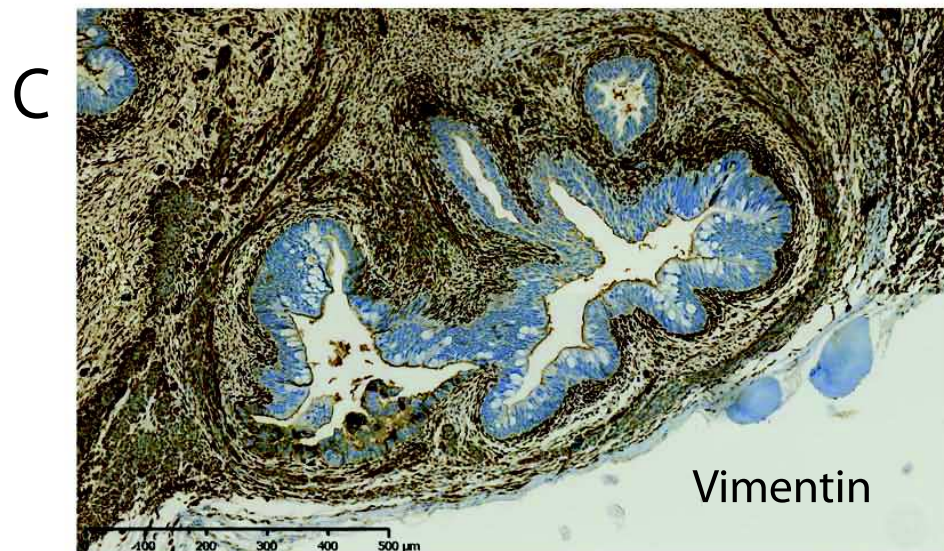
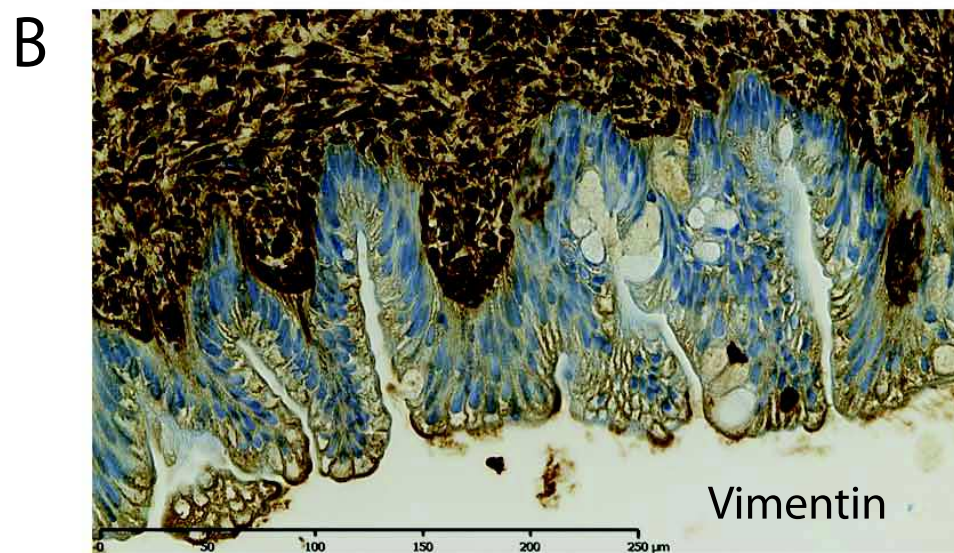
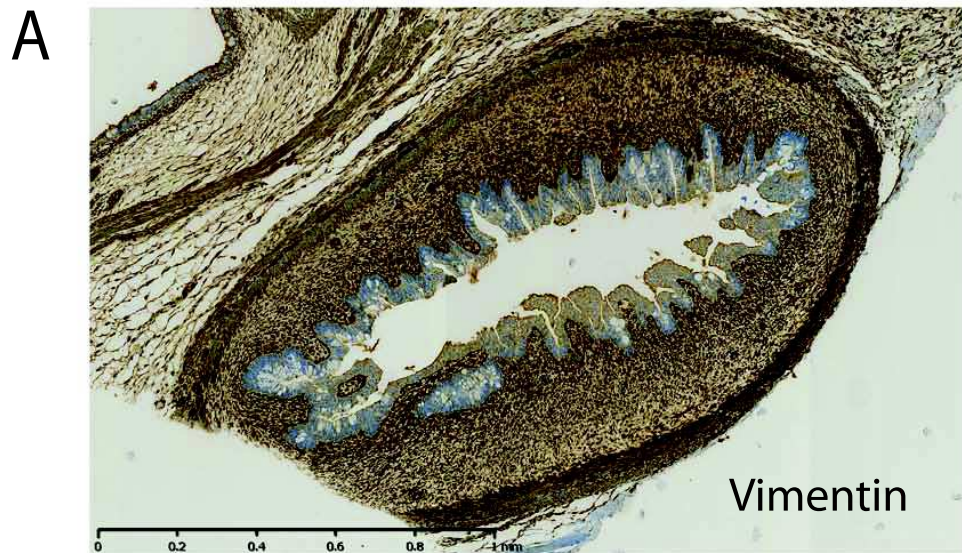




Figure 3

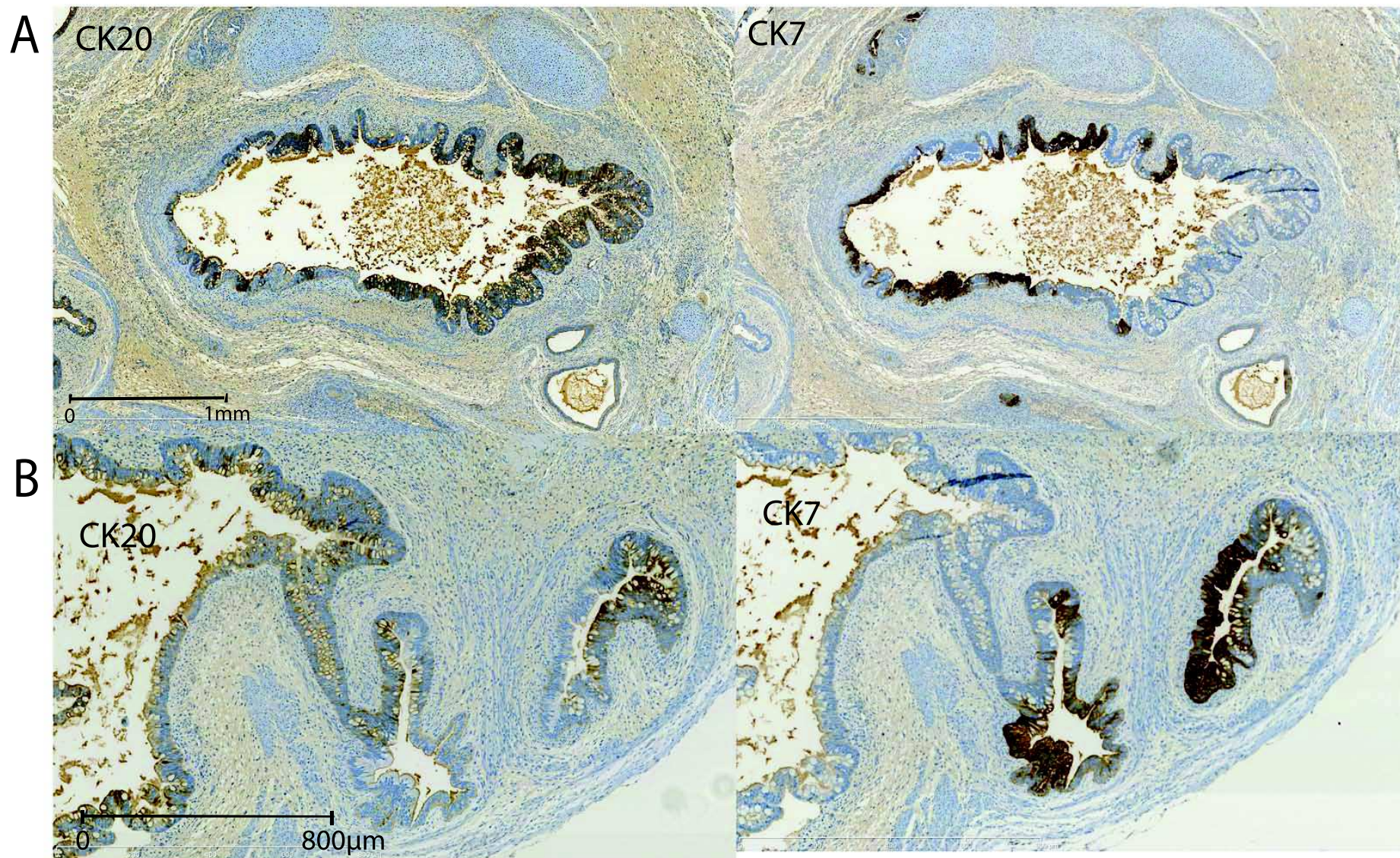


Figure 4

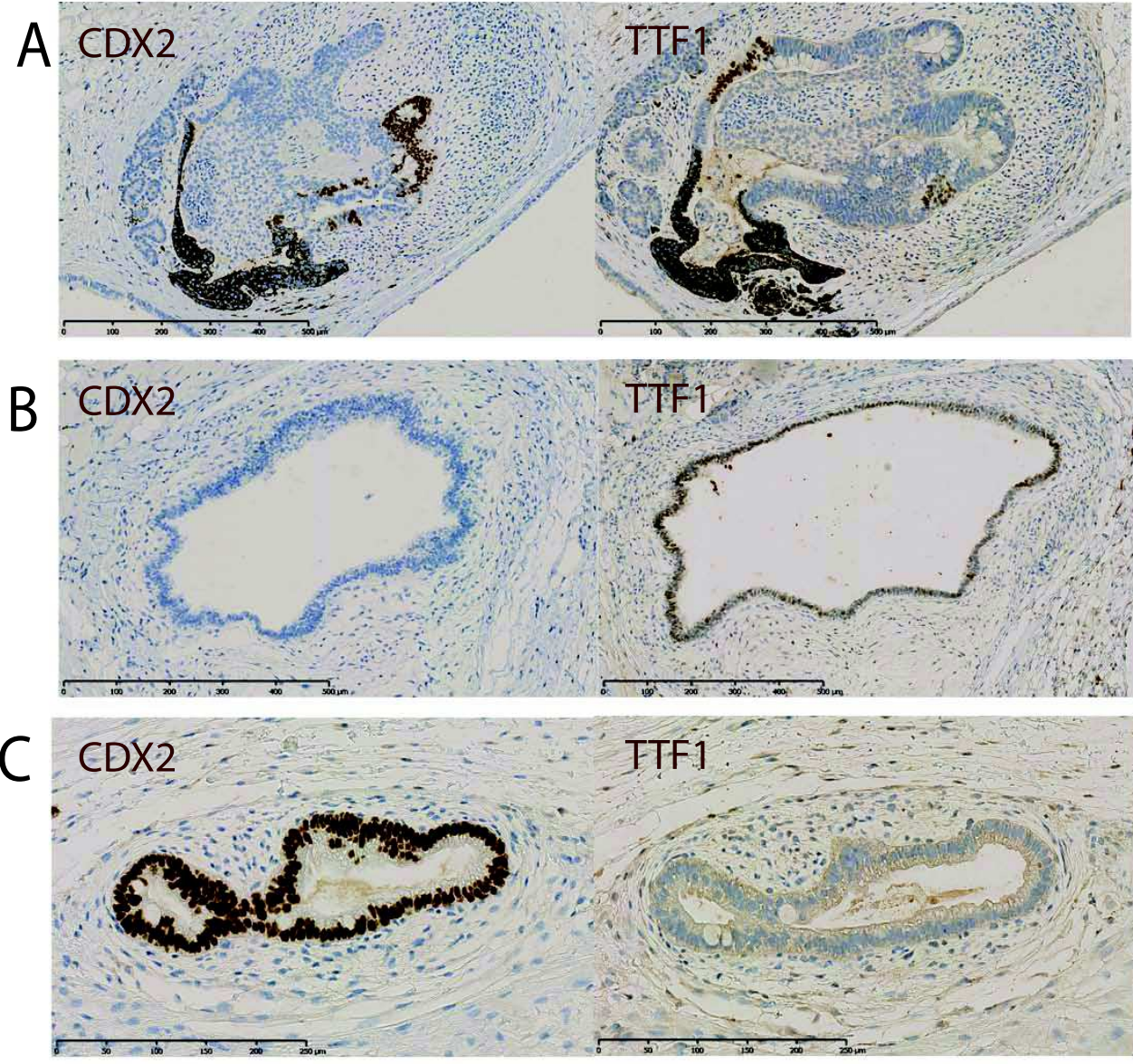
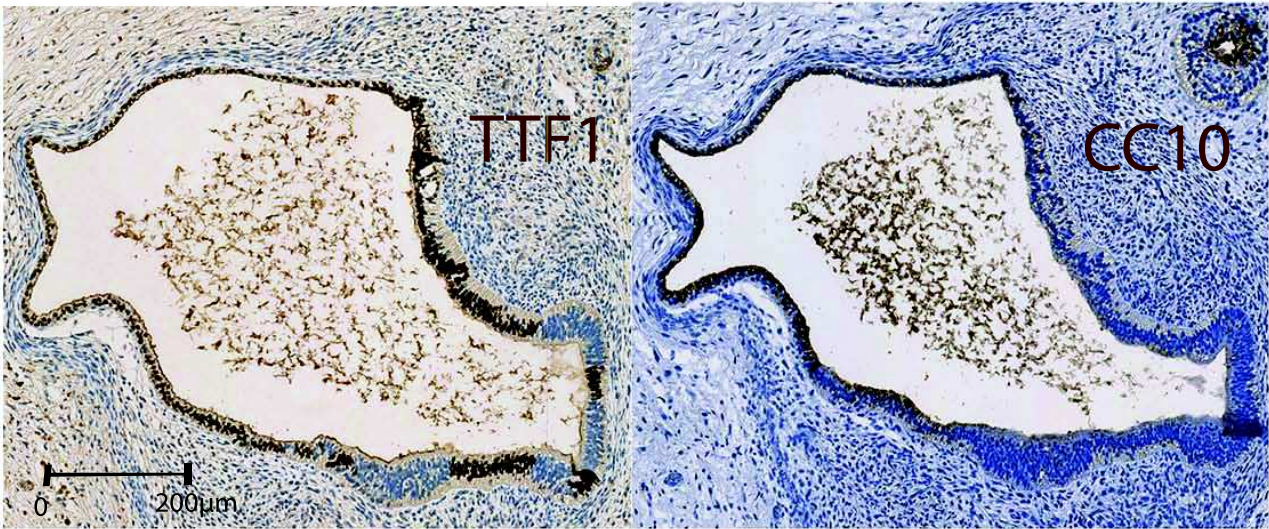
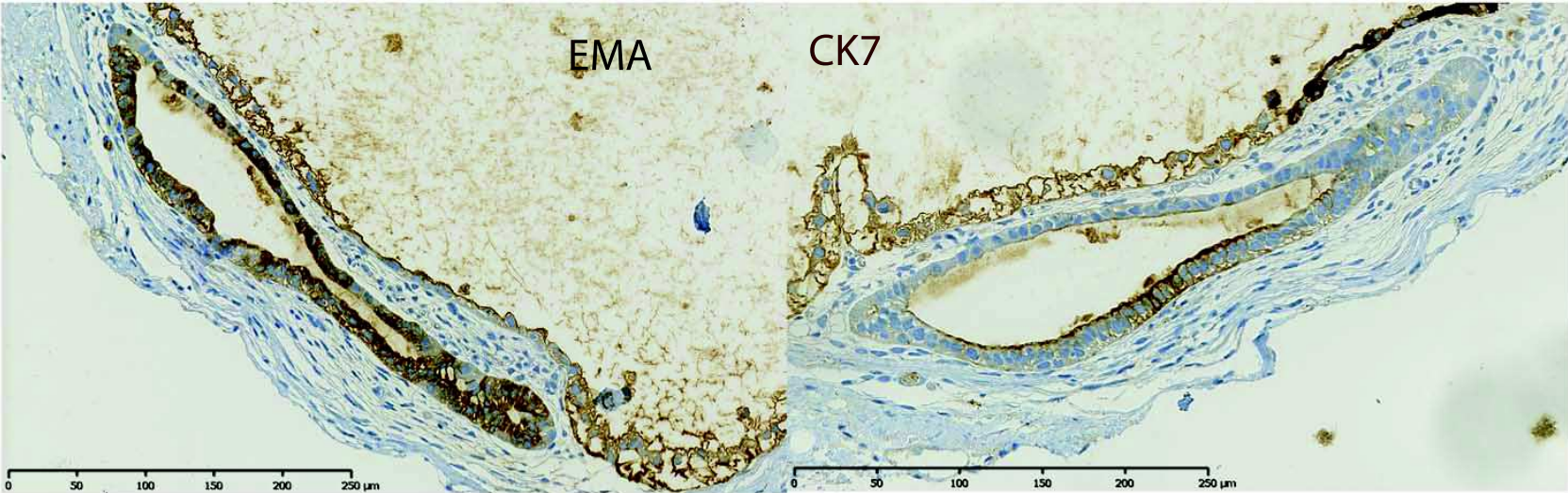


Figure 5

A



B



### 1.3 Discussion

La différenciation dirigée *in vitro*, en promouvant une différenciation en deux dimensions, permet d'obtenir un tissu spécifique. Mais ce type de différenciation est limité dans la maturation et dans les structures qu'elle peut reproduire. D'où l'intérêt du tératome qui offre une différenciation en trois dimensions et vascularisée (Eichelbaum and Evert, 1996).

Nous avons démontré que nos tératomes se différencient dans des tissus issus des trois feuillets embryonnaires (données non montrées). Nous nous sommes plus particulièrement intéressés à la différenciation dans la voie endodermique pulmonaire. Les tubes épithéliaux tubulaires retrouvés montrent soit une expression exclusive de marqueurs pulmonaires (TTF1/NKX2.1 ou CK7) ou de marqueurs digestifs (CDX2 ou CK20) soit une co-expression des deux dans le même tube, mais jamais dans les mêmes cellules. Cette cohabitation de cellules NKX2.1 et de cellules CDX2 n'existe pas à l'état physiologique puisque l'expression de NKX2.1 apparaît dans l'ébauche pulmonaire du tube digestif antérieur tandis que CDX2 est exprimé par le tube digestif moyen et postérieur. Ce profil atypique d'expression des marqueurs du tube digestif primitif montre la désorganisation du programme développemental dans les tératomes.

Le tératome est donc un moyen d'étudier *in vivo* le développement, mais cette technique a ses limites. Il n'est pas possible de suivre tous les marqueurs pulmonaires au cours de la formation d'un tératome. Il est difficile de guider la différenciation dans une voie en particulier, par exemple la voie pulmonaire. Une solution pour ce problème serait d'implémenter un système de polymère diffusant progressivement des cytokines d'intérêt tout au long de la formation du tératome.

Une alternative au tératome est l'organoïde. L'avantage de cette technique, outre d'économiser la vie des animaux de laboratoire, est sa disponibilité à l'intervention de l'expérimentateur, la facilité pour choisir l'arrêt de la manipulation, et bien sûr le nombre très supérieur d'expériences qui peut être entrepris par cette approche. Cependant, il convient de souligner que bien que cette différenciation se fasse en trois dimensions, la

non vascularisation de l'organoïde entraîne un arrêt assez rapide dans sa croissance. De plus, les cellules endothéliales sont un des constituants essentiels de la plupart de nos organes : il est vraisemblable, bien que ce point ait été peu abordé, qu'il soit important dans le développement des organes vascularisés.

Une autre alternative aux tératomes est la création d'un organe humain complet dans l'animal (Rashid et al., 2014). L'injection de CSP humaines dans un blastocyste d'un gros mammifère tel que le porc, et à condition de « vider » par manipulation génétique la niche d'accueil de l'organe, pourrait amener au développement de l'animal avec un organe humain. Le principe de cette approche a pu être démontré par la production d'un pancréas de rat dans une souris (Kobayashi et al., 2010). Cette stratégie soulève de sérieux problèmes éthiques, mais ces obstacles sont potentiellement solubles tant l'avancée thérapeutique serait immense (Bourret et al., 2016).

## **2. La différenciation dirigée en épithélium bronchique**

### **2.1 Production d'une lignée d'iPS d'un patient atteint de dyskinésie pulmonaire**

#### **2.1.1 La dyskinésie ciliaire primitive**

Pour démontrer l'intérêt des iPS pour modéliser les pathologies pulmonaires, il nous a paru intéressant de choisir (1) une pathologie génétique, et (2) une pathologie dont

l'anomalie serait simple à mettre en évidence. La dyskinesie ciliaire primitive (DCP) répond à ces deux conditions.

