



**HAL**  
open science

# Identification et caractérisation de nouveaux gènes soumis à l’empreinte parentale dans le placenta humain

Laïla El Khattabi

► **To cite this version:**

Laïla El Khattabi. Identification et caractérisation de nouveaux gènes soumis à l’empreinte parentale dans le placenta humain. Génétique. Université Sorbonne Paris Cité, 2015. Français. NNT : 2015US-PCB227 . tel-03617565

**HAL Id: tel-03617565**

**<https://theses.hal.science/tel-03617565>**

Submitted on 23 Mar 2022

**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.

UNIVERSITE RENE DESCARTES - PARIS 5  
ECOLE DOCTORALE BIO SPC

Année 2015

---

## **THESE DE DOCTORAT EN SCIENCES**

**Spécialité Génétique**

**Identification et caractérisation de  
nouveaux gènes soumis à l’empreinte  
parentale dans le placenta humain**

**Par Laïla EL KHATTABI**

**Sous la direction du Dr Sandrine BARBAUX**

Présentée et soutenue publiquement le 27 Novembre 2015

Composition du jury :

RAPPORTEURS	Madame le Professeur Marie-Laure KOTTLER Madame le Professeur Patricia FAUQUE
EXAMINATEURS	Monsieur le Professeur François GOFFINET Madame le Professeur Irène NETCHINE Madame le Docteur Hélène KIEFER
DIRECTEUR	Madame le Docteur Sandrine BARBAUX

## Remerciements

Aux **Professeurs Marie-Laure Kottler et Patricia Fauque**, pour m'avoir fait l'honneur d'accepter d'être rapporteurs de cette thèse de doctorat.

Aux **Professeur Irène Netchine, Docteur Hélène Kiefer** et au **Professeur François Goffinet** pour m'avoir fait l'honneur d'accepter d'évaluer ce travail en faisant partie du jury de cette thèse.

Au **Docteur Sandrine Barboux**, pour toutes les connaissances qu'elle m'a transmises en encadrant ce travail, sa générosité et sa patience durant ces trois années de thèse dans des conditions loin d'être optimales, ainsi que pour l'attention et le soutien qu'elle m'a apportés tout au long de la réalisation de cette thèse, pour ses précieux conseils et ses mots bienveillants quand il fallait.

Au **Docteur Daniel Vaiman**, pour m'avoir ouvert les portes de son laboratoire, son accueil dans son équipe et ses conseils.

Aux Docteurs Ahmed Ziyat, Céline Mehats, Julie Coquet, Aminata Touré et Brigitte Lefevre, ainsi que Patrick, Côme et Laurence, pour leur amitié, leur soutien scientifique et moral et pour les discussions très constructives sur fond scientifique ou pas ;-).

A Charlotte, Elma et Aurélien, mes « co », pour leur bonne humeur, leurs conseils et astuces techniques et leurs encouragements. A Jean et François pour leur aide et conseils.

Aux Docteurs Stéphanie Backer et Evelyne Bloch-Gallego, pour avoir fait de leur laboratoire mon 3<sup>ème</sup> QG !, pour leur aide et tout ce qu'elles m'ont appris sur le cerveau murin et autres.

A tout l'Institut Cochin pour la dynamique, l'ambiance « familiale » et esprit d'équipe.

Au **Professeur Jean Michel Dupont**, pour sa confiance en m'ayant recrutée en tant qu'Assistant Hospitalo-Universitaire et en me soutenant pour la suite de ma carrière, son écoute tout le long de ces trois années pas toujours faciles entre les obligations hospitalières, universitaires et de recherche.

A mes chères collègues, soutiens et amies de l'hôpital, notamment Dominique Letessier, Nathalie Le Dû, Aziza Lebbar, Audrey Coustier et le reste de l'équipe du Service de Cytogénétique.