La DCP, ou syndrome d'immobilité ciliaire primitive, est une maladie autosomique récessive entraînant un battement asynchrone des cils (Sturgess et al., 1986). Cette affection touche tous les organes porteurs de cellules ciliées : bronches, sinus, oreille interne, et gamètes en général. Au niveau pulmonaire, cela se traduit par une évacuation inefficace du mucus par l'appareil locomoteur des bronches, entraînant des infections à répétition et à terme une dilation des bronches ou bronchectasies. La colonisation microbiologique (bactéries, mycobactéries, champignons) et l'encombrement imposent des endoscopies itératives. L'impact de cette pathologie est lourd pour les patients : antibiothérapies répétées avec le risque d'émergence de germes plus résistants, et effets secondaires parfois inacceptables des antibiotiques, symptômes quotidiens respiratoires (toux, expectoration, rhinorrhée, anosmie, obstruction nasale...) et extraréspiratoires (otites, stérilité, surdit , situs inversus, etc.), et un d clin acc l r  de la fonction respiratoire conduisant   l'insuffisance respiratoire chronique (Popatia et al., 2014; Sturgess et al., 1986). En France, sa pr valence estim e est de 1/15 000   1/30 000 personnes. La dyskinesie ciliaire primitive est la deuxi me plus fr quente maladie g n tique apr s la mucoviscidose. Le diagnostique repose principalement sur des  tudes fonctionnelles : la clairance mucociliaire, le NO nasale et le mouvement ciliaire. Elle est  galement identifiable au niveau microscopique sur une biopsie de l' pith lium bronchique, soit par vid omicroscopie en mettant en  vidence un d faut de mobilit , soit par microscopie  lectronique par mise en  vidence d'un d faut dans l'assemblage des microtubules des cils.

L'analyse des causes g n tiques de la DCP a montr  que de nombreux g nes pouvaient  tre en cause. En effet, le cil motile est compos  de plus de 250 prot ines et, pour la plus part de ces prot ines, leur absence par mutation du g ne qui code pour elle, entra ne un dysfonctionnement du cil (Popatia et al., 2014). Les mutations les plus fr quentes touchent notamment les g nes codant pour les dyn ines (DNAI1, DNAI2, DNAH5, etc.), les prot ines d'assemblage des bras internes de dyn ines (CCDC39, CCDC40), ou bien d'autres prot ines

du cil (HYDIN, HEATR2, etc.). L'inventaire des mutations pouvant causer un DCP n'est pas terminé (Fliegau et al., 2007; Knowles et al., 2013).

### 2.1.2 Dérivation de deux lignées de fibroblastes cutanés et d'une lignée iPS de patients atteints de DCP

Dans le cadre de notre collaboration avec l'équipe de pneumologie du CHU de Montpellier, nous avons sélectionné deux patients, une femme de 30 ans et un homme de 35 ans atteints de DCP. Nous avons dérivé deux lignées de fibroblastes à partir d'une biopsie de peau à la face interne du bras pour chaque patient, FMO2 et FMO3 respectivement (Figure 29). Une lignée de fibroblastes d'un volontaire sain a également été obtenue (FMO1).

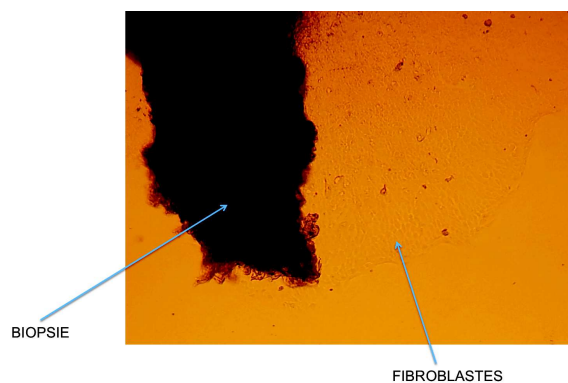
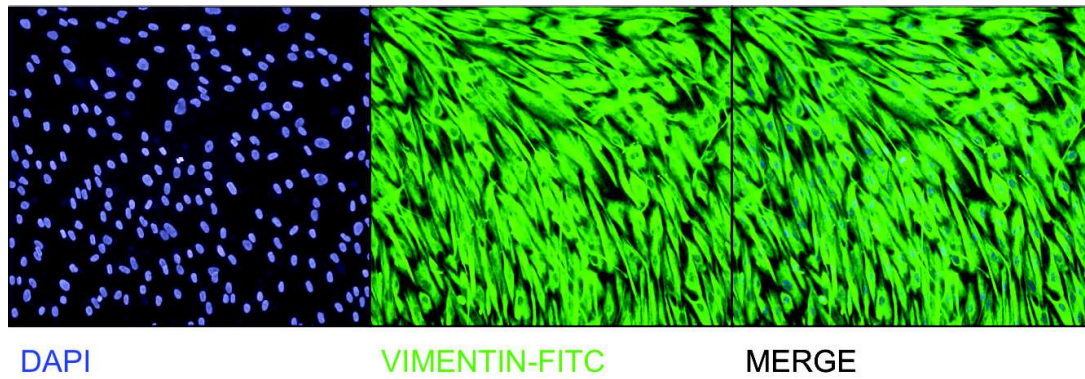


Figure 31 : Image de la dérivation de la lignée fibroblastique FMO3 9 jours après la biopsie.

Nous nous sommes assurés que les cellules dérivées expriment la vimentine, marqueur des fibroblastes (Figure 30). Ces lignées ont été congelées et sécurisées.



**Figure 32 : Immunofluorescence mettant en évidence l'expression par les fibroblastes de la lignée FMO2 de vimentine au cinquième passage.**

Nous avons ensuite réalisé une reprogrammation de la lignée de fibroblastes cutanés FMO2. Nous avons utilisé le kit de reprogrammation CytoTune™ (Life technologies™) (Lieu et al., 2013) qui permet l'expression des quatre facteurs de transcription du cocktail de Yamanaka via un virus de sendai (Figure 31). Ce type de vecteur est non intégratif et particulièrement efficace. En suivant les recommandations du fabricant, nous avons obtenu trente clones iPS à partir d'une seule reprogrammation. L'émergence du premier clone s'est faite en quinze jours (Figure 32), d'autres sont apparus jusqu'à plus de un mois après la reprogrammation. Brièvement, les fibroblastes humains sont cultivés en milieu DMEM. Dès qu'ils atteignent une confluence de 80-100%, ils sont dilués au 1/5. Deux jours après le repiquage, leur division cellulaire est théoriquement maximale. C'est à ce moment que les fibroblastes sont transfectés par les vecteurs Sendai. Il convient de noter que les fibroblastes transfectés non reprogrammés jouent le rôle de feeder (je n'ai pas utilisé de feeder murin). Après trois semaines de culture, les colonies d'iPS émergent et sont repiquées mécaniquement la première fois, puis en mécanique ou en passage non enzymatique en utilisant du Versène (chélateur du  $\text{Ca}^{++}$ ) les passages suivants. Les iPS sont cultivés en milieu E8 et sur matrice de Geltrex (Figure 32). Les clones ont été congelés précocement et nous nous sommes focalisés sur la lignée iPCD02-30. Celle-ci exprime les protéines de la pluripotence OCT4, SOX2 (Figure 36) et NANOG.



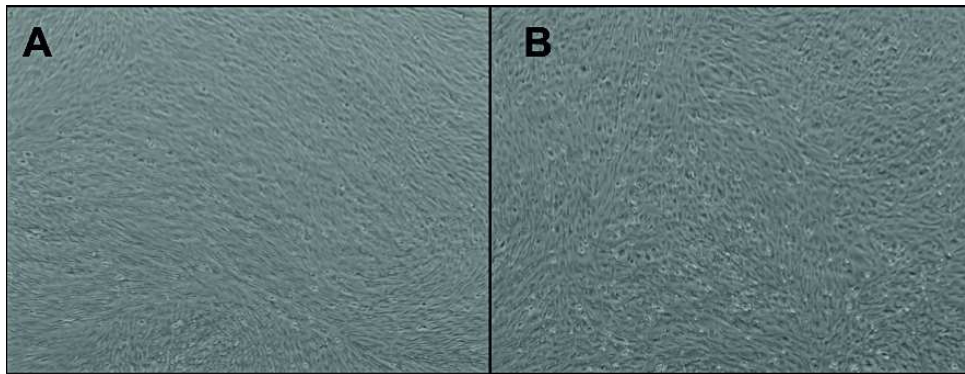


Figure 33 : (A) culture de fibroblastes FMO2 (B) culture de fibroblastes FMO2 48h après la reprogrammation.

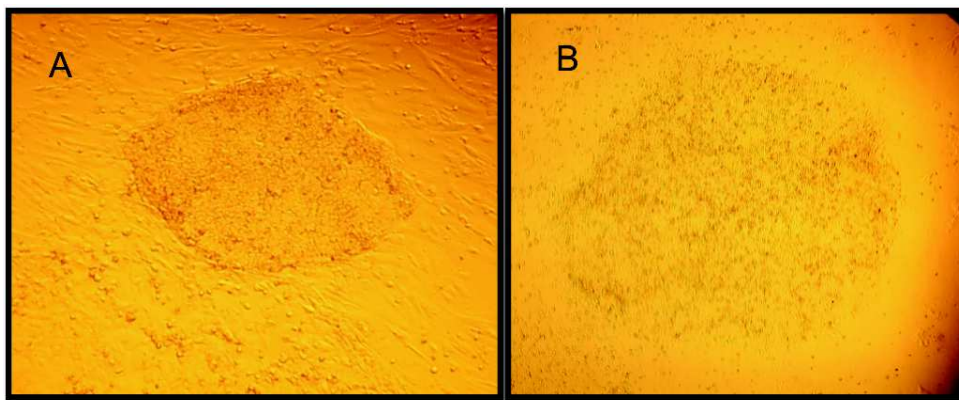


Figure 34 : (A) Premier clone apparu après reprogrammation : iPCD02-01 (P0), (B) iPCD02-30 adapté sur Geltrex dès son premier passage.

Des trente clones apparus, seul 60% ont pu donner une lignée. Les autres semblent avoir été mal reprogrammés et présentaient un aspect en grappe de raisin (Figure 33).

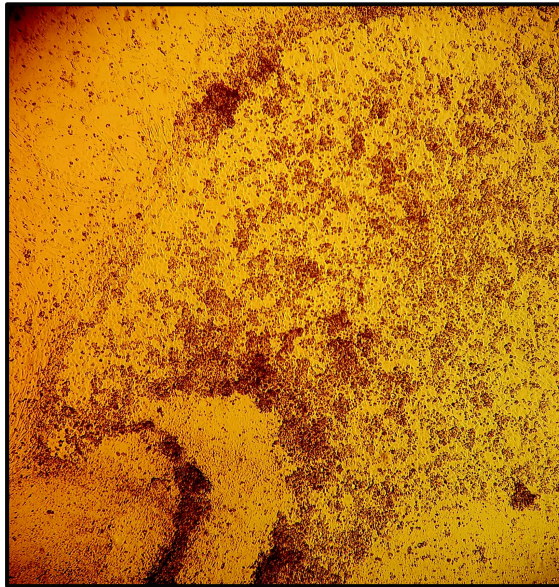


Figure 35 : Image d'une colonie iPS (iPCD02-26 (P1)) mal reprogrammée en aspect grappe de raisin.

Nous avons réalisé le caryotype de la lignée iPCD02-30 afin de s'assurer de l'absence d'anomalie chromosomique (Figure 34)

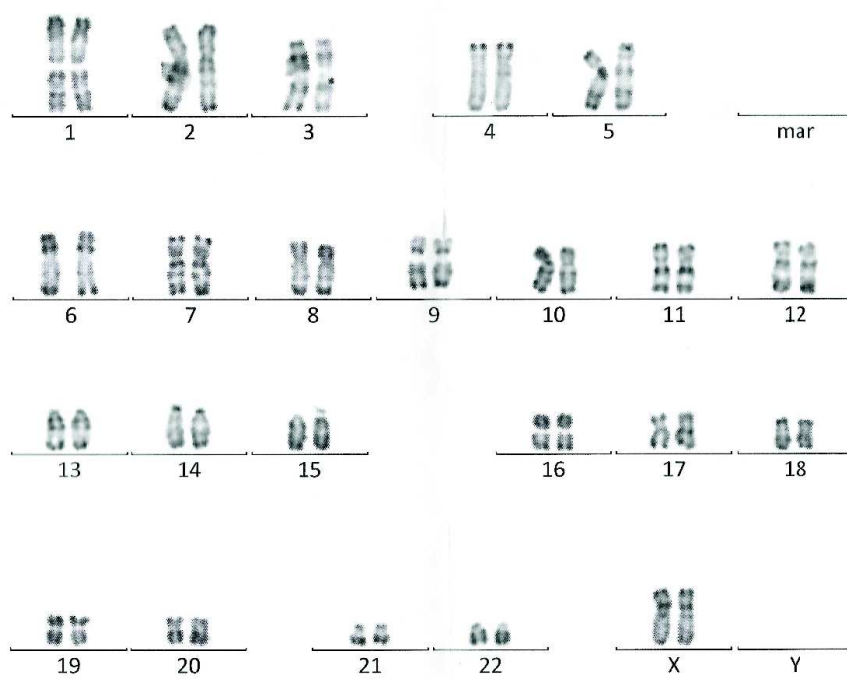


Figure 36 : Caryotype de la lignée iPCD02-30

La lignée iPS iPCD02-30 est capable de former des tératomes dans des souris immunodéprimées, et exprime les marqueurs de la pluripotence (Figure 35 et 36).

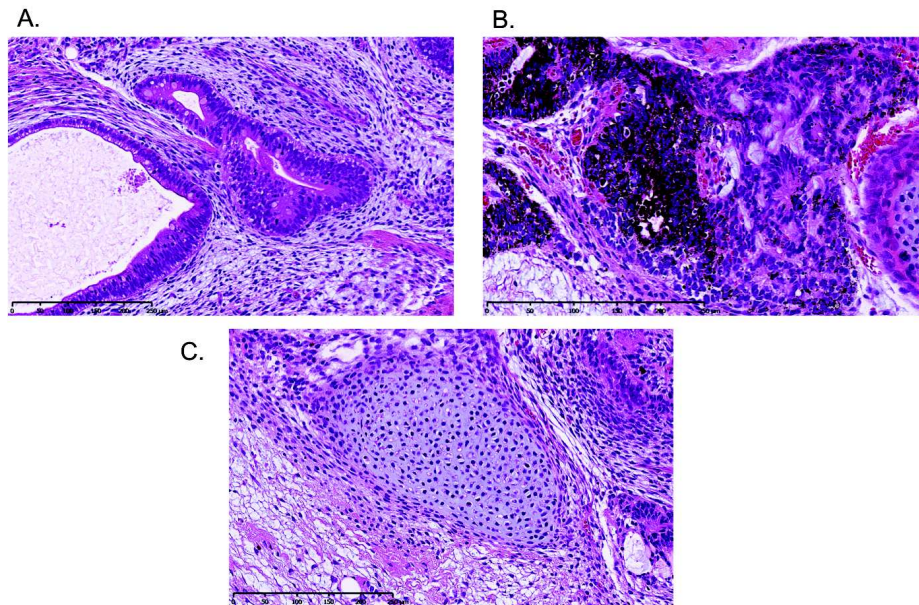


Figure 37 : Immunohistochimie de tératome issu d'une lignée iPS (iPCD02-30) montrant du tissu d'origine (A) endodermique, (B) ectodermique et (C) mésodermique.

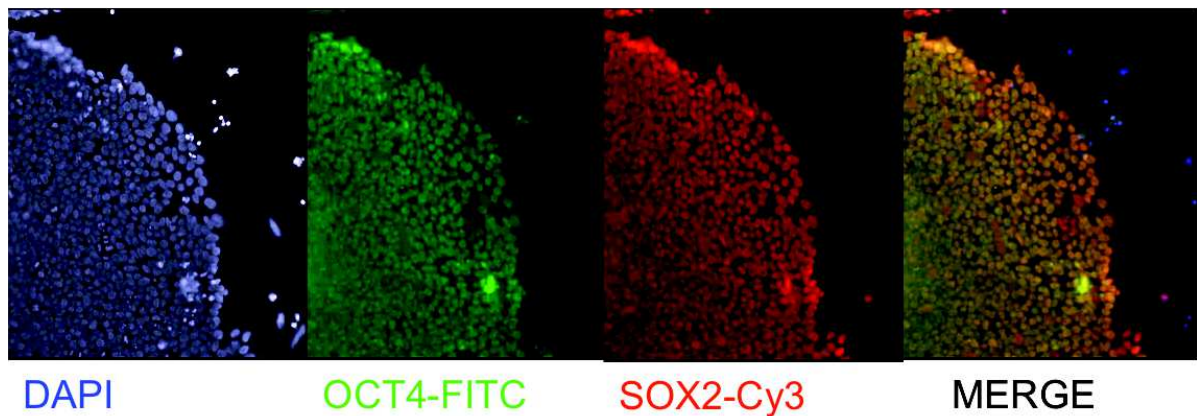


Figure 38 : Immunofluorescence montrant l'expression des marqueurs de la pluripotence OCT4 et SOX2 par iPCD02-30.

## **2.2 Protocoles de différenciation**

### **2.2.1 Identification des marqueurs à suivre**

Il est nécessaire de trouver les bons marqueurs afin de suivre l'évolution de la différenciation au cours du temps. Les marqueurs de la pluripotence disparaissent pour laisser place aux marqueurs de l'endoderme définitif qui laisseront place aux marqueurs de l'AFE, des progéniteurs bronchiques et enfin de l'épithélium bronchique.

Dans un chapitre précédent, nous avons vu les marqueurs de la pluripotence. La disparition des marqueurs nucléaires (NANOG, OCT4, SOX2) et des marqueurs membranaires (SSEA3, SSEA4, TRA-1-60, TRA-1-81) indique un engagement dans la différenciation de la part des CSP. L'apparition de l'endoderme définitif s'accompagne de l'expression des marqueurs nucléaires SOX17, FOXA2 et du marqueur membranaire CXCR4 (D'Amour et al., 2005). La différenciation vers l'épithélium bronchique se poursuit ensuite avec l'apparition des marqueurs de l'AFE représenté avec dans la Figure 37.

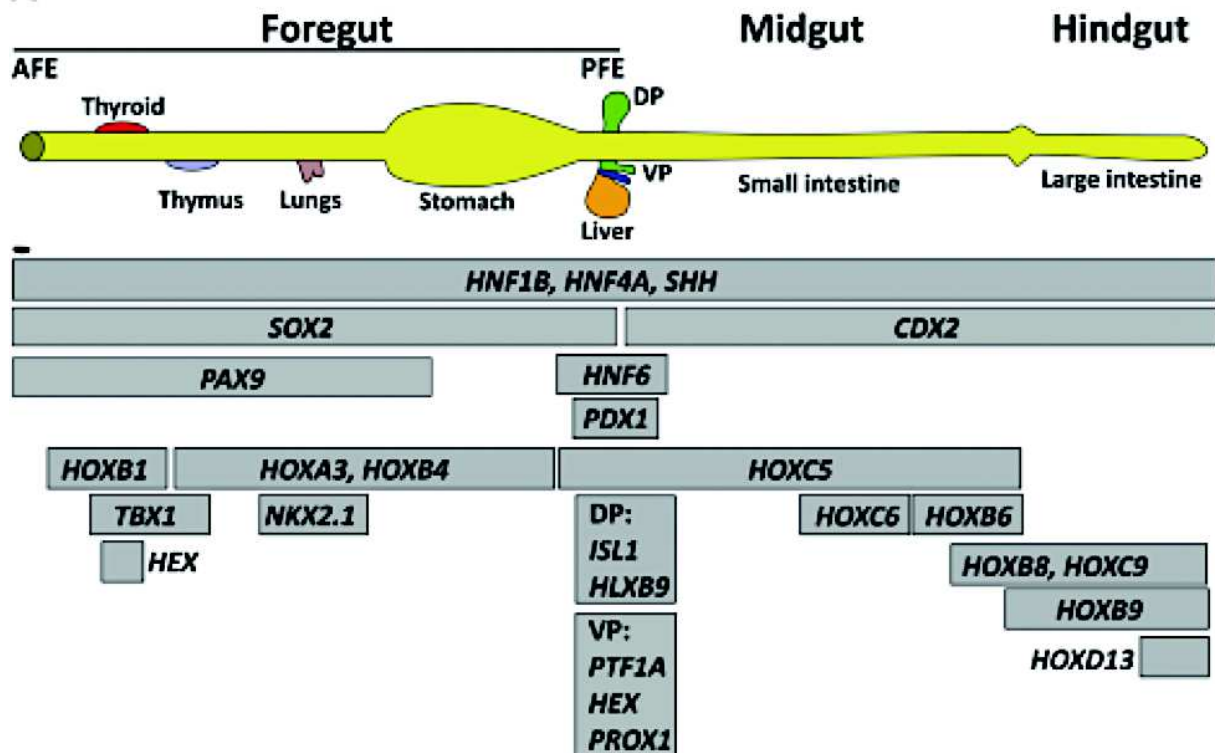


Figure 39 : Schéma de l'expression des facteurs de transcription le long du tube digestif d'un embryon de souris à E9.5. Abréviations: AFE, anterior foregut endoderm; DP, dorsal pancreatic bud; PFE, posterior foregut endoderm; VP, ventral pancreatic bud (d'après (Davenport et al., 2016b)).

Les marqueurs de l'appareil respiratoire diffèrent selon les parties de l'arbre bronchique et les types cellulaires. Au niveau des voies conductrices de l'air, on retrouve les cellules basales qui expriment p63, Forkhead box protein J1 (FOXJ1) et Tubuline IV (TubIV) pour les cellules ciliées, Mucin 5AC (MUC5AC), SCGB1A1 (ou CC10) et Cystic Fibrosis Transmembrane conductance Regulator (CFTR) pour les cellules caliciformes. SOX2 est un marqueur proximal des voies respiratoires alors que SOX9 est un marqueur distal. Au niveau alvéolaire, les pneumocytes II expriment le surfactant protein C (SFTPC).

L'épithélium bronchique exprime également des marqueurs épithéliaux comme Zona Occludens 1 (ZO-1), E-cadhérine et Epithelial calmoduline (EpCAM).

Les marqueurs que j'ai utilisés pendant mon travail de thèse sont récapitulés dans les tables 9 (primers qRT-PCR) et 10 (anticorps).

**Gene Forward primer (5'→3')**

**Reverse primer (5'→3')**

KRT5	GGAGTTGGACCACTCAACATC TGGAGTAGTAGCTT CCACTGC	SOX2	GCACATGAAGGAGCACCCGGATTA CGGGCAGCGTGTACTTATCCTTCTT
MUC5AC	CCATTGCTATTATGCCCTGTGT TGGTGGACGGACAG TCACT	FOXG1	CTCCGTC AACCTGCTCGCGGF CTGGCGCTCATGGA CGTGCT
NKX2-1	ACCAGGACACCATGAGGAAC CGCCGACAGGTA CTCTGT	PAX9	TGGTTATGTTGCTGGACATGGGTG GGAAGCCGTGACAG AATGACTACCT
FOXJ1	GAGCGGCGCTTCAAGAAG GGCCTCGGTATTCA CCGTC	SOX9	GAGGAAGTCGGTGAAGAACG CCAACATCGAGACC TTCGAT
SFTPC	CACCTGAAACGCCTTCTTATCG TGGCTCATGTGGAG ACCCAT	PAX6	TCTTTGCTTGGGAAATCCG CTGCCCGTTCAACA TCCTTAG
CCSP	TTCAGCGTGCATCGAAACCCC ACAGTGAGCTTTGG GCTATTTTT	OTX2	GTGGGCTACCCGGCCACCC GCACCCTCGACTCG GGCAAG
FOXA2	AGGAGGAAAACGGGAAAGAA CAACAACAGCAATG GAGGAG	KRT15	GGCTGGAGAACTCACTGGC CAGGCTGCGGTAAG TAGCG
BACT	CTGGAACGGTGAAGGTGACA AAGGGACTTCTGTGTA ACAATGCA	KRT16	GACCGGCGGAGATGTGAAC CTGCTCGTACTGGT CACGC
SOX17	AAGGGCGAGTCCCGTATC TTGTAGTTGGGGTG GTCCTG	MUC16	CCAGTCCTACATCTTCGGTTGT AGGGTAGTTCCTAG AGGGAGTT
TG	AGAAGAGCCTGTCGCTGAAA TTGGACCAGAAGGA GCAGTC	PDPN	GTCCACGCGCAAGAACAAAAG GGTCACTGTTGACA AACCATCT
NKX6-1	ATTCGTTGGGGATGACAGAG CGAGTCCTGCTTCT TCTTGG	P2X7	TATGAGACGAACAAAGTCACTCG GCAAAGCAAACGTA GAAAAAGAT
PDX1	CCCATGGATGAAGTCTACC GTCCTCCTCCTTTT TCCAC	FOXA1	CTCGCCTTACGGCTCTACG TACACACCTTGGA GTACGCC
AFP	TGGGACCCGAACTTTCCA GGCCACATCCAGGA CTAGTTTC	FOXP2	AATCTGCGACAGAGACAATAAGC TCCACTTGTTTGCT GCTGTAAA
CFTR	CTATGACCCGGATAACAAGGAGG CAAAAATGGCTGGG TGTAGGA	FOXE1	CACGGTGGACTTCTACGGG GGACACGAACCGAT CTATCCC
NANOG	TGATTTGTGGGCTGAAGAAA GAGGCATCTCAGCA GAAGACA	PITX1	CTAGAGGCCACGTTCCAGAG TGGTTACGCTCGCG CTTAC
TP63	ACTTCACGGTGTGCCACCCT GAGCTGGGGTTTCT ACGAAACGCT	DLX3	CTCGCCCAAGTCGGAATATAC CTGGTAGCTGGAGT AGATCGT
CDX2	CTGGAGCTGGAGAAGGAGTTTC ATTTTAACCTGCCT CTCAGAGAGC	MUC2	AGGATGACACCATCTACCTCAC CATCGCTCTTCTCA ATGAGCA
		TBX1	CGGCTCCTACGACTATTGCCC GGAACGTATTCTT GCTTGCCCT

**Tableau 9: Liste des amorces utilisées que nous avons utilisées pour les PCR et qui sont validées par A. Wong (Wong et al., 2015).**

Nom	Description	Clonal	Host	Cible	Fournisseur	No. Cat./Ref.	Dilution
CK20	mouse monoclonal	Monoclonal [SPM140]	Mouse	Semi-purified cytokeratin preparation (Human)	Abcam	ab1520	1/50
FOXA2 HNF-3B	= Human HNF-3 beta/FoxA2 Affinity Purified Polyclonal Ab, Goat IgG	Polyclonal	Goat IgG	Human HNF-3 beta/FoxA2	R&D	AF2400	1/100
OCT4	Oct-3/4 Antibody (H-134)	Polyclonal	Rabbit IgG	amino acids 1-134 of Oct-3/4 of human origin	Santa Cruz	sc-9081	1/300
SOX2	Human SOX2 Affinity Purified Polyclonal Ab, Goat IgG	Polyclonal	Goat IgG	human SOX2	R&D	AF2018	1/20
TTF1 NKX2.1	= mouse monoclonal	Monoclonal [8G7G3/1]	Mouse	TTF-1 humain, de rat et de souris	Dako	8G7G31	1/50
CC10	rabbit polyclonal	Polyclonal	Rabbit	human CC10	BioVendor	RD 181022220	1/10 000
Cy <sup>TM</sup> 3-conjugated* AffiniPure Donkey Anti-Goat	Secondaire		Donkey	Goat	Jackson Immunoresearch	705-165-147	1/1 000
Alexa Fluor® 488 Donkey Anti-Mouse IgG (H+L)	Secondaire		Donkey	Mouse	Molecular Probes / Invitrogen	A-21202	1/2 000

**Tableau 10 : Liste des anticorps utilisés pour les différentes immunofluorescences.**

## 2.2.2 Milieux

Les milieux de différenciation varient selon les protocoles utilisés et changent au cours de la différenciation. Tous ont en commun un milieu basique enrichi par d'autres facteurs indispensables à la survie des cellules.

A.	Composition	Volume
	KnockOut DMEM	450 ml
	KnockOut serum replacement	50 ml
	Penicillin-streptomycin	3 ml
	Glutamax	6 ml
	MTG	15 $\mu$ l
	Non-essential amino acid	3 ml

B.	Composition	Volume (ml)
	BEBM basal medium	500
	BPE	2
	Insulin	0.5
	Hydrocortisone	0.5
	GA-1000	0.5
	Transferrin	0.5
	Tri-iodothyronine	0.5
	Epinephrine	0.5
	Retinoic acid	0.5
	Human epidermal growth factor (hEGF)	0.5

C.	Composition	B-ALI differentiation medium (ml)	B-ALI growth medium (ml)
	Basal medium	500	250
	Bovine pituitary extract (BPE)	2	1
	Insulin	0.5	0.25
	Hydrocortisone	0.5	0.25
	GA-1000	0.5	0.25
	Retinoic acid	0.5	0.25
	Transferrin	0.5	0.25
	Tri-iodothyronine	0.5	0.25
	Epinephrine	0.5	0.25
	hEGF	0.5	0.25

Figure 40 : Composition des milieux de différenciation du protocole de (Wong et al., 2015) (A) pour la différenciation des CSP de l'endoderme définitif jusqu'aux progéniteurs bronchiques (B) pour l'expansion des progéniteurs bronchiques (C) pour la polarisation et la maturation de l'épithélium bronchique

Composition	Volume
<b><u>Milieu pour la production de DE</u></b>	
RPMI-1640	
B27 minus vitamine A	2%
<b><u>Milieu de différenciation des étapes suivantes</u></b>	
RPMI-1640	
B27	2%

Figure 41 : Composition des milieux de différenciation du protocole de (Mou et al., 2012).



Composition	Volume	Composition	Volume
<b>Milieu pour la différenciation</b>		<b>Milieu de maturation des cellules différenciées</b>	
IMDM	75%	Ham's F12 media	
Ham's Modified F12 medium	25%	HEPES pH 7.4	15mM
Supplémenté en N2, B27+RA		CaCl <sub>2</sub>	0.8mM
BSA	0.05%	BSA	0.25%
L-glutamine	200mM	Insulin	5µg/ml
Acide ascorbique	0.05mg/ml	Transferrin	5µg/ml
MTG	4.5x10 <sup>-4</sup>	Na selenite	5ng/ml
		Dexamethasone	50nM
		8-Br-cAMP	0.1mM
		IBMX	0.1mM

Figure 42 : Composition des milieux de différenciation du protocole de (Longmire et al., 2012).

Composition	Volume
<b>Milieu pour la différenciation</b>	
DMEM/F12	
Supplémenté en N2, B27	
Acide ascorbique	50µg/ml
Glutamax	2mM
MTG	0.4µM

Figure 43 : Composition des milieux de différenciation du protocole de (Green et al., 2011).

		Day 1	Day 2-3	Day 4	Day 5-9	Day 9-17	Day 17+
RPMI	Invitrogen	YES	YES				
FBS	Hyclone		1%	2%	2%	2%	
Ultra MEM - ITES				YES	YES	YES	
Pen/Strep	Invitrogen						100x
Wnt3a	R&D Systems	25ng/ml					
Activin A	Humanzyme	100ng/ml	100ng/ml				
SB431542	Sigma			500nM	500nM		
Noggin	Peptotech			20ng/ml			
BMP4	Humanzyme				5ng/ml	5ng/ml	
FGF2	Peptotech					10ng/ml	
FGF10	Peptotech					10ng/ml	
KGF	Peptotech					10ng/ml	
DMEM/F12	Invitrogen						YES
Ultrosor G	Pall						2%
Fetal Clone II	HyClone						2%
Insulin	Sigma						2.5µg/ml
Bovine Brain Extract	Lonza						22.5µg/ml
Transferrin	Sigma						2.5µg/ml
Hydrocortizone	Sigma						20nM
3,3',5-Triiodo-L-thyronine sodium salt	Sigma						500nM
Epinephrine	Sigma						1.5µM
Retinoic Acid	Sigma						10nM
Phosphorylethanolamine	Sigma						250nM
Ethanolamine	Sigma						250nM

Figure 44 : Composition des milieux de différenciation du protocole de (Firth et al., 2014).

### **2.2.3 Matrices**

Avant toute différenciation, il faut réaliser deux à trois passages sur MEC afin d'éliminer les feeders. Nous avons supprimé cette étape en cultivant nos CSP directement sur MEC.

Tout comme les milieux de différenciation, les matrices de culture sur lesquelles sont différenciées les cellules varient en fonction des protocoles.

Un protocole se démarque en induisant DE par la formation d'EB (Green et al., 2011). Les EB sont ensuite dissociées puis plâtées en couche unicellulaire sur une matrice de fibronectine à 0,33% puis à 0,2% lorsque la différenciation est plus avancée.

Nous avons utilisé une MEC (indifféremment Geltrex ou Matrigel) afin de différencier les CSP. Les cellules arrivées au stade DE sont passées sur Transwell précoâtés par avec du collagène IV à 0,06mg/ml (Wong et al., 2015). Nous avons apporté des modifications au coating des Transwells en augmentant le collagène IV à 1mg/ml et en le mélangeant à de la fibronectine 0,33% ou 0,22%. Il apparaît que le mélange de collagène 1mg/ml avec la fibronectine 0,33% est la meilleure combinaison pour l'attachement des cellules et leur différenciation.

### **2.2.4 Facteurs de croissance**

Chaque protocole utilise un cocktail de cytokines bien déterminé afin de différencier les cellules.

#### **2.2.4.1 Obtenir de l'endoderme définitif**

Un des premiers protocoles publiés dans la différenciation dirigée en épithélium bronchique (Mou et al., 2012), engage la différenciation des cellules par un cocktail composé d'Activine A (100ng/ml) et de LY294002 (5 $\mu$ M), un inhibiteur de la PI3kinase. Nous avons testé ce protocole et dû adapter la concentration de LY294002 en la diminuant. Malgré cela, la mortalité cellulaire très importante nous l'a fait abandonner.

L'obtention de l'endoderme définitif est une étape de différenciation commune avec la différenciation hépatique, à l'instar de l'équipe de D'Amour (D'Amour et al., 2005) qui utilise de l'Activine A (100ng/ml) avec une augmentation progressive de 0 à 2% de sérum, celle de l'équipe de Sabine Gerbal-Chaloin (U1183, non publié) qui utilise de l'Activine A (100ng/ml) avec une augmentation progressive de 0,1 à 0,2% de sérum et d'autres cytokines telles que du FGF2 (10ng/ml) et du Wnt3a (25ng/ml). Nous avons donc utilisé ce dernier protocole pour nos différenciations (Figures 43, 44).

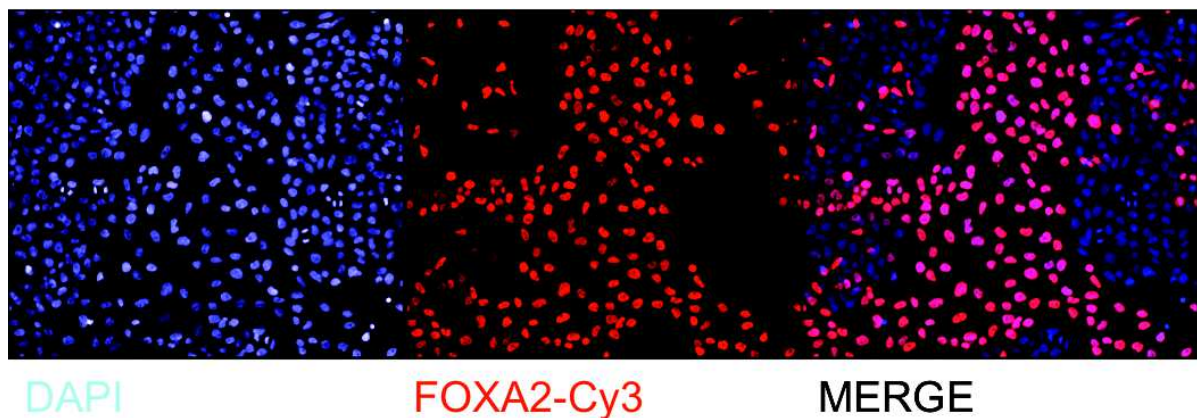


Figure 45 : Immunofluorescence mettant en évidence l'expression de FOXA2 dans des cellules (HS306) différenciées en DE par le protocole de S. Gerbal-Chaloin.

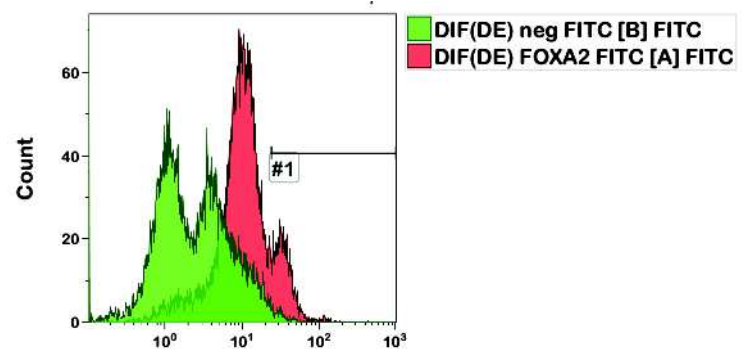


Figure 46 : Mise en évidence de l'expression de FOXA2, par FACS, dans des cellules (HS306) différenciées en DE par le protocole de S. Gerbal-Chaloin.

La commercialisation d'un kit de différenciation par Stemcell™, STEMdiff™ Definitive Endoderm Kit, également utilisé par d'autres équipes (Wong et al., 2015), permet d'obtenir une différenciation plus robuste, plus reproductible et moins couteuse *in fine* (Figures 45, 46, 47).

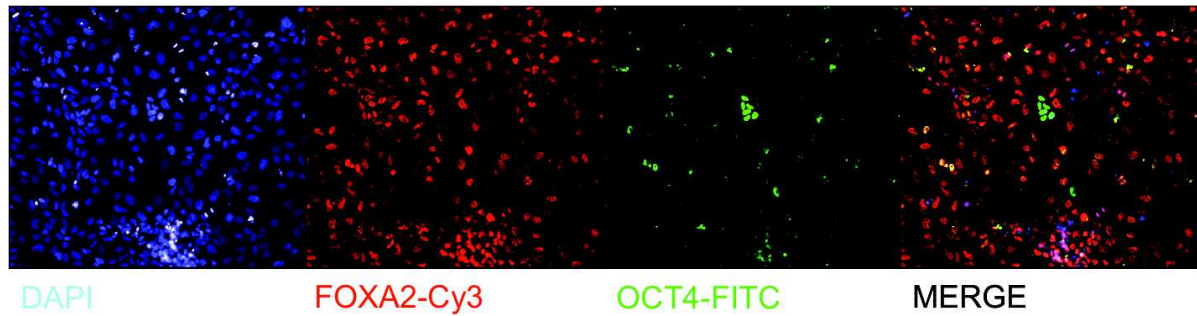


Figure 47 : Immunofluorescence mettant en évidence l'expression de FOXA2 dans les cellules (BJ-CTL-C1) différenciées en DE avec STEMdiff™ Definitive Endoderm Kit.

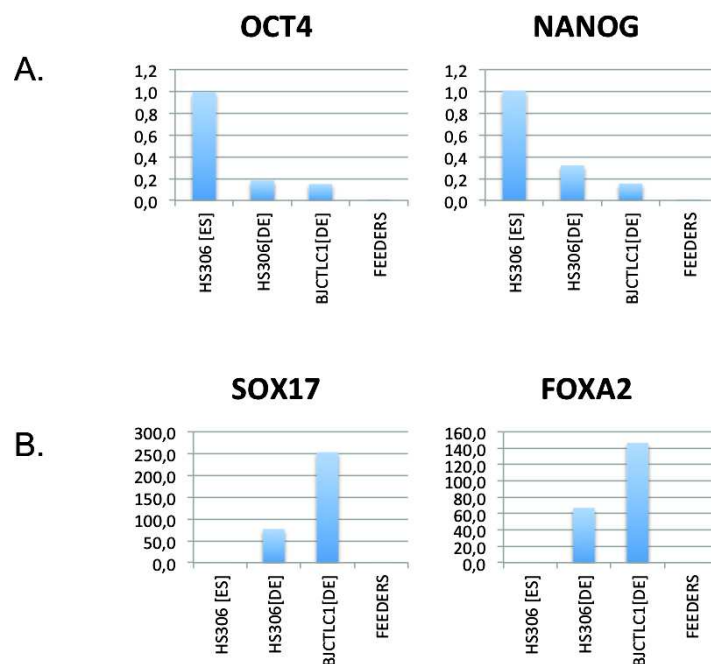


Figure 48 : Expression relative des marqueurs de (A) pluripotence, (B) DE après différenciation de CSP (lignée HS306 et BJCTL1) en DE avec STEMdiff™ Definitive Endoderm Kit.



Figure 49 : Microscopie optique de cellules (iPCD02-30) en cours de différenciation, après 3 jours de traitement par le STEMdiff Definitive Endoderm kit.

Nous avons également testé un autre protocole (Huang et al., 2015) mettant en place la différenciation en DE par le passage en EB (Figure 48). Cependant, la reproductibilité de cette technique sur la taille des EB et leur manutention ont posé problème. En effet, dans les mêmes conditions d'expérimentation, la taille des EB n'a pas pu être reproductible. En particulier, lorsque leur taille était trop faible, ils se désagrégeaient à chaque changement de milieu.