A l'équipe de Biochimie Génétique, Juliette Nectoux, Béatrice Parfait, Eric Pasmant, Magalie, Franck Leturcq, Thierry Bienvenu, Gilles, Aurélie, Philippe, Lucie, Corinne, Nathalie, JeanJean, Joelle.... Et sûr le Pr Michel Vidaud, pour m'avoir acceptée dans leur laboratoire surtout les fins de journées et les

weekends !, et pour m'avoir permis d'élargir ma palette de compétences durant ces trois années en parallèle de mon doctorat, en développant des tests de dépistage prénatal non-invasif en PCR digitale et en séquençage massivement parallèle.

A mes chers parents pour s'être tant investis en faisant de mes études l'une de leurs priorités, merci pour m'avoir offert la possibilité de poursuivre mes études en France, pour votre bienveillance et votre soutien indéfectible. Merci d'avoir toujours répondu présents et de m'avoir si bien soutenue et encouragée. Avec tout mon amour.

A Mayssa, ma fille adorée, ma raison de vivre, mon moteur dans la vie, sans qui ma vie serait bien plus triste.

A Emad, mon cher époux, pour son amour, soutien, sa patience tout le long de mon interminable parcours et ses précieux conseils.

A mes frères, Lotfi et Hicham, pour leur amour, leur soutien et leurs encouragements.

A mes amis d'ici et d'ailleurs, Charlotte D, Eva P, Céline F, Nathalie S, Adeline T, Aurélie M, Imane B, Sandra, Nadia H, Rabeaa A, Hanane J, Chahrazad M, pour leur amitié et leur soutien indéfectible.

A mes co-internes devenus mes amis : Nassera, Bassem, Marie-Laure, Ciri, Candice, Morgane, Maha, Jonathan, pour leur amitié et pour tous les agréables moments passés avec eux.

# TABLE DES MATIERES

<b>LISTE DES ABBREVIATIONS .....</b>	<b>7</b>
<b>RESUME .....</b>	<b>8</b>
<b>INTRODUCTION.....</b>	<b>10</b>
<b>Chapitre I : L’empreinte parentale chez les mammifères : définition et mécanismes moléculaires.....</b>	<b>10</b>
<b>1. Définition et mise en évidence de l’empreinte parentale .....</b>	<b>10</b>
1.1. Définition de l’empreinte parentale.....	10
1.2. Mise en évidence de l’empreinte parentale .....	11
1.2.1. Expériences de complémentation chromosomique .....	12
1.2.2. Expériences de transfert de pronoyaux .....	13
<b>2. Mécanismes moléculaires .....</b>	<b>14</b>
2.1. Méthylation de l’ADN .....	15
2.1.1. Ilots CpG .....	16
2.1.2. Régions différentiellement méthylées .....	17
2.1.3. Enzymes de méthylation de l’ADN .....	19
2.1.4. Signification biologique de la méthylation de l’ADN.....	23
2.2. Modifications d’histones .....	25
2.2.1. Les marques des histones.....	25
2.2.2. Héritabilité des marques d’histones.....	26
2.2.3. Signification biologique des modifications des histones.....	29
2.3. ARN non codants .....	33
2.3.1. Définitions et fonctions .....	33
2.3.2. Les ARNnc longs (lncRNA).....	34
<b>3. Le cycle de l’empreinte .....</b>	<b>36</b>
3.1. Etablissement de l’empreinte .....	38
3.2. Maintien de l’empreinte.....	39
3.3. Effacement de l’empreinte.....	41
<b>Chapitre II : Les gènes soumis à l’empreinte parentale .....</b>	<b>41</b>
<b>1. Caractéristiques des GSE.....</b>	<b>46</b>
<b>2. Régulation de l’empreinte.....</b>	<b>46</b>
2.1. Deux grands modèles de régulation de l’empreinte .....	46
2.2. Dynamique de l’empreinte parentale .....	49

2.3.	Réseau de GSE .....	52
<b>3.</b>	<b>Méthodes d'identification des GSE.....</b>	<b>53</b>
3.1.	L'approche clinique et expérimentale.....	53
3.2.	L'approche globale .....	54
3.2.1.	Analyse du transcriptome .....	54
3.2.2.	Analyse de l'épigénome .....	56
3.3.	L'approche <i>in silico</i> .....	56
	<b>Chapitre III : Fonctions et altérations de l'empreinte parentale .....</b>	<b>57</b>
<b>1.</b>	<b>L'empreinte parentale et le développement placentaire.....</b>	<b>58</b>
1.1.	Le placenta .....	58
	Le placenta humain versus le placenta murin.....	59
1.2.	L'empreinte parentale et le placenta .....	60
<b>2.</b>	<b>L'empreinte parentale et le développement embryonnaire .....</b>	<b>64</b>
<b>3.</b>	<b>L'empreinte parentale et le développement postnatal.....</b>	<b>66</b>
<b>4.</b>	<b>Intérêt évolutif de l'empreinte .....</b>	<b>67</b>
<b>5.</b>	<b>Pathologies humaines liées à l'empreinte parentale .....</b>	<b>69</b>
5.1.	Pathologies de cause endogène .....	69
5.1.1.	Les pathologies de l'empreinte .....	69
5.1.2.	Les pathologies de la grossesse.....	70
5.2.	Pathologies de cause exogène .....	71
5.2.1.	L'assistance médicale à la procréation.....	72
5.2.2.	L'alcool et le tabac.....	72
5.2.3.	La nutrition maternelle.....	72
5.2.4.	Les perturbateurs endocriniens .....	73
	<b>LE PROJET DE LA THESE .....</b>	<b>75</b>
<b>I.</b>	<b>Contexte du projet .....</b>	<b>75</b>
<b>II.</b>	<b>Objectif et organisation du projet .....</b>	<b>80</b>
1.	Objectif du projet .....	80
2.	Organisation du projet.....	81
<b>III.</b>	<b>Résultats .....</b>	<b>82</b>
1.	Article .....	82
2.	Résultats préliminaires .....	124
2.1.	Extension de l'étude du statut d'empreinte à d'autres mammifères .....	124
	Chez la souris.....	125

Chez le bovin.....	126
2.2. Analyse du profil d'expression de <i>DSCAM</i> dans l'embryon murin.....	126
2.3. Etudes fonctionnelles à partir des cellules HUVEC.....	127
Mise en silence de <i>DSCAM</i> par ARN interférence.....	127
Les paramètres cellulaires standards .....	128
Le test d'angiogenèse dans le Matrigel.....	129
Test de la blessure .....	130
2.4. Etude des modèles murins KO <i>Dcc</i> et KO <i>Ntn1</i> .....	132
<b>DISCUSSION .....</b>	<b>138</b>
<b>I. L'approche utilisée pour l'identification de nouveaux GSE.....</b>	<b>138</b>
<b>II. <i>DSCAM</i>, nouveau GSE .....</b>	<b>139</b>
<b>III. Le chromosome 21 et l'empreinte.....</b>	<b>140</b>
<b>IV. Rôle de <i>DSCAM</i> dans le placenta ? .....</b>	<b>143</b>
1. Fonctions connues de <i>DSCAM</i> .....	143
2. Hypothèses de fonctions placentaires de <i>DSCAM</i> .....	144
2.1. Rôle de <i>DSCAM</i> dans le développement de la structure placentaire .....	145
2.2. Rôle de <i>DSCAM</i> dans l'angiogenèse placentaire .....	146
2.3. <i>DSCAM</i> et pathologies placentaires .....	147
<b>PERSPECTIVES .....</b>	<b>148</b>
<b>BIBLIOGRAPHIE .....</b>	<b>150</b>
<b>ANNEXES.....</b>	<b>166</b>
<b>Communication 1 .....</b>	<b>166</b>
<b>Communication 2 .....</b>	<b>168</b>
<b>Bibliométrie .....</b>	<b>170</b>