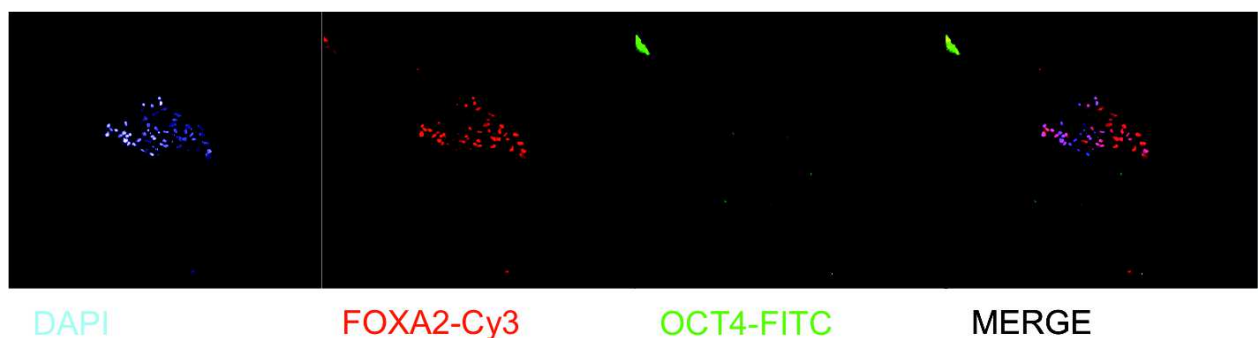


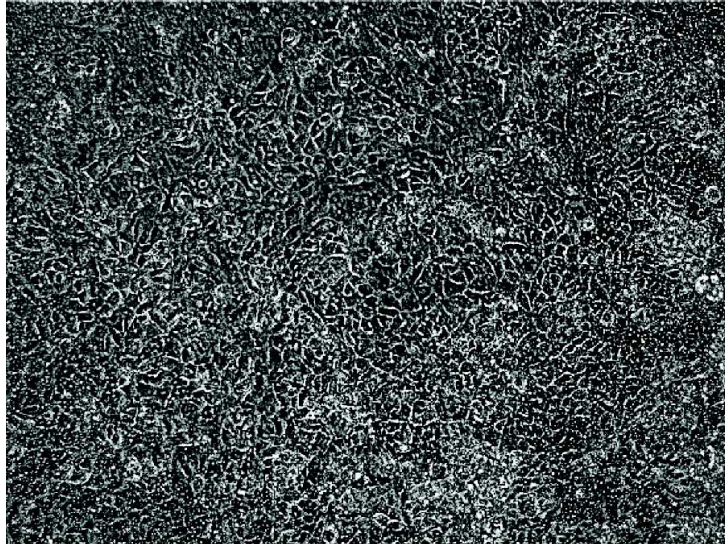
Figure 50 : Immunofluorescence mettant en évidence l'expression de FOXA2 dans les cellules (iPCD02-30) différenciées en DE par le protocole 25654758.

#### 2.2.4.2 Obtenir de l'AFE

Après avoir obtenu de l'endoderme définitif, l'étape suivante est l'obtention d'AFE.

L'équipe de Maharaj Rajagopal de Boston (Mou et al., 2012) utilise un inhibiteur sélectif des récepteurs du TGF- $\beta$  de type I ALK5 kinase, type I activin/nodal ALK4 et type I nodal ALK7, le A-83-01, à une concentration de 500nM. Cette molécule, même utilisée à des concentrations plus faibles allant de la dilution au demi à celle au un seizième, a entraîné une mortalité cellulaire importante. L'ajout de NOGGIN à 100ng/ml, un antagoniste de BMP4, améliore la survie des cellules en cours de différenciation mais leur nombre reste trop faible pour pouvoir poursuivre correctement la production d'épithélium bronchique.

L'équipe de Janet Rossant de Toronto (Wong et al., 2012; 2015) a opté pour la combinaison de deux cytokines, SHH à 50ng/ml et FGF2 à 500ng/ml, associé à un passage des cellules sur Transwell afin de différencier les cellules en AFE. Le passage sur Transwell provoque une mortalité cellulaire importante. Nous avons essayé plusieurs concentrations des cytokines sans succès. Lors du passage sur Transwell, les cellules se trouvent dans un état fragile car elles ont été isolées les unes des autres (passage single-cell). La molécule chimique Y-27632 est un inhibiteur de la kinase ROCK et inhibe l'apoptose cellulaire induite par ROCK lors du passage des CSP en single-cell (Kurosawa, 2012). L'ajout de cette molécule à 10 $\mu$ M lors du passage des cellules DE sur Transwell augmente considérablement la survie cellulaire et permet de d'aborder cette étape avec les concentrations préconisées (Figure 49). Nous nous confrontons alors à une autre difficulté : la densité cellulaire. Cela fera l'objet d'un chapitre plus loin.



**Figure 51 : Microscopie optique de cellules (iPCD02-30) différenciées en AFE (après 8 jours de différenciation selon le protocole (Wong et al., 2015)).**

#### **2.2.4.3 Obtenir des PL**

Amy Wong, de l'équipe de Janet Rossant, (Wong et al., 2012; 2015) obtient des progéniteurs bronchiques par la succession de deux étapes de cinq jours. La première associe les cytokines suivantes : FGF10, FGF7, aussi connue comme KGF, et BMP4, respectivement aux concentrations de 50ng/ml, 50ng/ml et 5ng/ml. La seconde combine FGF10, FGF7 et FGF18, toutes trois à 10ng/ml. Nous avons appliqué ce protocole sans difficulté majeure, sinon la densité cellulaire qui a été un point critique de cette étape et qui sera également développée dans un chapitre ultérieur (Figure 50).

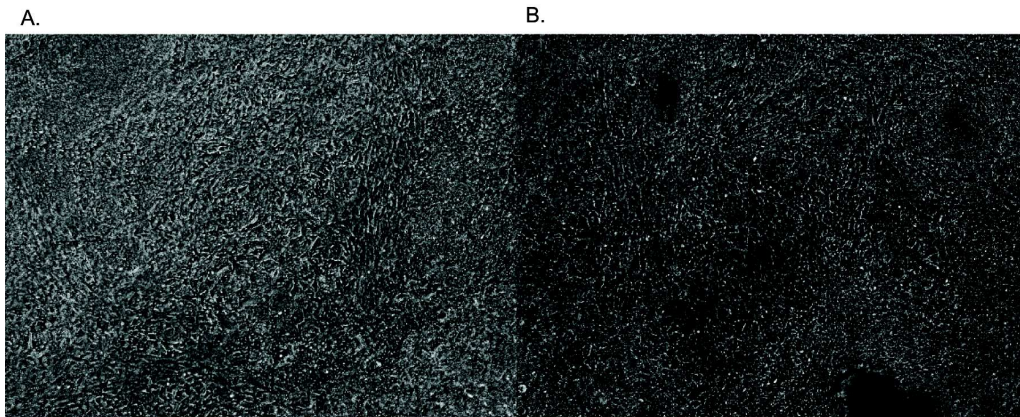


Figure 52 : Microscopie optique des cellules (iPCD02-30) au stade PL (après (A) 24 et (B) 29 jours de différenciation selon le protocole (Wong et al., 2015)).

#### 2.2.4.4 Polarisation, Maturation

Enfin, la dernière étape qui permet la croissance des progéniteurs suivie de leur polarisation et de leur maturation (maturité d'un épithélium foetal), est réalisée par l'équipe de Toronto (Wong et al., 2012; 2015) par la succession de deux milieux de culture, décrits dans le Chapitre 2.2.2 (Figure 51).

Nous avons été confronté à plusieurs difficultés au cours de cette étape. (i) la densité cellulaire, qui sera abordée dans le chapitre suivant ; (ii) le décollement de l'épithélium bronchique, ce point a été résolu par l'adaptation du coating des Transwells comme décrit précédemment dans le Chapitre 2.2.3 ; (iii) certaines cellules différenciées présentent des contractions rythmées spontanées. Néanmoins, nous avons pu constater l'expression de marqueurs pulmonaires avancés par qRT-PCR (Figure 52).



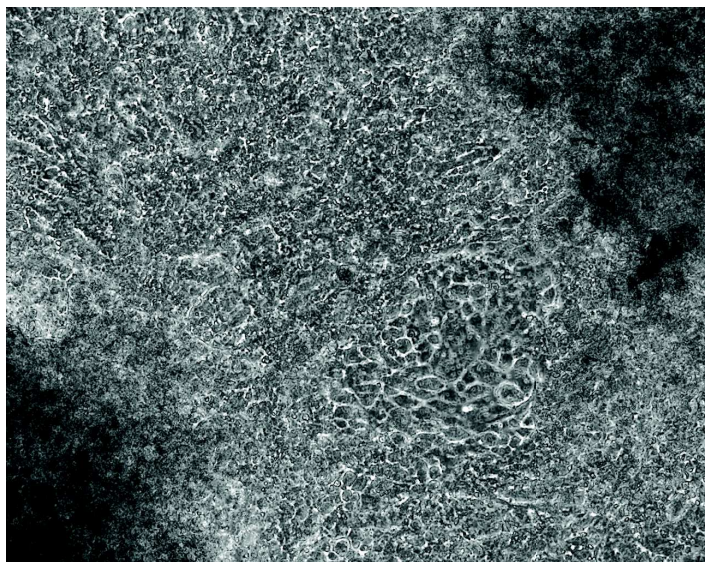


Figure 53 : Microscopie optique des cellules (iPCD02-30) au stade PL (après 35 jours de différenciation selon le protocole (Wong et al., 2015)).

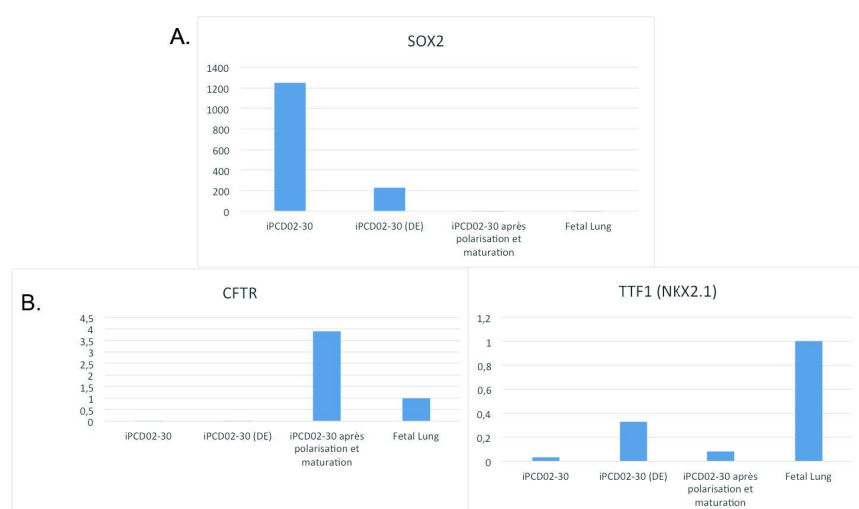


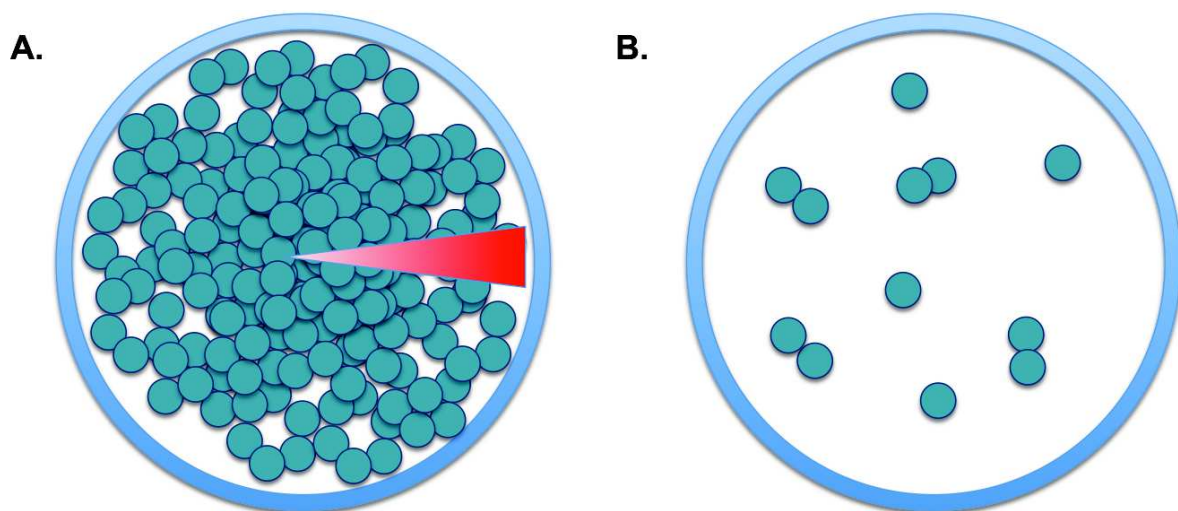
Figure 54 : Expression relative des marqueurs (A) de pluripotence, (B) d'un épithélium bronchique polarisé et mature après différenciation de CSP (lignée iPCD02-30) selon le protocole (Wong et al., 2015).

## 2.2.5 Densité cellulaire et passage

### 2.2.5.1 Densité cellulaire

Au cours de ce travail de recherche, il est apparu qu'adapter la densité cellulaire après chaque passage est un élément clé de la différenciation au même titre que la concentration des cytokines. En effet, ces deux paramètres vont de pair. S'il y a trop de

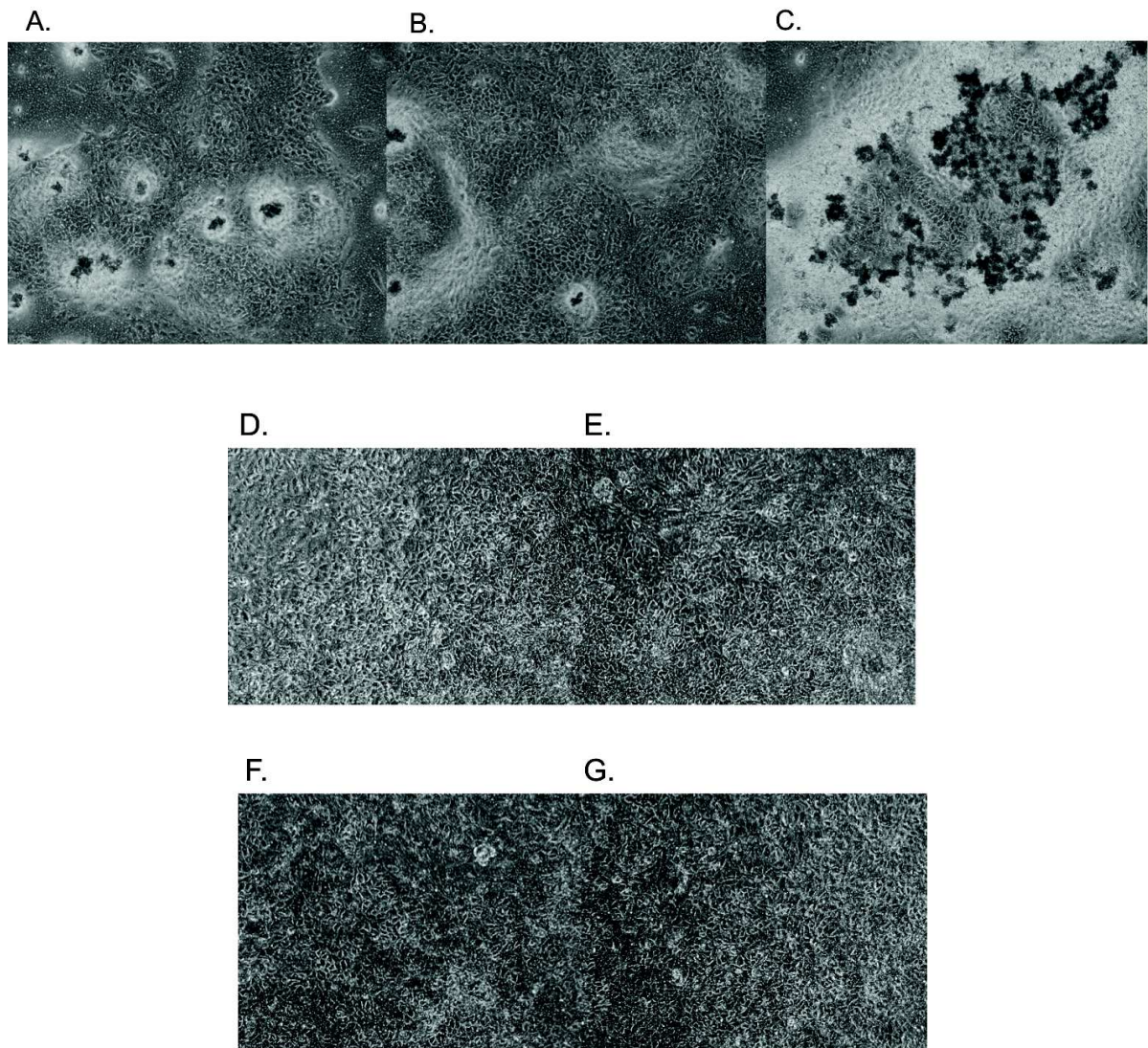
cellules dans le puits de culture (Figure 53-A), il se forme un complexe en trois dimensions et donc un gradient de concentration de diffusion des cytokines. Les cellules en périphérie ne recevront pas les cytokines à la même concentration que les cellules au centre de l'agrégat. Au contraire, si la densité est trop faible (Figure 53-B), les cellules sont trop dispersées. L'altération de la communication paracrine entraîne une surmortalité cellulaire.



**Figure 55 : (A) Représentation d'une densité cellulaire (A) trop importante entraînant un gradient de concentration de cytokines (B) trop faible avec absence de communication paracrine.**

De plus, chaque lignée cellulaire a un comportement différent vis à vis de la concentration cellulaire lors des passages. Par exemple, une lignée peut avoir besoin d'êtreensemencée à XXX cellules/cm<sup>2</sup>, alors qu'une autre aura besoin du double, voire plus, ou au contraire de la moitié, voire moins, de cellules/cm<sup>2</sup>. Donc chaque étape dans la différenciation nécessitant un passage cellulaire a nécessité une adaptation de la densité cellulaire. Les cellules en cours de différenciation ne cessent pas de proliférer. A chaque passage, il a donc fallu optimiser soigneusement la densité cellulaire, afin d'avoir un épithélium unicellulaire tout au long de la différenciation (Figure 54). Malgré tout, lorsque la prolifération de l'épithélium est trop importante au cours de la différenciation, il est possible d'avoir recours à un passage cellulaire de Transwell à Transwell avec une dilution de 1:1 ou 1:2. Cela est possible uniquement aux étapes permettant d'obtenir des

progéniteurs bronchiques (PL), ensuite l'épithélium est trop cohésif et il est impossible de réaliser un passage.



**Figure 56 : Microscopie cellulaire des cellules (iPCD02-30) après 9 jours de différenciation selon le protocole (Wong et al., 2015) et après plating au stade AFE à (A) 25 000 , (B) 35 000, (C) 50 000, (D) 100 000, (E) 150 000, (F) 175 000, (G) 200 000 cellules.**

### **2.2.5.2 Passage**

Il existe plusieurs méthodes de passage des cellules. Historiquement, les colonies de CSP sont passées mécaniquement, une partie (ou clump) est découpée à la pointe effilée d'une pipette pasteur, au scalpel ou au tips de pipette, et déposée sur un nouveau tapis de

feeder. Cette méthode a été abandonnée dans le laboratoire au profit d'autres moins fastidieuses. Le Versène (réactif à base d'EDTA, un chélateur de  $Ca^{++}$ ), avec une utilisation à température ambiante, permet le passage des colonies en clumps.

Outre les passages mécaniques, ou en clumps, il existe les passages grâce à l'utilisation d'enzymes. Selon l'enzyme employée et le temps d'incubation, cela permet d'isoler les cellules clumps ou en single-cell. La collagénase IV ne permet que de dissocier les colonies des cellules feeders : cette enzyme ne permet qu'un passage en clumps. Dans le cas d'une dissociation complète des cellules, en single-cell, il est possible de dénombrer les cellules avant de les passer, ce qui est impossible avec les clumps. Plusieurs enzymes sont utilisables dans ce cadre, notamment l'acutase, la trypsine et le TrypLE. Cependant, après plusieurs passages les cellules peuvent présenter des anomalies génomiques (Bai et al., 2013; 2015). L'utilisation du Versène à 37°C permet également un passage en single-cell.

Enfin, l'utilisation de la molécule chimique Y-27632, décrite plus haut, est une aide précieuse pour diminuer considérablement la mortalité après un passage en single-cell et donc augmenter la densité cellulaire après passage sur Transwell des cellules différenciées en DE.

## 2.3 Discussion

Au cours de mon travail de thèse, j'ai pu obtenir du DE, de l'AFE et des PL (Figures 46 et 52). Cependant, je n'ai pu obtenir des épithélia homogènes exprimant en protéine des marqueurs de cellules bronchiques différenciées, notamment des cils. Moi-même et d'autres dans le laboratoire, poursuivons des expériences de différenciation pour obtenir un épithélium qui serait similaire à ce que l'on peut obtenir à partir d'une biopsie de bronche adulte (Gras et al., 2012). La difficulté à obtenir un épithélium comportant des cellules bronchiques différenciées peut être expliqué par plusieurs facteurs.

Tous les êtres humains, malgré une grande diversité génétique, ont deux yeux, un nez, une bouche et un appareil respiratoire complet ! Si l'on met à part les malformations graves – qui sont rares - la variabilité génétique humaine ne se répercute pas sur le

développement embryonnaire ou fœtal : les voies développementales des organes sont très stables, sans doute par l'intermédiaire d'une redondance de gènes et d'un tamponnement (« buffering ») des programmes d'expression génétiques mis en jeu. Cette observation élémentaire contraste avec la grande variabilité des lignées dans l'efficacité de la différenciation dirigée de nombreux tissus (Cahan and Daley, 2013; Kajiwara et al., 2012). La raison de cette variabilité est sans doute expliquée par les conditions de différenciation actuelles, encore largement imparfaites car elles n'offrent pas une robustesse équivalente à la différenciation *in vivo*. De ce fait, dans l'état actuel des protocoles de différenciation, certaines lignées sont incapables de se différencier correctement en tel ou tel tissu. Il apparaît ainsi que le succès de production d'un tissu est très lignée-dépendant. Je n'ai testé qu'un nombre limité de lignées CSP (n= 3) et il est possible que ces lignées ne soient pas optimales pour obtenir de l'épithélium bronchique.

Reproduire le développement embryonnaire n'est pas si aisé. Ce phénomène physiologique dure neuf mois chez l'être humain. Le compresser en quelques semaines entraîne de facto une perte d'information et de précision dans la différenciation. Il est peut-être nécessaire de prolonger les protocoles de différenciation. En pratique, dans l'état actuel des protocoles que j'ai testé, cela ne semble pas envisageable car j'ai fréquemment observé un décollement des épithéliums au delà de huit semaines.

De plus, le programme d'expression du développement pulmonaire humain est encore relativement mal connu. Nous savons, comme nous l'avons décrit plus haut, que l'endoderme définitif s'antériorise, se ventralise puis donne naissance au bourgeon pulmonaire qui s'agrandit et forme plusieurs branches. Cependant, nous ne savons pas précisément le destin de différenciation chaque cellule. La connaissance du profil d'expression ARN et protéomique, cellule par cellule, du développement pulmonaire serait une ressource unique pour mieux guider les cellules au cours des protocoles de différenciation *in vitro*, notamment en caractérisant les facteurs de croissance et les protéines de la matrice extracellulaire mis en jeu à chaque étape. Le séquençage RNA single cell est une avancée technique majeure pour atteindre cet objectif.

Chacune des étapes de différenciation en épithélium bronchique que nous avons décrites précédemment correspond à une durée très précise. Elles doivent s'enchaîner

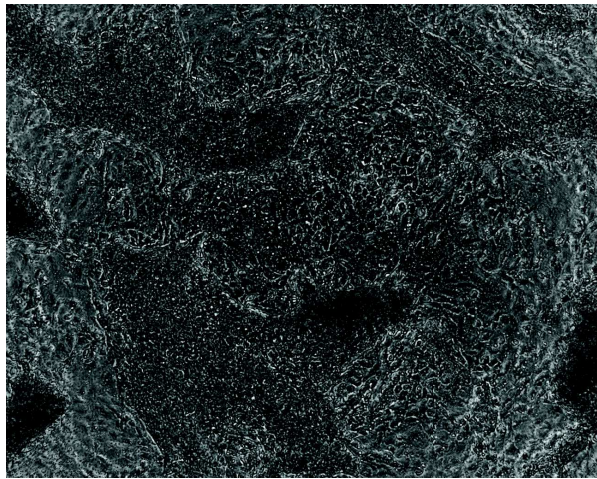
avec une rigueur parfaite. C'est ainsi que l'équipe de H.W Snoeck a remarqué que le timing d'induction d'une étape par rapport à une autre est critique (Green et al., 2011). La différenciation en AFE doit être engagée entre le cinquième jour et le sixième de différenciation en DE. Avant, la différenciation échoue. Après, la répression des marqueurs du hindgut n'est pas suffisante par les cytokines qui induisent l'AFE.

Comme le soulève l'équipe de S. Nedelec (Maury et al., 2015), les différents protocoles aujourd'hui publiés n'étudient qu'une fraction infime des combinaisons possibles. En effet, si l'on associe la matrice de différenciation, le milieu, les cytokines, la densité cellulaire au moment du passage, le nombre de passages à réaliser et le timing d'induction des étapes, la combinatoire pour trouver le meilleur protocole augmente de façon exponentielle. De plus, S. Nedelec souligne que des changements subtils affectent grandement la différenciation. Il sera donc essentiel de mettre en place des systèmes permettant de tester de vastes panoplies de combinaisons de paramètres, ce qui nécessite de miniaturiser les conditions de culture et la mise en place d'un test le plus rapide et le moins onéreux pour déterminer le taux de succès de la différenciation.

Aussi, afin de suivre l'avancée de la différenciation, il est impératif de caractériser les meilleurs marqueurs. Dans le Chapitre 2.2.1, nous avons évoqué les marqueurs qu'il nous a parus judicieux de suivre. Cependant, il n'existe pas un marqueur unique que l'on pourrait quantifier tout au long de la différenciation. Pareillement, il manque des marqueurs intermédiaires. Ceux là-même qui seraient nécessaires à identifier correctement l'entre-deux de l'endoderme définitif à l'épithélium bronchique. L'absence de marqueurs très spécifiques empêche de suivre et quantifier correctement les différentes étapes de la différenciation. L'approche par RNAseq single cell évoquée ci-dessus pourrait contribuer à mettre en évidence des marqueurs, ou des combinaison de marqueurs, notamment membranaires (pour envisager une approche par cytométrie de flux).

Parvenir à quantifier les cellules différenciées à telle ou telle étape du processus permettrait de pouvoir optimiser les protocoles de différenciation. En effet, chaque étape de différenciation offre un patchwork de cellules très hétérogène (Figure 55). Dès la première étape toutes les cellules ne se dirigent pas vers l'endoderme définitif. De 80% (Longmire et al., 2012; Mou et al., 2012; Wong et al., 2012) à 95% (Green et al., 2011;

Wong et al., 2015) des cellules se différencient en DE selon le protocole utilisé. Si dès le départ toutes les cellules ne se dirigent dans la bonne voie de différenciation, comment améliorer le pourcentage de cellules obtenues finalement dans le tissu épithélial désiré ? Une équipe japonaise trie les cellules de l'AFE ventralisé grâce au double marquage TTF1 et carboxypeptidase M (CPM) (Konishi et al., 2016). Les cellules obtenues sont alors engagées dans un protocole de différenciation en 3D afin d'obtenir des cellules épithéliales des voies aériennes proximales. Un autre marqueur pan-épithélial CXCR4 peut également être utilisé pour sélectionner les cellules parvenues au stade d'endoderme définitif afin de promouvoir l'efficacité de la différenciation (Davenport et al., 2016a).



**Figure 57 : Microscopie optique de cellules (iPCD02-30) différenciées très hétérogènes à l'étape AFE après 9 jours de différenciation selon le protocole (Wong et al., 2015).**

Les protocoles de différenciation en 2D tels que ceux que nous avons mis en œuvre visent à obtenir une population cellulaire d'intérêt la plus pure possible. Or les tissus en cours de développement ou au stade adulte sont composés de nombreux types cellulaires différents. En particulier, le poumon est constitué non seulement de cellules épithéliales bronchiques et pulmonaires, mais également de cellules fibroblastiques et de cellules endothéliales. Ces types cellulaires, en particulier le tissu mésenchymateux, jouent un rôle essentiel dans le développement de l'arbre respiratoire. Par exemple, la structuration de l'axe antéro-postérieur de l'endoderme est sous l'influence directe de facteurs de croissance sécrétés par le mésoderme, notamment les FGF, WNT and BMP (Zorn and Wells, 2009). Le rôle des vaisseaux est plus discutable, puisque le développement pulmonaire ne semble pas être entravé par l'ablation des cellules endothéliales dans des

cultures organotypiques de poumon de rongeur (Havrilak and Shannon, 2015). Il sera donc important de considérer ce point dans la mise en place de différenciation bronchique *in vitro* : faut-il inclure des cellules de type mésenchymateuses ? C'est en tout cas le parti pris de Gotoh et al. qui ont produit des organoïdes d'épithélium pulmonaire en passant par une étape de co-culture avec des fibroblastes humains foëtaux (Gotoh et al., 2014).

L'aspect dimensionnel de la culture – 2D vs 3D – est également un paramètre fondamental. Comme discuté plus haut, le développement normal se fait en 3D. Or de nombreux protocoles de différenciation proposent des cultures en 2D. Le tératome a montré l'intérêt d'une culture en 3D pour obtenir une structure mimant par certains aspects le développement initial de l'arbre bronchique. Cependant, le tératome présente de nombreuses contraintes – difficulté à intervenir pour orienter la différenciation, impossibilité d'évaluer le processus de différenciation en cours d'expérimentation, hétérogénéité des tissus et surtout, impossibilité de développer cette technologie pour effectuer des criblages à haut débit de molécules chimiques. Aussi, les organoïdes seraient mieux appropriés car il est possible de diriger la différenciation en un tissu défini et à un niveau de maturité très avancé malgré l'absence de vascularisation (Dye et al., 2015; Konishi et al., 2016). Ces structures sont organisées en 3D, et surtout, composées de différents types cellulaires, y compris de cellules de type fibroblastiques (Dye et al., 2015). Il est envisageable de les utiliser pour procéder à des criblages de médicaments (Dekkers et al., 2013).



# DISCUSSION

# 1. Perspectives

Les CSP ouvrent de nombreuses possibilités dans le domaine biomédical.

Ces cellules singulières permettent l'étude du développement embryonnaire normal ou pathologique, et de la physiopathologie de tissus sains ou malades. En effet, le patrimoine génétique des CSP de départ peut être choisi, exempt ou non de mutation, ce qui permet donc de pouvoir étudier une maladie génétique particulière.

Reproduire *in vitro* le phénotype de la DCP en modélisant l'épithélium bronchique est une preuve de principe qu'il est possible de mimer une pathologie pulmonaire avec une composante génétique dans un puits à partir de CSP choisies. Cette démonstration de principe a déjà été réalisée pour de nombreux tissus, y compris le cœur et le cerveau (Marchetto et al., 2010; Moretti et al., 2010). En se basant sur cette constatation qu'une pathologie simple se modélise *in vitro* par différenciation dirigée en 2D ou par la production d'organoïdes, il sera possible d'envisager de modéliser des pathologies plus complexes comme la BPCO.

La BPCO est une pathologie très invalidante pour les patients dont les thérapeutiques actuelles ne sont que symptomatiques. Aucun traitement curatif n'existe à ce jour. Les enjeux sont donc importants d'un point de vue médical et pharmaceutique pour les patients. Reproduire le tissu pathologique sera une aide précieuse dans le développement des thérapeutiques. En effet, les modèles animaux sont peu adaptés aux tests pharmacologiques car il existe des différences d'espèces et les modèles *in vitro* actuels – culture de biopsie de bronche de patient en interface air/liquide – n'offrent qu'une quantité limitée de tissus.

Les iPS ouvrent la possibilité de pouvoir reproduire en quantité quasi illimitée un tissu voulu avec un patrimoine génétique choisi. Il sera donc possible de cribler des molécules

candidat-médicaments. Non seulement il sera réalisable de découvrir des traitements avec une efficacité maximale et une toxicité minimale, mais également des thérapeutiques curatives en reversant le phénotype pathologique.

## 2. Limites

Malgré les nombreuses perspectives offertes, les CSP présentent également plusieurs limitations.

En effet, avant de pouvoir utiliser dans un but thérapeutique les tissus et organoïdes dérivés des CSP, il faudra (i) une production de grade clinique, (ii) s'assurer qu'aucune cellule qui compose ces nouveaux traitements ne soit atteinte d'anomalie génétique. Il est effectivement inconcevable d'implanter des cellules anormales risquant de dériver en tumeur dans un patient, quelque soit sa pathologie. Un premier essai clinique utilisant des iPS se déroule au Japon (Reardon and Cyranoski, 2014), mais cet essai a été stoppé par principe de précaution après la découverte de modifications génétiques dans la lignée d'iPS par rapport aux cellules de départ, bien que la tumorigénicité de ces mutations ne soit pas prouvée (Garber, 2015).

Une autre limite dans la différenciation des CSP est leur variabilité dans leur capacité à se différencier en tel ou tel tissu. Les protocoles actuels sont rudimentaires et ne reproduisent qu'imparfaitement les signaux de différenciation qui sont plus robustes *in vivo*. Ainsi des différences de génotype mêmes faibles donnent des disparités très marquées au niveau de la capacité des lignées de CSP à produire un tissu donné. Les aptitudes des CSP à la différenciation sont donc encore très lignée-dépendante et tissu-dépendant.

Bien qu'il soit un outil de référence pour vérifier la pluripotence des CSP, le tératome présente des limites dans la différenciation. Les cellules sont implantées dans un

microenvironnement murin : sang, cellules endothéliales et fragments d'autres tissus sur le site à injection. Mais cette technique ne permet aucune intervention du manipulateur pour privilégier une voie de différenciation. De plus, du fait notamment de l'absence de gastrulation au cours de la croissance du tératome, celui-ci ne reproduit qu'incomplètement le développement embryonnaire. Il s'agit malgré tout d'un modèle très intéressant de différenciation en 3D.

### 3. Alternative

Une alternative à la modélisation *in vitro* de l'épithélium bronchique par les CSP est le modèle de culture *ex vivo* de biopsie de bronche (Gras et al., 2012). La mise en culture d'une biopsie endobronchique permet de reproduire, après une étape de dédifférenciation puis de différenciation et grâce à une interface air-liquide (ALI), un épithélium bronchique mature. Chaque épithélium présente les caractéristiques morphologiques et fonctionnelles des pathologies dont sont atteints les patients donateurs. Les grandes restrictions de cette technique sont (i) l'acte médical pour obtenir la biopsie qui n'est pas bénin et présente comme tout acte médical un risque pour le patient, (ii) la limite dans la quantité d'épithélium produit qui, contrairement aux iPS, n'est pas illimité. Ces restrictions rendent difficile l'utilisation de ces épithélia dans le processus de criblage à large échelle de molécules candidat-médicament.

# Bibliographie

Aasen, T., Raya, A., Barrero, M.J., Garreta, E., Consiglio, A., Gonzalez, F., Vassena, R., Bilic, J., Pekarik, V., Tiscornia, G., et al. (2008). Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat Biotechnol* 26, 1276–1284.

Abubakar, M.B., and Gan, S.H. (2016). Molecular Targets in Advanced Therapeutics of Cancers: The Role of Pharmacogenetics. *Oncology* 91, 3–12.

Adhikary, S., and Eilers, M. (2005). Transcriptional regulation and transformation by Myc proteins. *Nat Rev Mol Cell Biol* 6, 635–645.

Aird, W.C. (2007). Phenotypic heterogeneity of the endothelium: II. Representative vascular beds. *Circulation Research* 100, 174–190.

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.

Aoi, T., Yae, K., Nakagawa, M., Ichisaka, T., Okita, K., Takahashi, K., Chiba, T., and Yamanaka, S. (2008). Generation of pluripotent stem cells from adult mouse liver and stomach cells. *Science* 321, 699–702.

Asprer, J.S.T., and Lakshmipathy, U. (2015). Current methods and challenges in the comprehensive characterization of human pluripotent stem cells. *Stem Cell Rev and Rep* 11, 357–372.

Assou, S., Cerecedo, D., Tondeur, S., Pantesco, V., Hovatta, O., Klein, B., Hamamah, S., and De Vos, J. (2009). A gene expression signature shared by human mature oocytes and embryonic stem cells. *BMC Genomics* 10, 10.

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* 17, 126–140.

Bai, Q., Desprat, R., Klein, B., Lemaitre, J.-M., and De Vos, J. (2013). Embryonic stem cells or induced pluripotent stem cells? A DNA integrity perspective. *Curr Gene Ther* 13, 93–98.

Bai, Q., Ramirez, J.M., Becker, F., Pantesco, V., Lavabre-Bertrand, T., Hovatta, O., Lemaitre, J.-M., Pellestor, F., and De Vos, J. (2015). Temporal analysis of genome alterations induced by single-cell passaging in human embryonic stem cells. *Stem Cells Dev* 24, 653–662.

Baker, M. (2012). Optimal stoichiometry for pluripotency. *Nat Meth* 9, 122–123.

Bao, L., He, L., Chen, J., Wu, Z., Liao, J., Rao, L., Ren, J., Li, H., Zhu, H., Qian, L., et al. (2011).

Reprogramming of ovine adult fibroblasts to pluripotency via drug-inducible expression of defined factors. *Cell Res* 21, 600–608.

Baranska, W., and Koprowski, H. (1970). Fusion of unfertilized mouse eggs with somatic cells. *J. Exp. Zool.* 174, 1–14.

Barrero, M.J., and Izpisua Belmonte, J.C. (2011). iPS cells forgive but do not forget. *Nat Cell Biol* 13, 523–525.

Beers, M.F., and Morrisey, E.E. (2011). The three R's of lung health and disease: repair, remodeling, and regeneration. *J Clin Invest* 121, 2065–2073.

Bellusci, S., Grindley, J., Emoto, H., Itoh, N., and Hogan, B.L. (1997). Fibroblast growth factor 10 (FGF10) and branching morphogenesis in the embryonic mouse lung. *Development* 124, 4867–4878.

Boland, M.J., Hazen, J.L., Nazor, K.L., Rodriguez, A.R., Gifford, W., Martin, G., Kupriyanov, S., and Baldwin, K.K. (2009). Adult mice generated from induced pluripotent stem cells. *Nature* 461, 91–94.

Bourret, R., Martinez, E., Violla, F., Giquel, C., Thonnat-Marin, A., and De Vos, J. (2016). Human-animal chimeras: ethical issues about farming chimeric animals bearing human organs. *Stem Cell Res Ther* 7, 87.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122, 947–956.

Braam, S.R., Zeinstra, L., Litjens, S., Ward-van Oostwaard, D., van den Brink, S., van Laake, L., Lebrin, F., Kats, P., Hochstenbach, R., Passier, R., et al. (2008). Recombinant vitronectin is a functionally defined substrate that supports human embryonic stem cell self-renewal via  $\alpha 5 \beta 1$  integrin. *Stem Cells* 26, 2257–2265.

Briggs, R., and King, T.J. (1952). Transplantation of Living Nuclei From Blastula Cells into Enucleated Frogs' Eggs. *Proc Natl Acad Sci USA* 38, 455–463.

BurrIDGE, P.W., Li, Y.F., Matsa, E., Wu, H., Ong, S.-G., Sharma, A., Holmström, A., Chang, A.C., Coronado, M.J., Ebert, A.D., et al. (2016). Human induced pluripotent stem cell-derived cardiomyocytes recapitulate the predilection of breast cancer patients to doxorubicin-induced cardiotoxicity. *Nat Med* 22, 547–556.

Cacabelos, R., Torrellas, C., Tejjido, O., and Carril, J.C. (2016). Pharmacogenetic considerations in the treatment of Alzheimer's disease. *Pharmacogenomics* 17, 1041–1074.

Cahan, P., and Daley, G.Q. (2013). Origins and implications of pluripotent stem cell variability and heterogeneity. *Nat Rev Mol Cell Biol* 14, 357–368.

Cao, F., van der Bogt, K.E.A., Sadrzadeh, A., Xie, X., Sheikh, A.Y., Wang, H., Connolly, A.J., Robbins, R.C., and Wu, J.C. (2007). Spatial and temporal kinetics of teratoma formation from murine embryonic stem cell transplantation. *Stem Cells Dev* 16, 883–891.

Carey, B.W., Markoulaki, S., Hanna, J.H., Faddah, D.A., Buganim, Y., Kim, J., Ganz, K., Steine, E.J., Cassady, J.P., Creighton, M.P., et al. (2011). Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. *Cell Stem Cell* 9, 588–598.

Cartwright, P., McLean, C., Sheppard, A., Rivett, D., Jones, K., and Dalton, S. (2005). LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. *Development* 132, 885–896.

Carvajal-Vergara, X., Sevilla, A., D'Souza, S.L., Ang, Y.-S., Schaniel, C., Lee, D.-F., Yang, L., Kaplan, A.D., Adler, E.D., Rozov, R., et al. (2010). Patient-specific induced pluripotent stem-cell-derived models of LEOPARD syndrome. *Nature* 465, 808–812.

Cebrian-Serrano, A., Stout, T., and Dinnyes, A. (2013). Veterinary applications of induced pluripotent stem cells: regenerative medicine and models for disease? *Vet. J.* 198, 34–42.