## LISTE DES ABBREVIATIONS

<b>ADNc</b>	ADN complémentaire
<b>ADNg</b>	ADN génomique
<b>AMP</b>	Assistance médicale à la procréation
<b>ARNm</b>	ARN messenger
<b>ARNnc</b>	ARN non codant
<b>CTCF</b>	<i>CCCTC-binding factor</i>
<b>DMR</b>	<i>Differentially methylated region</i>
<b>DNMT</b>	<i>DNA Methyl Transferase</i>
<b>DUP</b>	Disomie uniparentale
<b>GSE</b>	Gène(s) Soumis à l'Empreinte
<b>HAT</b>	<i>Histone acetyltransferase</i>
<b>HDAC</b>	<i>Histone deacetylase</i>
<b>HMT</b>	<i>Histone methyltransferase</i>
<b>HUVEC</b>	<i>Human Umbilical Vein Endothelial Cell</i>
<b>ICR</b>	<i>Imprinting control region</i>
<b>lncRNA</b>	ARN non codant long
<b>LINE</b>	<i>Long Interspersed Nuclear Elements</i>
<b>LTR</b>	Long terminal repeat
<b>5mC</b>	5-méthyl-Cytosine
<b>miRNA</b>	micro RNA
<b>PE</b>	Pré-éclampsie
<b>piRNA</b>	<i>piwi interacting RNA</i>
<b>PRC</b>	<i>Polycomb repressive complex</i>
<b>RCIU</b>	Retard de croissance intra-utérin
<b>SINE</b>	<i>Short Interspersed Nuclear Elements</i>
<b>siRNA</b>	<i>Small interfering RNA</i>
<b>snoRNA</b>	<i>Short nucleolar RNA</i>
<b>SNP</b>	<i>Single nucleotide polymorphism</i>



## RESUME

L'empreinte parentale est un phénomène épigénétique complexe qui régule l'expression mono-allélique de certains gènes en fonction de l'origine parentale de l'allèle. Les gènes soumis à empreinte (GSE) participent au développement placentaire et fœtal. Ces gènes présentent un intérêt du point de vue fondamental notamment pour leur régulation épigénétique, leur rôle dans les échanges fœto-maternels et le développement du système nerveux. Du point de vue clinique, ils sont impliqués dans les pathologies de la grossesse comme la pré-éclampsie et le retard de croissance intra-utérin ainsi que des pathologies neuro-développementales. De nombreux syndromes cliniques associent retard de croissance intra-utérin et déficience intellectuelle suggérant des voies de signalisation communes dans le développement des tissus placentaire et nerveux. La caractérisation de ces GSE pourrait donc permettre la mise en évidence de ces voies.

Un peu moins de 100 GSE ont été identifiés à ce jour chez l'homme. Nous avons donc mené un crible systématique, en utilisant une approche haut débit, pour identifier sans *a priori* de nouveaux GSE dans le placenta humain. Cette étude pilote a permis dans un premier temps de mettre en évidence 7 nouveaux GSE, dont deux présentent, d'après la littérature, une fonction dans le système nerveux : *NTM* et *MAGI2*. Dans le présent travail, nous avons étudié 23 nouveaux gènes candidats afin de valider leur statut de GSE, ce qui a conduit à l'identification de trois nouveaux GSE (*DSCAM*, *AIM1* et *NCAM2*). Nous avons approfondi l'étude pour l'un d'entre eux, à savoir le gène *DSCAM* qui est localisé sur le chromosome 21, faisant de lui le premier GSE connu sur ce chromosome. *DSCAM* participe au guidage des neurones lors du développement du système nerveux central et est un gène candidat pour la déficience intellectuelle associée au Down Syndrome. Nous présentons ici les résultats de la caractérisation de ce GSE dans le placenta humain par l'étude (1) des mécanismes épigénétiques notamment la méthylation des îlots CpG qui pourrait gouverner la mise en place de l'empreinte, (2) de l'expression de *DSCAM* dans les placentas normaux et pathologiques, ainsi que dans les cellules trophoblastiques humaines avant et après syncytialisation. Par ailleurs, nous présentons les résultats préliminaires d'études chez des modèles animaux et cellulaires. Nous avons évalué le statut d'empreinte dans d'autres espèces (souris et bovin) et d'autres tissus (notamment le cerveau), et étudié l'effet d'un déficit des partenaires moléculaires de *DSCAM* (souris Knockout *Dcc* et *Netrin-1*). Enfin, nous avons utilisé le modèle des cellules endothéliales immortalisées HUVEC pour étudier le rôle de *DSCAM* dans l'angiogenèse placentaire.