Chal, J., Oginuma, M., Tanoury, Al, Z., Gobert, B., Sumara, O., Hick, A., Bousson, F., Zidouni, Y., Mursch, C., Moncuquet, P., et al. (2015). Differentiation of pluripotent stem cells to muscle fiber to model Duchenne muscular dystrophy. *Nature Publishing Group* 33, 962–969.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and Smith, A.G. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* 113, 643–655.

Chambers, S.M., Fasano, C., Papapetrou, E., Tomishima, M., Sadelain, M., and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol.*

Chasman, D.I., Posada, D., Subrahmanyam, L., Cook, N.R., Stanton, V.P., and Ridker, P.M. (2004). Pharmacogenetic study of statin therapy and cholesterol reduction. *Jama* 291, 2821–2827.

Chen, F., Cao, Y., Qian, J., Shao, F., Niederreither, K., and Cardoso, W.V. (2010). A retinoic acid-dependent network in the foregut controls formation of the mouse lung primordium. *J Clin Invest* 120, 2040–2048.

Chen, G., Gulbranson, D.R., Hou, Z., Bolin, J.M., Ruotti, V., Probasco, M.D., Smuga-Otto, K., Howden, S.E., Diol, N.R., Propson, N.E., et al. (2011). Chemically defined conditions for human iPSC derivation and culture. *Nat Meth* 8, 424–429.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133, 1106–1117.

Chew, J.-L., Loh, Y.-H., Zhang, W., Chen, X., Tam, W.-L., Yeap, L.-S., Li, P., Ang, Y.-S., Lim, B., Robson, P., et al. (2005). Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol Cell Biol* 25, 6031–6046.

- Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., Ambartsumyan, G., Aimiwu, O., Richter, L., Zhang, J., et al. (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* 5, 111–123.
- Choi, J., Lee, S., Mallard, W., Clement, K., Tagliazucchi, G.M., Lim, H., Choi, I.Y., Ferrari, F., Tsankov, A.M., Pop, R., et al. (2015). A comparison of genetically matched cell lines reveals the equivalence of human iPSCs and ESCs. *Nature Publishing Group* 33, 1173–1181.
- Cooke, M.J., Stojković, M., and Przyborski, S.A. (2006). Growth of teratomas derived from human pluripotent stem cells is influenced by the graft site. *Stem Cells Dev* 15, 254–259.
- D'Amour, K.A., Agulnick, A.D., Eliazer, S., Kelly, O.G., Kroon, E., and Baetge, E.E. (2005). Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nat Biotechnol* 23, 1534–1541.
- Damjanov, I., and Andrews, P.W. (2007). The terminology of teratocarcinomas and teratomas. *Nat Biotechnol* 25, 1212–discussion1212.
- Dang, D.T., Pevsner, J., and Yang, V.W. (2000). The biology of the mammalian Krüppel-like family of transcription factors. *The International Journal of Biochemistry & Cell Biology* 32, 1103–1121.
- Davenport, C., Diekmann, U., and Naujok, O. (2016a). A Quick and Efficient Method for the Purification of Endoderm Cells Generated from Human Embryonic Stem Cells. *J Vis Exp*.
- Davenport, C., Diekmann, U., Budde, I., Detering, N., and Naujok, O. (2016b). Anterior-Posterior Patterning of Definitive Endoderm Generated from Human Embryonic Stem Cells Depends on the Differential Signaling of Retinoic Acid, Wnt-, and BMP-Signaling. *Stem Cells*.
- David, L., and Polo, J.M. (2014). Phases of reprogramming. *Stem Cell Res* 12, 754–761.
- De Vos, J., Bouckenheimer, J., Sansac, C., Lemaître, J.-M., and Assou, S. (2016). Human induced pluripotent stem cells: A disruptive innovation. *Curr Res Transl Med* 64, 91–96.
- Dekkers, J.F., Wiegerinck, C.L., de Jonge, H.R., Bronsveld, I., Janssens, H.M., de Winter-de Groot, K.M., Brandsma, A.M., de Jong, N.W.M., Bijvelds, M.J.C., Scholte, B.J., et al. (2013). A functional CFTR assay using primary cystic fibrosis intestinal organoids. *Nat Med* 19, 939–945.
- Deng, Y., Liu, Q., Luo, C., Chen, S., Li, X., Wang, C., Liu, Z., Lei, X., Zhang, H., Sun, H., et al. (2012). Generation of induced pluripotent stem cells from buffalo (*Bubalus bubalis*) fetal fibroblasts with buffalo defined factors. *Stem Cells Dev* 21, 2485–2494.
- Desai, N., Rambhia, P., and Gishto, A. (2015). Human embryonic stem cell cultivation: historical perspective and evolution of xeno-free culture systems. *Reprod. Biol. Endocrinol.* 13, 9.
- Dodsworth, B.T., Flynn, R., and Cowley, S.A. (2015). The Current State of Naïve Human



Pluripotency. *Stem Cells* 33, 3181–3186.

Dye, B.R., Hill, D.R., Ferguson, M.A.H., Tsai, Y.-H., Nagy, M.S., Dyal, R., Wells, J.M., Mayhew, C.N., Nattiv, R., Klein, O.D., et al. (2015). In vitro generation of human pluripotent stem cell derived lung organoids. *Elife* 4.

Ebert, A.D., Yu, J., Rose, F.F., Mattis, V.B., Lorson, C.L., Thomson, J.A., and Svendsen, C.N. (2009). Induced pluripotent stem cells from a spinal muscular atrophy patient. *Nature* 457, 277–280.

Eichelbaum, M., and Evert, B. (1996). Influence of pharmacogenetics on drug disposition and response. *Clin. Exp. Pharmacol. Physiol.* 23, 983–985.

Eiraku, M., Takata, N., Ishibashi, H., Kawada, M., Sakakura, E., Okuda, S., Sekiguchi, K., Adachi, T., and Sasai, Y. (2011). Self-organizing optic-cup morphogenesis in three-dimensional culture. *Nature* 472, 51–56.

Eiraku, M., Watanabe, K., Matsuo-Takasaki, M., Kawada, M., Yonemura, S., Matsumura, M., Wataya, T., Nishiyama, A., Muguruma, K., and Sasai, Y. (2008). Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals. *Cell Stem Cell* 3, 519–532.

Eminli, S., Foudi, A., Stadtfeld, M., Maherali, N., Ahfeldt, T., Mostoslavsky, G., Hock, H., and Hochedlinger, K. (2009). Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. *Nature Genetics* 41, 968–976.

Eminli, S., Utikal, J., Arnold, K., Jaenisch, R., and Hochedlinger, K. (2008). Reprogramming of neural progenitor cells into induced pluripotent stem cells in the absence of exogenous Sox2 expression. *Stem Cells* 26, 2467–2474.

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

Feng, B., Jiang, J., Kraus, P., Ng, J.-H., Heng, J.-C.D., Chan, Y.-S., Yaw, L.-P., Zhang, W., Loh, Y.-H., Han, J., et al. (2009a). Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nat Cell Biol* 11, 197–203.

Feng, B., Ng, J.-H., Heng, J.-C.D., and Ng, H.-H.H. (2009b). Molecules that promote or enhance reprogramming of somatic cells to induced pluripotent stem cells. *Cell Stem Cell* 4, 301–312.

Firth, A.L., Dargitz, C.T., Qualls, S.J., Menon, T., Wright, R., Singer, O., Gage, F.H., Khanna, A., and Verma, I.M. (2014). Generation of multiciliated cells in functional airway epithelia from human induced pluripotent stem cells. *Proceedings of the National Academy of Sciences* 111, E1723–E1730.

Fliegauf, M., Benzing, T., and Omran, H. (2007). When cilia go bad: cilia defects and ciliopathies. *Nat Rev Mol Cell Biol* 8, 880–893.

Friedrich Ben-Nun, I., Montague, S.C., Houck, M.L., Tran, H.T., Garitaonandia, I., Leonardo,

T.R., Wang, Y.-C., Charter, S.J., Laurent, L.C., Ryder, O.A., et al. (2011). Induced pluripotent stem cells from highly endangered species. *Nat Meth* 8, 829–831.

Fusaki, N., Ban, H., Nishiyama, A., Saeki, K., and Hasegawa, M. (2009). Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proc. Jpn. Acad., Ser. B, Phys. Biol. Sci.* 85, 348–362.

Garber, K. (2015). RIKEN suspends first clinical trial involving induced pluripotent stem cells. *Nature Publishing Group* 33, 890–891.

Gertow, K., Wolbank, S., Rozell, B., Sugars, R., Andäng, M., Parish, C.L., Imreh, M.P., Wendel, M., and Ahrlund Richter, L. (2004). Organized development from human embryonic stem cells after injection into immunodeficient mice. *Stem Cells Dev* 13, 421–435.

Ghosh, Z., Wilson, K.D., Wu, Y., Hu, S., Quertermous, T., and Wu, J.C. (2010). Persistent donor cell gene expression among human induced pluripotent stem cells contributes to differences with human embryonic stem cells. *PLoS ONE* 5, e8975.

Gomperts, B.N. (2014). Induction of multiciliated cells from induced pluripotent stem cells. *Proceedings of the National Academy of Sciences* 111, 6120–6121.

Gonzalez, F., Boué, S., and Izpisua Belmonte, J.C. (2011). Methods for making induced pluripotent stem cells: reprogramming à la carte. *Nat Rev Genet* 12, 231–242.

Gotoh, S., Ito, I., Nagasaki, T., Yamamoto, Y., Konishi, S., Korogi, Y., Matsumoto, H., Muro, S., Hirai, T., Funato, M., et al. (2014). Generation of alveolar epithelial spheroids via isolated progenitor cells from human pluripotent stem cells. *Stem Cell Reports* 3, 394–403.

Graham, C.F. (1969). The fusion of cells with one- and two-cell mouse embryos. *Wistar Inst Symp Monogr* 9, 19–35.

Gras, D., Bourdin, A., Vachier, I., de Senneville, L., Bonnans, C., and Chanez, P. (2012). An ex vivo model of severe asthma using reconstituted human bronchial epithelium. *J. Allergy Clin. Immunol.* 129, 1259–1266.e1.

Greber, B., Lehrach, H., and Adjaye, J. (2007). Fibroblast growth factor 2 modulates transforming growth factor beta signaling in mouse embryonic fibroblasts and human ESCs (hESCs) to support hESC self-renewal. *Stem Cells* 25, 455–464.

Green, J.B., and Smith, J.C. (1990). Graded changes in dose of a *Xenopus* activin A homologue elicit stepwise transitions in embryonic cell fate. *Nature* 347, 391–394.

Green, M.D., Chen, A., Nostro, M.-C., D'Souza, S.L., Schaniel, C., Lemischka, I.R., Gouon-Evans, V., Keller, G., and Snoeck, H.-W. (2011). Generation of anterior foregut endoderm from human embryonic and induced pluripotent stem cells. *Nat Biotechnol* 29, 267–272.

Grskovic, M., Javaherian, A., Strulovici, B., and Daley, G.Q. (2011). Induced pluripotent stem cells--opportunities for disease modelling and drug discovery. *Nat Rev Drug Discov* 10, 915–

929.

Gurdon, J.B. (1962). The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *J Embryol Exp Morphol* 10, 622–640.

Gurdon, J.B., and Uehlinger, V. (1966). “Fertile” intestine nuclei. *Nature* 210, 1240–1241.

Gurdon, J.B., ELSDALE, T.R., and FISCHBERG, M. (1958). Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. *Nature* 182, 64–65.

Gutierrez-Aranda, I., Ramos-Mejía, V., Bueno, C., Muñoz-López, M., Real, P.J., Macía, A., Sánchez, L., Ligeró, G., García-Pérez, J.L., and Menéndez, P. (2010). Human Induced Pluripotent Stem Cells Develop Teratoma More Efficiently and Faster than Human Embryonic Stem Cells Regardless the Site of Injection. *Stem Cells*.

Han, X., Han, J., Ding, F., Cao, S., Lim, S.S., Dai, Y., Zhang, R., Zhang, Y., Lim, B., and Li, N. (2011). Generation of induced pluripotent stem cells from bovine embryonic fibroblast cells. *Cell Res* 21, 1509–1512.

Hanna, J.H., Markoulaki, S., Schorderet, P., Carey, B.W., Beard, C., Wernig, M., Creighton, M.P., Steine, E.J., Cassady, J.P., Foreman, R., et al. (2008). Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell* 133, 250–264.

Hanna, J.H., Wernig, M., Markoulaki, S., Sun, C.-W., Meissner, A., Cassady, J.P., Beard, C., Brambrink, T., Wu, L.-C., Townes, T.M., et al. (2007). Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. *Science* 318, 1920–1923.

Hannan, N.R.F., Segeritz, C.-P., Touboul, T., and Vallier, L. (2013). Production of hepatocyte-like cells from human pluripotent stem cells. *Nat Protoc* 8, 430–437.

Havrilak, J.A., and Shannon, J.M. (2015). Branching of lung epithelium in vitro occurs in the absence of endothelial cells. *Dev Dyn* 244, 553–563.

Henderson, J.K., Draper, J.S., Baillie, H.S., Fishel, S., Thomson, J.A., Moore, H., and Andrews, P.W. (2002). Preimplantation human embryos and embryonic stem cells show comparable expression of stage-specific embryonic antigens. *Stem Cells* 20, 329–337.

Heng, J.-C.D., Feng, B., Han, J., Jiang, J., Kraus, P., Ng, J.-H., Orlov, Y.L., Huss, M., Yang, L., Lufkin, T., et al. (2010). The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. *Cell Stem Cell* 6, 167–174.

Hewitt, K.J., and Garlick, J.A. (2013). Cellular reprogramming to reset epigenetic signatures. *Mol. Aspects Med.* 34, 841–848.

Hiyama, E., and Hiyama, K. (2007). Telomere and telomerase in stem cells. *Br J Cancer* 96, 1020–1024.

Ho, R., Chronis, C., and Plath, K. (2011). Mechanistic insights into reprogramming to induced pluripotency. *J Cell Physiol* 226, 868–878.

Honda, A., Hirose, M., Hatori, M., Matoba, S., Miyoshi, H., Inoue, K., and Ogura, A. (2010). Generation of induced pluripotent stem cells in rabbits: potential experimental models for human regenerative medicine. *Journal of Biological Chemistry* *285*, 31362–31369.

Hovatta, O., Mikkola, M., Gertow, K., Strömberg, A.-M., Inzunza, J., Hreinsson, J., Rozell, B., Blennow, E., Andäng, M., and Ahrlund Richter, L. (2003). A culture system using human foreskin fibroblasts as feeder cells allows production of human embryonic stem cells. *Hum Reprod* *18*, 1404–1409.

Hrvatin, S., O'Donnell, C.W., Deng, F., Millman, J.R., Pagliuca, F.W., Dilorio, P., Rezanian, A., Gifford, D.K., and Melton, D.A. (2014). Differentiated human stem cells resemble fetal, not adult,  $\beta$  cells. *Proceedings of the National Academy of Sciences* *111*, 3038–3043.

Huang, S.X.L., Green, M.D., de Carvalho, A.T., Mumau, M., Chen, Y.-W., D'Souza, S.L., and Snoeck, H.-W. (2015). The in vitro generation of lung and airway progenitor cells from human pluripotent stem cells. *Nat Protoc* *10*, 413–425.

Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., and Melton, D.A. (2008a). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nat Biotechnol* *26*, 795–797.

Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., Muhlestein, W., and Melton, D.A. (2008b). Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nature Publishing Group* *26*, 1269–1275.

Huch, M., and Koo, B.-K. (2015). Modeling mouse and human development using organoid cultures. *Development* *142*, 3113–3125.

Hughes, C.S., Postovit, L.M., and Lajoie, G.A. (2010). Matrigel: a complex protein mixture required for optimal growth of cell culture. *Proteomics* *10*, 1886–1890.

Ichida, J.K., Blanchard, J., Lam, K., Son, E.Y., Chung, J.E., Egli, D., Loh, K.M., Carter, A.C., Di Giorgio, F.P., Koszka, K., et al. (2009). A small-molecule inhibitor of tgf-Beta signaling replaces sox2 in reprogramming by inducing nanog. *Cell Stem Cell* *5*, 491–503.

Jaenisch, R. (2012). Nuclear cloning and direct reprogramming: the long and the short path to Stockholm. *Cell Stem Cell* *11*, 744–747.

Jaenisch, R., and Young, R.A. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* *132*, 567–582.

Jeffery, P.K. (1983). Morphologic features of airway surface epithelial cells and glands. *Am. Rev. Respir. Dis.* *128*, S14–S20.

Jiang, J., Chan, Y.-S., Loh, Y.-H., Cai, J., Tong, G.-Q., Lim, C.-A., Robson, P., Zhong, S., and Ng, H.-H.H. (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. *Nat Cell Biol* *10*, 353–360.

Judson, R.L., Babiarz, J.E., Venere, M., and Blelloch, R.H. (2009). Embryonic stem cell-specific microRNAs promote induced pluripotency. *Nature Publishing Group* *27*, 459–461.

- Kadzik, R.S., and Morrisey, E.E. (2012). Directing lung endoderm differentiation in pluripotent stem cells. *Cell Stem Cell* *10*, 355–361.
- Kajiwara, M., Aoi, T., Okita, K., Takahashi, R., Inoue, H., Takayama, N., Endo, H., Eto, K., Toguchida, J., Uemoto, S., et al. (2012). Donor-dependent variations in hepatic differentiation from human-induced pluripotent stem cells. *Proceedings of the National Academy of Sciences* *109*, 12538–12543.
- Kamao, H., Mandai, M., Okamoto, S., Sakai, N., Suga, A., Sugita, S., Kiryu, J., and Takahashi, M. (2014). Characterization of human induced pluripotent stem cell-derived retinal pigment epithelium cell sheets aiming for clinical application. *Stem Cell Reports* *2*, 205–218.
- Kanemura, H., Go, M.J., Nishishita, N., Sakai, N., Kamao, H., Sato, Y., Takahashi, M., and Kawamata, S. (2013). Pigment epithelium-derived factor secreted from retinal pigment epithelium facilitates apoptotic cell death of iPSC. *Sci Rep* *3*, 2334.
- Kehler, J., Tolkunova, E., Koschorz, B., Pesce, M., Gentile, L., Boiani, M., Lomelí, H., Nagy, A., McLaughlin, K.J., Schöler, H.R., et al. (2004). Oct4 is required for primordial germ cell survival. *EMBO Rep* *5*, 1078–1083.
- Kim, J.B., Greber, B., Araúzo-Bravo, M.J., Meyer, J., Park, K.I., Zaehres, H., and Schöler, H.R. (2009a). Direct reprogramming of human neural stem cells by OCT4. *Nature* *461*, 649–643.
- Kim, J.B., Sebastiano, V., Wu, G., Araúzo-Bravo, M.J., Sasse, P., Gentile, L., Ko, K., Ruau, D., Ehrlich, M., van den Boom, D., et al. (2009b). Oct4-induced pluripotency in adult neural stem cells. *Cell* *136*, 411–419.
- Kim, J.B., Zaehres, H., Wu, G., Gentile, L., Ko, K., Sebastiano, V., Araúzo-Bravo, M.J., Ruau, D., Han, D.W., Zenke, M., et al. (2008). Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature* *454*, 646–650.
- Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M.J., Ji, H., Ehrlich, L.I.R., et al. (2010). Epigenetic memory in induced pluripotent stem cells. *Nature*.
- Knowles, M.R., Daniels, L.A., Davis, S.D., Zariwala, M.A., and Leigh, M.W. (2013). Primary Ciliary Dyskinesia: Recent Advances in Diagnostics, Genetics, and Characterization of Clinical Disease. *Am J Respir Crit Care Med* *130624114343007*.
- Kobari, L., Yates, F., Oudrhiri, N., Francina, A., Kiger, L., Mazurier, C., Rouzbeh, S., Nemer, El, W., Hebert, N., Giarratana, M.-C.C., et al. (2012). Human induced pluripotent stem cells can reach complete terminal maturation: in vivo and in vitro evidence in the erythropoietic differentiation model. *Haematologica*.
- Kobayashi, T., Yamaguchi, T., Hamanaka, S., Kato-Itoh, M., Yamazaki, Y., Ibata, M., Sato, H., Lee, Y.-S., Usui, J.-I., Knisely, A.S., et al. (2010). Generation of rat pancreas in mouse by interspecific blastocyst injection of pluripotent stem cells. *Cell* *142*, 787–799.
- Koivisto, H., Hyvärinen, M., Strömberg, A.-M., Inzunza, J., Matilainen, E., Mikkola, M.,

Hovatta, O., and Teerijoki, H. (2004). Cultures of human embryonic stem cells: serum replacement medium or serum-containing media and the effect of basic fibroblast growth factor. *Reprod Biomed Online* 9, 330–337.

Kolaja, K. (2014). Stem cells and stem cell-derived tissues and their use in safety assessment. *Journal of Biological Chemistry* 289, 4555–4561.

Kondoh, H., and Kamachi, Y. (2010). SOX-partner code for cell specification: Regulatory target selection and underlying molecular mechanisms. *The International Journal of Biochemistry & Cell Biology* 42, 391–399.

Konishi, S., Gotoh, S., Tateishi, K., Yamamoto, Y., Korogi, Y., Nagasaki, T., Matsumoto, H., Muro, S., Hirai, T., Ito, I., et al. (2016). Directed Induction of Functional Multi-ciliated Cells in Proximal Airway Epithelial Spheroids from Human Pluripotent Stem Cells. *Stem Cell Reports* 6, 18–25.

Kumar, D., Talluri, T.R., Anand, T., and Kues, W.A. (2015). Induced pluripotent stem cells: Mechanisms, achievements and perspectives in farm animals. *World J Stem Cells* 7, 315–328.

Kumar, R., DiMenna, L., Schrode, N., Liu, T.-C., Franck, P., Muñoz-Descalzo, S., Hadjantonakis, A.-K., Zarrin, A.A., Chaudhuri, J., Elemento, O., et al. (2013). AID stabilizes stem-cell phenotype by removing epigenetic memory of pluripotency genes. *Nature* 500, 89–92.

Kurmann, A.A., Serra, M., Hawkins, F., Rankin, S.A., Mori, M., Astapova, I., Ullas, S., Lin, S., Bilodeau, M., Rossant, J., et al. (2015). Regeneration of Thyroid Function by Transplantation of Differentiated Pluripotent Stem Cells. *Cell Stem Cell* 17, 527–542.

Kuroda, T., Tada, M., Kubota, H., Kimura, H., Hatano, S.-Y., Suemori, H., Nakatsuji, N., and Tada, T. (2005). Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. *Mol Cell Biol* 25, 2475–2485.

Kurosawa, H. (2007). Methods for inducing embryoid body formation: in vitro differentiation system of embryonic stem cells. *J. Biosci. Bioeng.* 103, 389–398.

Kurosawa, H. (2012). Application of Rho-associated protein kinase (ROCK) inhibitor to human pluripotent stem cells. *J. Biosci. Bioeng.* 114, 577–581.

Kyttälä, A., Moraghebi, R., Valensisi, C., Kettunen, J., Andrus, C., Pasumarthy, K.K., Nakanishi, M., Nishimura, K., Ohtaka, M., Weltner, J., et al. (2016). Genetic Variability Overrides the Impact of Parental Cell Type and Determines iPSC Differentiation Potential. *Stem Cell Reports* 6, 200–212.

Lapasset, L., Milhavet, O., Prieur, A., Besnard, E., Babled, A., Aït-Hamou, N., Leschik, J., Pellestor, F., Ramirez, J.M., De Vos, J., et al. (2011). Rejuvenating senescent and centenarian human cells by reprogramming through the pluripotent state. *Genes Dev* 25, 2248–2253.

- Lazzaro, D., Price, M., de Felice, M., and Di Lauro, R. (1991). The transcription factor TTF-1 is expressed at the onset of thyroid and lung morphogenesis and in restricted regions of the foetal brain. *Development* *113*, 1093–1104.
- Lee, A.S., Tang, C., Cao, F., Xie, X., van der Bogt, K., Hwang, A., Connolly, A.J., Robbins, R.C., and Wu, J.C. (2009a). Effects of cell number on teratoma formation by human embryonic stem cells. *Cell Cycle* *8*, 2608–2612.
- Lee, G., Papapetrou, E.P., Kim, H., Chambers, S.M., Tomishima, M.J., Fasano, C.A., Ganat, Y.M., Menon, J., Shimizu, F., Viale, A., et al. (2009b). Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. *Nature* *461*, 402–406.
- Lewis, B.D., Hurt, R.D., Payne, W.S., Farrow, G.M., Knapp, R.H., and Muhm, J.R. (1983). Benign teratomas of the mediastinum. *J. Thorac. Cardiovasc. Surg.* *86*, 727–731.
- Li, C., Zhou, J., Shi, G., Ma, Y., Yang, Y., Gu, J., Yu, H., Jin, S., Wei, Z., Chen, F., et al. (2009). Pluripotency can be rapidly and efficiently induced in human amniotic fluid-derived cells. *Hum Mol Genet* *18*, 4340–4349.
- Li, R., Liang, J., Ni, S., Zhou, T., Qing, X., Li, H., He, W., Chen, J., Li, F., Zhuang, Q., et al. (2010). A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell* *7*, 51–63.
- Li, Y., McClintick, J., Zhong, L., Edenberg, H.J., Yoder, M.C., and Chan, R.J. (2005). Murine embryonic stem cell differentiation is promoted by SOCS-3 and inhibited by the zinc finger transcription factor Klf4. *Blood* *105*, 635–637.
- Liang, P., Lan, F., Lee, A.S., Gong, T., Sanchez-Freire, V., Wang, Y., Diecke, S., Sallam, K., Knowles, J.W., Wang, P.J., et al. (2013). Drug screening using a library of human induced pluripotent stem cell-derived cardiomyocytes reveals disease-specific patterns of cardiotoxicity. *Circulation* *127*, 1677–1691.
- Liao, J., Cui, C., Chen, S., Ren, J., Chen, J., Gao, Y., Li, H., Jia, N., Cheng, L., Xiao, H., et al. (2009). Generation of induced pluripotent stem cell lines from adult rat cells. *Cell Stem Cell* *4*, 11–15.
- Lieu, P.T., Fontes, A., Vemuri, M.C., and Macarthur, C.C. (2013). Generation of induced pluripotent stem cells with CytoTune, a non-integrating Sendai virus. *Methods Mol Biol* *997*, 45–56.
- Liu, H., Zhu, F., Yong, J., Zhang, P., Hou, P., Li, H., Jiang, W., Cai, J., Liu, M., Cui, K., et al. (2008). Generation of induced pluripotent stem cells from adult rhesus monkey fibroblasts. *Cell Stem Cell* *3*, 587–590.
- Loh, Y.-H., Agarwal, S., Park, I.-H., Urbach, A., Huo, H., Heffner, G.C., Kim, K., Miller, J.D., Ng, K., and Daley, G.Q. (2009). Generation of induced pluripotent stem cells from human blood. *Blood* *113*, 5476–5479.
- Loh, Y.-H., Wu, Q., Chew, J.-L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J.,

Leong, B., Liu, J., et al. (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nature Genetics* 38, 431–440.

Longmire, T.A., Ikonomidou, L., Hawkins, F., Christodoulou, C., Cao, Y., Jean, J.C., Kwok, L.W., Mou, H., Rajagopal, J., Shen, S.S., et al. (2012). Efficient derivation of purified lung and thyroid progenitors from embryonic stem cells. *Cell Stem Cell* 10, 398–411.

Loring, J.F., and Rao, M.S. (2006). Establishing standards for the characterization of human embryonic stem cell lines. *Stem Cells* 24, 145–150.

Ludwig, T.E., Levenstein, M.E., Jones, J.M., Berggren, W.T., Mitchen, E.R., Frane, J.L., Crandall, L.J., Daigh, C.A., Conard, K.R., Piekarczyk, M.S., et al. (2006). Derivation of human embryonic stem cells in defined conditions. *Nat Biotechnol* 24, 185–187.

Luo, J., Suhr, S., Chang, E.A., Wang, K., Ross, P.J., Nelson, L., Venta, P., Knott, J., and Cibelli, J.B. (2011). Generation of LIF and bFGF-Dependent Induced Pluripotent Stem Cells from Canine Adult Somatic Cells. *Stem Cells Dev.*

Maherali, N., and Hochedlinger, K. (2008). Guidelines and techniques for the generation of induced pluripotent stem cells. *Cell Stem Cell* 3, 595–605.

Maherali, N., and Hochedlinger, K. (2009). Tgfbeta signal inhibition cooperates in the induction of iPSCs and replaces Sox2 and cMyc. *Curr Biol* 19, 1718–1723.

Marchetto, M.C.N., Carromeu, C., Acab, A., Yu, D., Yeo, G.W., Mu, Y., Chen, G., Gage, F.H., and Muotri, A.R. (2010). A model for neural development and treatment of rett syndrome using human induced pluripotent stem cells. *Cell* 143, 527–539.

Marchetto, M.C.N., Yeo, G.W., Kainohana, O., Marsala, M., Gage, F.H., and Muotri, A.R. (2009). Transcriptional signature and memory retention of human-induced pluripotent stem cells. *PLoS ONE* 4, e7076.

Marion, R.M., Strati, K., Li, H., Tejera, A., Schoeftner, S., Ortega, S., Serrano, M., and Blasco, M.A. (2009). Telomeres acquire embryonic stem cell characteristics in induced pluripotent stem cells. *Cell Stem Cell* 4, 141–154.

Marson, A., Foreman, R., Chevalier, B., Bilodeau, S., Kahn, M., Young, R.A., and Jaenisch, R. (2008). Wnt signaling promotes reprogramming of somatic cells to pluripotency. *Cell Stem Cell* 3, 132–135.

Marx, V. (2015). Tissue engineering: Organs from the lab. *Nature* 522, 373–377.

Matsa, E., Ahrens, J.H., and Wu, J.C. (2016). Human Induced Pluripotent Stem Cells as a Platform for Personalized and Precision Cardiovascular Medicine. *Physiological Reviews* 96, 1093–1126.

Maury, Y., Côme, J., Piskorowski, R.A., Salah-Mohellibi, N., Chevaleyre, V., Peschanski, M., Martinat, C., and Nedelec, S. (2015). Combinatorial analysis of developmental cues efficiently converts human pluripotent stem cells into multiple neuronal subtypes. *Nature Publishing Group* 33, 89–96.



McConnell, B.B., Ghaleb, A.M., Nandan, M.O., and Yang, V.W. (2007). The diverse functions of Krüppel-like factors 4 and 5 in epithelial biology and pathobiology. *Bioessays* 29, 549–557.

McCracken, K.W., Catá, E.M., Crawford, C.M., Sinagoga, K.L., Schumacher, M., Rockich, B.E., Tsai, Y.-H., Mayhew, C.N., Spence, J.R., Zavros, Y., et al. (2014). Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. *Nature* 516, 400–404.

McNeish, J., Roach, M., Hambor, J., Mather, R.J., Weibley, L., Lazzaro, J., Gazard, J., Schwarz, J., Volkmann, R., Machacek, D., et al. (2010). High-throughput screening in embryonic stem cell-derived neurons identifies potentiators of alpha-amino-3-hydroxyl-5-methyl-4-isoxazolepropionate-type glutamate receptors. *Journal of Biological Chemistry* 285, 17209–17217.

Mikkelsen, T., Hanna, J.H., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. *Nature*.

Min, H., Danilenko, D.M., Scully, S.A., Bolon, B., Ring, B.D., Tarpley, J.E., DeRose, M., and Simonet, W.S. (1998). Fgf-10 is required for both limb and lung development and exhibits striking functional similarity to *Drosophila* branchless. *Genes Dev* 12, 3156–3161.

Mitalipova, M.M., Rao, R.R., Hoyer, D.M., Johnson, J.A., Meisner, L.F., Jones, K.L., Dalton, S., and Stice, S.L. (2005). Preserving the genetic integrity of human embryonic stem cells. *Nat Biotechnol* 23, 19–20.

Miyazaki, T., Futaki, S., Suemori, H., Taniguchi, Y., Yamada, M., Kawasaki, M., Hayashi, M., Kumagai, H., Nakatsuji, N., Sekiguchi, K., et al. (2012). Laminin E8 fragments support efficient adhesion and expansion of dissociated human pluripotent stem cells. *Nature Communications* 3, 1236.

Miyoshi, K., Tsuji, D., Kudoh, K., Satomura, K., Muto, T., Itoh, K., and Noma, T. (2010). Generation of human induced pluripotent stem cells from oral mucosa. *J. Biosci. Bioeng.* 110, 345–350.

Mo, X., Li, N., and Wu, S. (2014). Generation and characterization of bat-induced pluripotent stem cells. *Theriogenology* 82, 283–293.

Moeller, K.H., Rosado-de-Christenson, M.L., and Templeton, P.A. (1997). Mediastinal mature teratoma: imaging features. *AJR Am J Roentgenol* 169, 985–990.

Mohsen-Kanson, T., Hafner, A.-L., Wdziekonski, B., Takashima, Y., Villageois, P., Carrière, A., Svensson, M., Bagnis, C., Chignon-Sicard, B., Svensson, P.-A., et al. (2014). Differentiation of human induced pluripotent stem cells into brown and white adipocytes: role of Pax3. *Stem Cells* 32, 1459–1467.

Mordwinkin, N.M., Lee, A.S., and Wu, J.C. (2013). Patient-specific stem cells and cardiovascular drug discovery. *Jama* 310, 2039–2040.

Moretti, A., Bellin, M., Welling, A., Jung, C.B., Lam, J.T., Bott-Flügel, L., Dorn, T., Goedel, A., Höhnke, C., Hofmann, F., et al. (2010). Patient-specific induced pluripotent stem-cell models for long-QT syndrome. *N Engl J Med* 363, 1397–1409.

Mou, H., Zhao, R., Sherwood, R., Ahfeldt, T., Lapey, A., Wain, J., Sicilian, L., Izvolsky, K., Musunuru, K., Cowan, C.A., et al. (2012). Generation of Multipotent Lung and Airway Progenitors from Mouse ESCs and Patient-Specific Cystic Fibrosis iPSCs. *Cell Stem Cell* 10, 385–397.

Muotri, A.R., Marchetto, M.C.N., Coufal, N.G., Oefner, R., Yeo, G., Nakashima, K., and Gage, F.H. (2010). L1 retrotransposition in neurons is modulated by MeCP2. *Nature* 468, 443–446.

Nagy, K., Sung, H.-K., Zhang, P., Laflamme, S., Vincent, P., Agha-Mohammadi, S., Woltjen, K., Monetti, C., Michael, I.P., Smith, L.C., et al. (2011). Induced pluripotent stem cell lines derived from equine fibroblasts. *Stem Cell Rev and Rep* 7, 693–702.

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nature Publishing Group* 26, 101–106.

Nakagawa, M., Takizawa, N., Narita, M., Ichisaka, T., and Yamanaka, S. (2010). Promotion of direct reprogramming by transformation-deficient Myc. *Proceedings of the National Academy of Sciences* 107, 14152–14157.

Nakanishi, M., and Otsu, M. (2012). Development of Sendai virus vectors and their potential applications in gene therapy and regenerative medicine. *Curr Gene Ther* 12, 410–416.

Nakatake, Y., Fukui, N., Iwamatsu, Y., Masui, S., Takahashi, K., Yagi, R., Yagi, K., Miyazaki, J.-I., Matoba, R., Ko, M.S.H., et al. (2006). Klf4 cooperates with Oct3/4 and Sox2 to activate the Lefty1 core promoter in embryonic stem cells. *Mol Cell Biol* 26, 7772–7782.

Narsinh, K.H., Sun, N., Sanchez-Freire, V., Lee, A.S., Almeida, P., Hu, S., Jan, T., Wilson, K.D., Leong, D., Rosenberg, J., et al. (2011). Single cell transcriptional profiling reveals heterogeneity of human induced pluripotent stem cells. *J Clin Invest* 121, 1217–1221.

Naryshkin, N.A., Weetall, M., Dakka, A., Narasimhan, J., Zhao, X., Feng, Z., Ling, K.K.Y., Karp, G.M., Qi, H., Woll, M.G., et al. (2014). Motor neuron disease. SMN2 splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. *Science* 345, 688–693.

Newman, A.M., and Cooper, J.B. (2010). Lab-specific gene expression signatures in pluripotent stem cells. *Cell Stem Cell* 7, 258–262.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Schöler, H.R., and Smith, A.G. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* 95, 379–391.

Nishimoto, M., Fukushima, A., Okuda, A., and Muramatsu, M. (1999). The gene for the embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2. *Mol Cell Biol* *19*, 5453–5465.

Nishishita, N., Takenaka, C., Fusaki, N., and Kawamata, S. (2011). Generation of human induced pluripotent stem cells from cord blood cells. *J Stem Cells* *6*, 101–108.

Nissan, X., Larribere, L., Saidani, M., Hurbain, I., Delevoye, C., Feteira, J., Lemaitre, G., Peschanski, M., and Baldeschi, C. (2011). Functional melanocytes derived from human pluripotent stem cells engraft into pluristratified epidermis. *Proc Natl Acad Sci USA*.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature Genetics* *24*, 372–376.

Odorico, J.S., Kaufman, D.S., and Thomson, J.A. (2001). Multilineage differentiation from human embryonic stem cell lines. *Stem Cells* *19*, 193–204.

Ohi, Y., Qin, H., Hong, C., Blouin, L., Polo, J.M., Guo, T., Qi, Z., Downey, S.L., Manos, P.D., Rossi, D.J., et al. (2011). Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nat Cell Biol* *13*, 541–549.

Okumura-Nakanishi, S., Saito, M., Niwa, H., and Ishikawa, F. (2005). Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. *J Biol Chem* *280*, 5307–5317.

Orlova, V.V., van den Hil, F.E., Petrus-Reurer, S., Drabsch, Y., Dijke, Ten, P., and Mummery, C.L. (2014). Generation, expansion and functional analysis of endothelial cells and pericytes derived from human pluripotent stem cells. *Nat Protoc* *9*, 1514–1531.

Ozolek, J.A., and Castro, C.A. (2011). Teratomas Derived from Embryonic Stem Cells as Models for Embryonic Development, Disease, and Tumorigenesis (InTech).

Papp, B., and Plath, K. (2013). Epigenetics of Reprogramming to Induced Pluripotency. *Cell* *152*, 1324–1343.

Peng, T., and Morrisey, E.E. (2013). Development of the pulmonary vasculature: Current understanding and concepts for the future. *Pulm Circ* *3*, 176–178.

Pera, M.F., and Tam, P.P.L. (2010). Extrinsic regulation of pluripotent stem cells. *Nature* *465*, 713–720.

Pera, M.F., Andrade, J., Houssami, S., Reubinoff, B.E., Trounson, A., Stanley, E.G., Ward-van Oostwaard, D., and Mummery, C. (2004). Regulation of human embryonic stem cell differentiation by BMP-2 and its antagonist noggin. *J Cell Sci* *117*, 1269–1280.

Petit, I., Kesner, N.S., Karry, R., Robicsek, O., Aberdam, E., Müller, F.J., Aberdam, D., and Ben-Shachar, D. (2012). Induced pluripotent stem cells from hair follicles as a cellular model for neurodevelopmental disorders. *Stem Cell Res* *8*, 134–140.

Peura, T.T., Bosman, A., and Stojanov, T. (2007). Derivation of human embryonic stem cell lines. *Theriogenology* *67*, 32–42.

- Polo, J.M., Liu, S., Figueroa, M.E., Kulalert, W., Eminli, S., Tan, K.Y., Apostolou, E., Stadtfeld, M., Li, Y., Shioda, T., et al. (2010). Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nature Publishing Group* 28, 848–855.
- Popatia, R., Haver, K., and Casey, A. (2014). Primary Ciliary Dyskinesia: An Update on New Diagnostic Modalities and Review of the Literature. *Pediatr Allergy Immunol Pulmonol* 27, 51–59.
- Prokhorova, T.A., Harkness, L.M., Frandsen, U., Ditzel, N., Schrøder, H.D., Burns, J.S., and Kassem, M. (2009). Teratoma formation by human embryonic stem cells is site dependent and enhanced by the presence of Matrigel. *Stem Cells Dev* 18, 47–54.
- Que, J., Choi, M., Ziel, J.W., Klingensmith, J., and Hogan, B.L.M. (2006). Morphogenesis of the trachea and esophagus: current players and new roles for noggin and Bmps. *Differentiation* 74, 422–437.
- Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., Mccuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc regulates transcriptional pause release. *Cell* 141, 432–445.
- Rashid, S.T., Corbineau, S., Hannan, N., Marciniak, S.J., Miranda, E., Alexander, G., Huang-Doran, I., Griffin, J., Ahrlund Richter, L., Skepper, J., et al. (2010). Modeling inherited metabolic disorders of the liver using human induced pluripotent stem cells. *J Clin Invest* 120, 3127–3136.
- Rashid, T., Kobayashi, T., and Nakauchi, H. (2014). Revisiting the flight of Icarus: making human organs from PSCs with large animal chimeras. *Cell Stem Cell* 15, 406–409.
- Raya, A., Rodríguez-Pizà, I., Guenechea, G., Vassena, R., Navarro, S., Barrero, M.J., Consiglio, A., Castellà, M., Ríó, P., Sleep, E., et al. (2009). Disease-corrected haematopoietic progenitors from Fanconi anaemia induced pluripotent stem cells. *Nature* 460, 53–59.
- Reardon, S., and Cyranoski, D. (2014). Japan stem-cell trial stirs envy. *Nature* 513, 287–288.
- Ren, J., Pak, Y., He, L., Qian, L., Gu, Y., Li, H., Rao, L., Liao, J., Cui, C., Xu, X., et al. (2011). Generation of hircine-induced pluripotent stem cells by somatic cell reprogramming. *Cell Res* 21, 849–853.
- Rezania, A., Bruin, J.E., Arora, P., Rubin, A., Batushansky, I., Asadi, A., O'Dwyer, S., Quiskamp, N., Mojibian, M., Albrecht, T., et al. (2014). Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. *Nature Publishing Group* 32, 1121–1133.
- Richards, M., Tan, S.-P., Tan, J.-H., Chan, W.-K., and Bongso, A. (2004). The transcriptome profile of human embryonic stem cells as defined by SAGE. *Stem Cells* 22, 51–64.
- Rizzino, A. (2009). Sox2 and Oct-3/4: a versatile pair of master regulators that orchestrate the self-renewal and pluripotency of embryonic stem cells. *Wiley Interdiscip Rev Syst Biol Med* 1, 228–236.
- Rodda, D.J., Chew, J.-L., Lim, L.-H., Loh, Y.-H., Wang, B., Ng, H.-H.H., and Robson, P. (2005).