# INTRODUCTION

# INTRODUCTION

## Chapitre I : L'empreinte parentale chez les mammifères : définition et mécanismes moléculaires

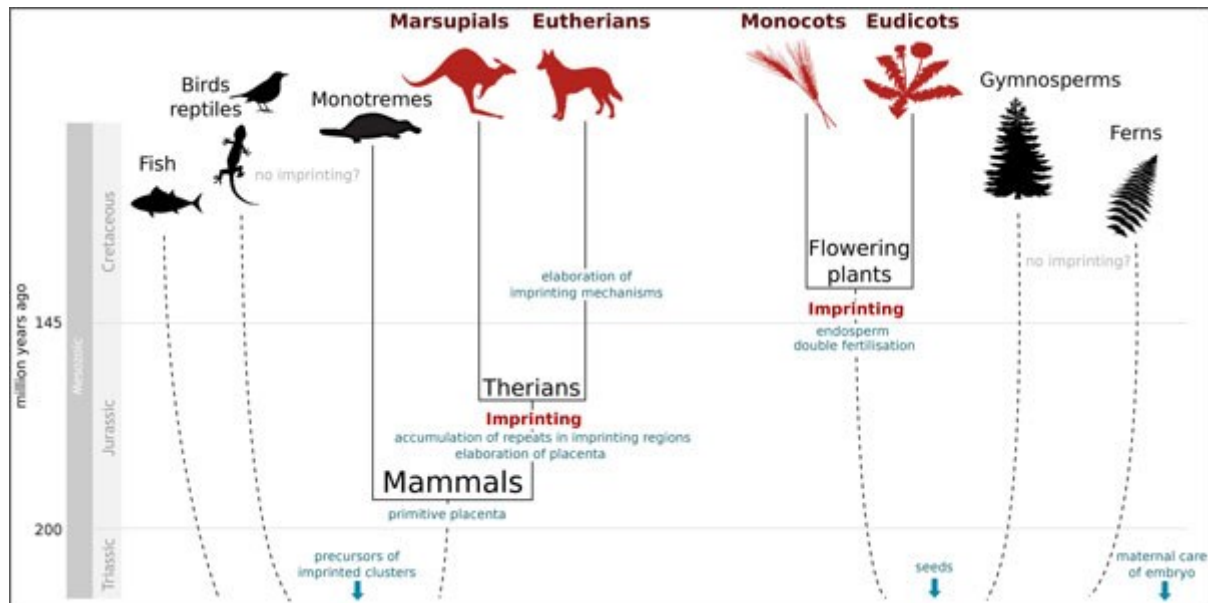
### 1. Définition et mise en évidence de l'empreinte parentale

#### 1.1. Définition de l'empreinte parentale

L'empreinte parentale caractérise une régulation épigénétique particulière de l'expression de certains gènes se traduisant par la répression de l'expression de l'un des deux allèles selon son origine parentale. Un gène soumis à empreinte (GSE) est un gène caractérisé par une **expression monoallélique dépendante de l'origine parentale de l'allèle**, ainsi seul l'allèle maternel ou l'allèle paternel est exclusivement exprimé malgré la présence des deux allèles. L'allèle non exprimé est dit soumis à empreinte car il porte des marques épigénétiques qui le rendent fonctionnellement inactif.