Transcriptional regulation of nanog by OCT4 and SOX2. *J Biol Chem* **280**, 24731–24737.

Rodin, S., Antonsson, L., Niaudet, C., Simonson, O.E., Salmela, E., Hansson, E.M., Domogatskaya, A., Xiao, Z., Damdimopoulou, P., Sheikhi, M., et al. (2014). Clonal culturing of human embryonic stem cells on laminin-521/E-cadherin matrix in defined and xeno-free environment. *Nature Communications* **5**, 3195.

Rodin, S., Domogatskaya, A., Ström, S., Hansson, E.M., Chien, K.R., Inzunza, J., Hovatta, O., and Tryggvason, K. (2010). Long-term self-renewal of human pluripotent stem cells on human recombinant laminin-511. *Nature Publishing Group* **28**, 611–615.

Rosen, C., Shezen, E., Aronovich, A., Klionsky, Y.Z., Yaakov, Y., Assayag, M., Biton, I.E., Tal, O., Shakhar, G., Ben-Hur, H., et al. (2015). Preconditioning allows engraftment of mouse and human embryonic lung cells, enabling lung repair in mice. *Nat Med* **21**, 869–879.

Rouhani, F., Kumasaka, N., de Brito, M.C., Bradley, A., Vallier, L., and Gaffney, D. (2014). Genetic background drives transcriptional variation in human induced pluripotent stem cells. *PLoS Genet* **10**, e1004432.

Ruiz, S., Brennand, K., Panopoulos, A.D., Herrerías, A., Gage, F.H., and Izpisua Belmonte, J.C. (2010). High-efficient generation of induced pluripotent stem cells from human astrocytes. *PLoS ONE* **5**, e15526.

Ryan, A.K., and Rosenfeld, M.G. (1997). POU domain family values: flexibility, partnerships, and developmental codes. *Genes Dev* **11**, 1207–1225.

Samavarchi-Tehrani, P., Golipour, A., David, L., Sung, H.-K., Beyer, T.A., Datti, A., Woltjen, K., Nagy, A., and Wrana, J.L. (2010). Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* **7**, 64–77.

Sato, N., Sanjuan, I.M., Heke, M., Uchida, M., Naef, F., and Brivanlou, A.H. (2003). Molecular signature of human embryonic stem cells and its comparison with the mouse. *Dev Biol* **260**, 404–413.

Schepers, G.E., Teasdale, R.D., and Koopman, P. (2002). Twenty pairs of sox: extent, homology, and nomenclature of the mouse and human sox transcription factor gene families. *Dev Cell* **3**, 167–170.

Schmidt, R., and Plath, K. (2012). The roles of the reprogramming factors Oct4, Sox2 and Klf4 in resetting the somatic cell epigenome during induced pluripotent stem cell generation. *Genome Biol* **13**, 251.

Seki, T., and Fukuda, K. (2015). Methods of induced pluripotent stem cells for clinical application. *World J Stem Cells* **7**, 116–125.

Shalom-Feuerstein, R., Serror, L., Aberdam, E., Muller, F.-J., van Bokhoven, H., Wiman, K.G., Zhou, H., Aberdam, D., and Petit, I. (2013). Impaired epithelial differentiation of induced pluripotent stem cells from ectodermal dysplasia-related patients is rescued by the small

compound APR-246/PRIMA-1MET. *Proceedings of the National Academy of Sciences* 110, 2152–2156.

Sherwood, R.I., Maehr, R., Mazzone, E.O., and Melton, D.A. (2011). Wnt signaling specifies and patterns intestinal endoderm. *Mech Dev* 128, 387–400.

Shi, Y., Desponts, C., Do, J.T., Hahm, H.S., Schöler, H.R., and Ding, S. (2008a). Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. *Cell Stem Cell* 3, 568–574.

Shi, Y., Do, J.T., Desponts, C., Hahm, H.S., Schöler, H.R., and Ding, S. (2008b). A combined chemical and genetic approach for the generation of induced pluripotent stem cells. *Cell Stem Cell* 2, 525–528.

Shultz, L.D., Ishikawa, F., and Greiner, D.L. (2007). Humanized mice in translational biomedical research. *Nat Rev Immunol* 7, 118–130.

Shultz, L.D., Lyons, B.L., Burzenski, L.M., Gott, B., Chen, X., Chaleff, S., Kotb, M., Gillies, S.D., King, M., Mangada, J., et al. (2005). Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2R gamma null mice engrafted with mobilized human hemopoietic stem cells. *J Immunol* 174, 6477–6489.

Si-Tayeb, K., Noto, F.K., Nagaoka, M., Li, J., Battle, M.A., Duris, C., North, P.E., Dalton, S., and Duncan, S.A. (2010). Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells. *Hepatology* 51, 297–305.

Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T.W., and Smith, A.G. (2008). Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biol* 6, e253.

Singh, U., Quintanilla, R.H., Grecian, S., Gee, K.R., Rao, M.S., and Lakshminpathy, U. (2012). Novel live alkaline phosphatase substrate for identification of pluripotent stem cells. *Stem Cell Rev and Rep* 8, 1021–1029.

Smith, M.J., Webber, B.R., Mohtashami, M., Stefanski, H.E., Zúñiga-Pflücker, J.C., and Blazar, B.R. (2015). In Vitro T-Cell Generation From Adult, Embryonic, and Induced Pluripotent Stem Cells: Many Roads to One Destination. *Stem Cells* 33, 3174–3180.

Snoeck, H.-W. (2015). Progenitor cell engraftment in the lung: at last? *Nat Med* 21, 842–844.

Sommer, C.A., Stadtfeld, M., Murphy, G.J., Hochedlinger, K., Kotton, D.N., and Mostoslavsky, G. (2009). Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. *Stem Cells* 27, 543–549.

Spence, J.R., Mayhew, C.N., Rankin, S.A., Kuhar, M.F., Vallance, J.E., Tolle, K., Hoskins, E.E., Kalinichenko, V.V., Wells, S.I., Zorn, A.M., et al. (2011). Directed differentiation of human pluripotent stem cells into intestinal tissue in vitro. *Nature* 470, 105–109.

Spinelli, V., Guillot, P.V., and de Coppi, P. (2013). Induced pluripotent stem (iPS) cells from

human fetal stem cells (hFSCs). *Organogenesis* 9, 101–110.

Sridharan, R., Tchieu, J., Mason, M.J., Yachechko, R., Kuoy, E., Horvath, S., Zhou, Q., and Plath, K. (2009). Role of the murine reprogramming factors in the induction of pluripotency. *Cell* 136, 364–377.

Stadtfield, M., Apostolou, E., Akutsu, H., Fukuda, A., Follett, P., Natesan, S., Kono, T., Shioda, T., and Hochedlinger, K. (2010). Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature*.

Stadtfield, M., Brennand, K., and Hochedlinger, K. (2008a). Reprogramming of pancreatic beta cells into induced pluripotent stem cells. *Curr Biol* 18, 890–894.

Stadtfield, M., Maherali, N., Breault, D.T., and Hochedlinger, K. (2008b). Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell* 2, 230–240.

Stevens, L.C., and Little, C.C. (1954). Spontaneous Testicular Teratomas in an Inbred Strain of Mice. *Proc Natl Acad Sci USA* 40, 1080–1087.

Sturgess, J.M., Thompson, M.W., Czegledy-Nagy, E., and Turner, J.A. (1986). Genetic aspects of immotile cilia syndrome. *Am. J. Med. Genet.* 25, 149–160.

Suga, H., Kadoshima, T., Minaguchi, M., Ohgushi, M., Soen, M., Nakano, T., Takata, N., Wataya, T., Muguruma, K., Miyoshi, H., et al. (2011). Self-formation of functional adenohypophysis in three-dimensional culture. *Nature*.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663–676.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861–872.

Takasato, M., Er, P.X., Becroft, M., Vanslambrouck, J.M., Stanley, E.G., Elefanty, A.G., and Little, M.H. (2014). Directing human embryonic stem cell differentiation towards a renal lineage generates a self-organizing kidney. *Nat Cell Biol* 16, 118–126.

Takebe, T., Sekine, K., Enomura, M., Koike, H., Kimura, M., Ogaeri, T., Zhang, R.-R., Ueno, Y., Zheng, Y.-W., Koike, N., et al. (2013). Vascularized and functional human liver from an iPSC-derived organ bud transplant. *Nature* 499, 481–484.

Tamaoki, N., Takahashi, K., Tanaka, T., Ichisaka, T., Aoki, H., Takeda-Kawaguchi, T., Iida, K., Kunisada, T., Shibata, T., Yamanaka, S., et al. (2010). Dental pulp cells for induced pluripotent stem cell banking. *Journal of Dental Research* 89, 773–778.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science* 282, 1145–1147.

- Thomson, J.A., Kalishman, J., Golos, T.G., Durning, M., Harris, C.P., Becker, R.A., and Hearn, J.P. (1995). Isolation of a primate embryonic stem cell line. *Proc Natl Acad Sci USA* *92*, 7844–7848.
- Tiemann, U., Wu, G., Marthaler, A.G., Schöler, H.R., and Tapia, N. (2016). Epigenetic Aberrations Are Not Specific to Transcription Factor-Mediated Reprogramming. *Stem Cell Reports* *6*, 35–43.
- Tokuzawa, Y., Kaiho, E., Maruyama, M., Takahashi, K., Mitsui, K., Maeda, M., Niwa, H., and Yamanaka, S. (2003). Fbx15 is a novel target of Oct3/4 but is dispensable for embryonic stem cell self-renewal and mouse development. *Mol Cell Biol* *23*, 2699–2708.
- Tomioka, M., Nishimoto, M., Miyagi, S., Katayanagi, T., Fukui, N., Niwa, H., Muramatsu, M., and Okuda, A. (2002). Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. *Nucleic Acids Res* *30*, 3202–3213.
- Umeda, K., Oda, H., Yan, Q., Matthias, N., Zhao, J., Davis, B.R., and Nakayama, N. (2015). Long-term expandable SOX9+ chondrogenic ectomesenchymal cells from human pluripotent stem cells. *Stem Cell Reports* *4*, 712–726.
- Vallier, L., Alexander, M., and Pedersen, R.A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J Cell Sci* *118*, 4495–4509.
- Vallier, L., Mendjan, S., Brown, S., Chng, Z., Teo, A., Smithers, L.E., Trotter, M.W.B., Cho, C.H.-H., Martinez, A., Rugg-Gunn, P., et al. (2009). Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. *Development* *136*, 1339–1349.
- van den Berg, D.L.C., Zhang, W., Yates, A., Engelen, E., Takacs, K., Bezstarosti, K., Demmers, J., Chambers, I., and Poot, R.A. (2008). Estrogen-related receptor beta interacts with Oct4 to positively regulate Nanog gene expression. *Mol Cell Biol* *28*, 5986–5995.
- Verma, R., Holland, M.K., Temple-Smith, P., and Verma, P.J. (2012). Inducing pluripotency in somatic cells from the snow leopard (*Panthera uncia*), an endangered felid. *Theriogenology* *77*, 220–8–228.e1–2.
- Watson, C.L., Mahe, M.M., Múnera, J., Howell, J.C., Sundaram, N., Poling, H.M., Schweitzer, J.I., Vallance, J.E., Mayhew, C.N., Sun, Y., et al. (2014). An in vivo model of human small intestine using pluripotent stem cells. *Nat Med* *20*, 1310–1314.
- West, F.D., Terlouw, S.L., Kwon, D.J., Mumaw, J.L., Dhara, S.K., Hasneen, K., Dobrinsky, J.R., and Stice, S.L. (2010). Porcine Induced Pluripotent Stem Cells Produce Chimeric Offspring. *Stem Cells Dev*.
- Willyard, C. (2015). The boom in mini stomachs, brains, breasts, kidneys and more. *Nature* *523*, 520–522.
- Wong, A.P., Bear, C.E., Chin, S., Pasceri, P., Thompson, T.O., Huan, L.-J., Ratjen, F., Ellis, J., and Rossant, J. (2012). Directed differentiation of human pluripotent stem cells into mature



airway epithelia expressing functional CFTR protein. *Nat Biotechnol*.

Wong, A.P., Chin, S., Xia, S., Garner, J., Bear, C.E., and Rossant, J. (2015). Efficient generation of functional CFTR-expressing airway epithelial cells from human pluripotent stem cells. *Nat Protoc* *10*, 363–381.

Wright, A.J., and Andrews, P.W. (2009). Surface marker antigens in the characterization of human embryonic stem cells. *Stem Cell Res* *3*, 3–11.

Xiao, L., Yuan, X., and Sharkis, S.J. (2006). Activin A maintains self-renewal and regulates fibroblast growth factor, Wnt, and bone morphogenic protein pathways in human embryonic stem cells. *Stem Cells* *24*, 1476–1486.

Xu, R.-H., Peck, R.M., Li, D.S., Feng, X., Ludwig, T., and Thomson, J.A. (2005). Basic FGF and suppression of BMP signaling sustain undifferentiated proliferation of human ES cells. *Nat Meth* *2*, 185–190.

Xu, R.-H., Sampsel-Barron, T.L., Gu, F., Root, S., Peck, R.M., Pan, G., Yu, J., Antosiewicz-Bourget, J., Tian, S., Stewart, R., et al. (2008). NANOG is a direct target of TGFbeta/activin-mediated SMAD signaling in human ESCs. *Cell Stem Cell* *3*, 196–206.

Yang, S., Lin, G., Deng, L., and Lu, G.X. (2012). Tumorigenic characteristics of embryonal carcinoma cells as a model for studying tumour progression of human embryonic stem cells. *Cell Proliferation* *45*, 299–310.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J.E., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* *318*, 1917–1920.

Zhang, J., Ratanasirintrao, S., Chandrasekaran, S., Wu, Z., Ficarro, S.B., Yu, C., Ross, C.A., Cacchiarelli, D., Xia, Q., Seligson, M., et al. (2016). LIN28 Regulates Stem Cell Metabolism and Conversion to Primed Pluripotency. *Cell Stem Cell* *19*, 66–80.

Zhang, J., Lian, Q., Zhu, G., Zhou, F., Sui, L., Tan, C., Mutalif, R.A., Navasankari, R., Zhang, Y., Tse, H.-F., et al. (2011). A human iPSC model of Hutchinson Gilford Progeria reveals vascular smooth muscle and mesenchymal stem cell defects. *Cell Stem Cell* *8*, 31–45.

Zhao, X.-Y., Li, W., Lv, Z., Liu, L., Tong, M., Hai, T., Hao, J., Guo, C.-L., Ma, Q.-W., Wang, L., et al. (2009). iPSC cells produce viable mice through tetraploid complementation. *Nature* *461*, 86–90.

Zhao, Y., Yin, X., Qin, H., Zhu, F., Liu, H., Yang, W., Zhang, Q., Xiang, C., Hou, P., Song, Z., et al. (2008). Two supporting factors greatly improve the efficiency of human iPSC generation. *Cell Stem Cell* *3*, 475–479.

Zorn, A.M., and Wells, J.M. (2009). Vertebrate Endoderm Development and Organ Formation. *Annu Rev Cell Dev Biol* *25*, 221–251.

**Titre : « MODELISATION DE L'EPITHELIUM BRONCHIQUE HUMAIN PAR LA TECHNOLOGIE DES CELLULES SOUCHES PLURIPOTENTES INDUITES (iPS) »**

**Mots clés :** épithélium bronchique, iPS, tératome, différenciation

Les cellules souches pluripotentes (CSP) incluent les cellules souches embryonnaires (ES) et les cellules souches pluripotentes induites (iPS). Elles sont définies par deux propriétés fondamentales : l'auto-renouvellement et la capacité à se différencier dans tous les types cellulaires. Les ES sont dérivées de la masse cellulaire de l'embryon. Elles soulèvent l'intérêt de la communauté scientifique du fait de leur capacité à générer tous les tissus. Il s'agit d'un outil biotechnologique majeur dont les applications thérapeutiques et pharmacologiques comporteront notamment la médecine régénératrice, la modélisation *in vitro* de maladies humaines et le criblage de candidat-médicaments. Cependant l'utilisation d'embryons humains pour générer les ES soulève des problèmes éthiques. Les iPS contournent ces difficultés car elles sont dérivées de cellules somatiques différenciées. En effet, S. Yamanaka, qui a reçu le prix Nobel en 2012, a découvert en 2006 une technique simple de reprogrammation cellulaire. L'expression transitoire de quatre gènes (OCT4, SOX2, c-MYC and KLF4) est suffisante pour reprogrammer des fibroblastes murins en iPS. Ces cellules iPS ont la même morphologie et les mêmes propriétés que les cellules ES. L'année suivante, S. Yamanaka a appliqué avec succès son cocktail à des fibroblastes humains pour produire des iPS humaines (hiPS). Les hiPS peuvent également dépasser les problèmes immunologiques soulevés par l'utilisation d'ES dans la thérapie cellulaire, par le simple fait que les hiPS pourront être dérivées du patient à traiter. De plus, parce qu'il est possible de choisir les cellules du donneur à reprogrammer selon son génotype, il est plus facile de modéliser des maladies génétiques à partir d'hiPS que d'ES. Enfin, d'un point de vue pharmaceutique, les hiPS peuvent fournir une plateforme quasi-infinie pour le criblage de molécules afin de traiter diverses pathologies.

Le but de mon projet de recherche est l'utilisation de la technologie hiPS afin de modéliser le développement de l'épithélium bronchique. Premièrement, *in vivo*, des tératomes ont été générés après injection d'hiPS dans des souris immunodéficientes. Les tératomes démontrent la capacité de nos hiPS à se différencier en épithélium bronchique. Secondairement, *in vitro*, reproduire le développement embryonnaire et fœtal permet d'offrir une méthode simple pour modéliser l'épithélium bronchique dans un puits.

Cette technologie ouvre la voie vers de nombreuses recherches, du criblage de molécules à la production de cellules souches pour réparer l'épithélium bronchique, et *in fine* à la promotion de nouveaux traitements pharmacologiques ou de thérapie innovante pour les maladies respiratoires.

**Title : « MODELING HUMAN BRONCHIAL EPITHELIUM BY INDUCED PLURIPOTENT STEM CELL (iPS) TECHNOLOGY »**

**Keywords :** bronchial epithelium, iPS, teratoma, differentiation

Pluripotent stem cells (PSC) include embryonic stem cells (ES) and induced pluripotent stem cells (iPS). They are defined by two fundamental properties: self-renewal and the capacity to differentiate into all cell types. ES cells are derived from the inner cell mass of embryos. They arouse the interest of the scientific community in particular for their ability to generate all tissues. They provide major therapeutic and pharmacological applications, including regenerative medicine, *in vitro* modelling of human diseases and molecular screening. However, the use of human blastocysts to generate ES cells raises many ethical problems. iPS circumvent these ethical issues as they can be derived from differentiated somatic tissues. Indeed, S. Yamanaka, Nobel Prize in 2012, discovered in 2006 a simple technique of cellular reprogramming. The transient expression of four genes (OCT4, SOX2, c-MYC and KLF4) is sufficient to reprogram mouse fibroblasts into iPS. These iPS cells have the same morphology and the same properties than ES cells. The following year, S. Yamanaka applied successfully his cocktail to human fibroblasts to produce human iPS (hiPS). hiPS may also overcome immunological problems raised by the use of ES cell for cellular therapy, as hiPS can be derived from the patient to be treated. In addition, it is easier to model genetic diseases from hiPS than ES, because it is possible to choose the donor cells to reprogram according to its genotype. Finally, from a pharmacological point of view, hiPS can provide a broad platform of molecular screening to treat various diseases.

The aim of my research project is to use the hiPS technology to model the development of bronchial epithelium. First, *in vivo*, teratomas were formed by the injection of hiPS into immunodeficient mice. Teratomas highlight the ability of differentiation of our hiPS into bronchial epithelium. Second, *in vitro*, reproducing embryonic and foetal bronchial development provides a way to model bronchial epithelium in a dish.

These techniques open the door to many potential research avenues from screening small molecules to engineering stem cells to repair bronchial epithelium, and will *in fine* promote new pharmacologic or cell-based treatments for respiratory diseases.