Comme tout processus **épigénétique**, l'empreinte parentale répond à la définition suivante : il s'agit d'une modification de l'expression des gènes sans modification du code génétique porté par les nucléotides. Ses modifications doivent être **héritables** au cours de la mitose et **réversibles**. L'empreinte est ainsi transmise de la cellule mère à la cellule fille au cours des divisions cellulaires. Néanmoins, elle est effacée dans la lignée germinale afin d'apposer la nouvelle empreinte selon qu'il s'agisse des gamètes paternels ou maternels. L'empreinte peut être restreinte à des tissus ou labile limitée à un stade particulier du développement.

En termes d'évolution, l'apparition de l'empreinte est concomitante avec l'apparition de la **viviparité** chez les mammifères. Ainsi, ce phénomène épigénétique est observé chez les marsupiaux et les euthériens. Il faut néanmoins préciser que l'empreinte est également observée chez certains insectes et plantes mais cet aspect ne sera pas développé dans ce manuscrit (Figure 1).



**Figure 1 :** Chronologie d'apparition de l'empreinte parentale au cours de l'évolution (Pires and Grossniklaus, 2014).

## 1.2. Mise en évidence de l'empreinte parentale

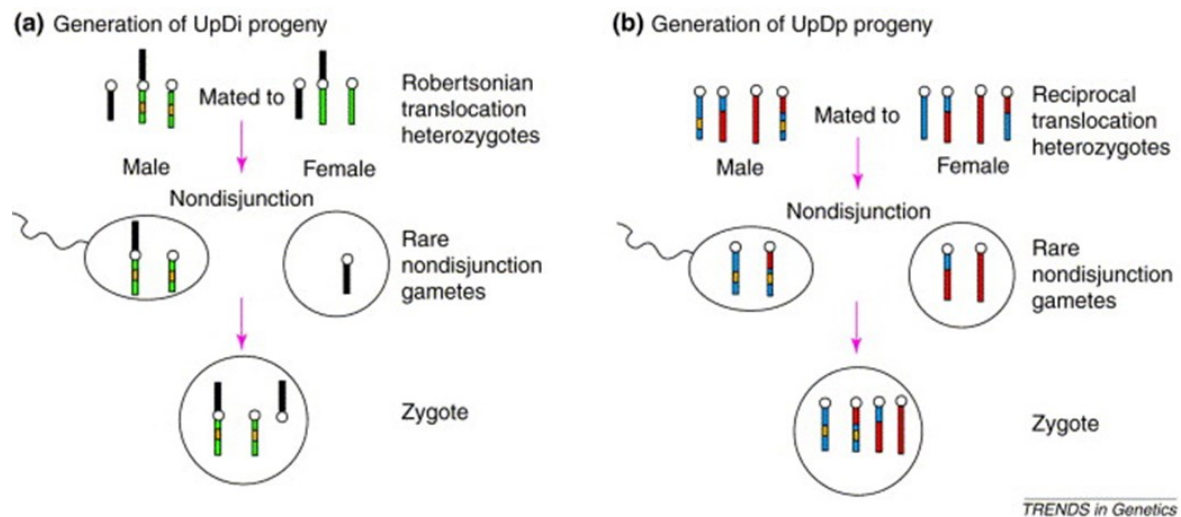
L'absence de reproduction asexuée chez les mammifères a longtemps été une intrigue biologique. De nombreuses tentatives de parthénogenèse chez la souris ont échoué. Divers postulats ont été avancés pour expliquer ces échecs, qui aujourd'hui semblent improbables avec l'évolution des connaissances : l'asynchronie entre le stade de développement du produit de parthénogenèse et la muqueuse utérine au moment de l'implantation; les différences de ratio nucléo-cytoplasmique, l'absence d'un facteur spermatique, la présence de mutations récessives létales, un déséquilibre du dosage génique des gènes du chromosome X (Graham, 1974).

Ce sont des analyses anatomo-pathologiques de tumeurs humaines qui ont suggéré une nouvelle hypothèse basée sur une différence fonctionnelle des génomes paternel et maternel. L'étude de coupes issues de tératomes ovariens et de moles hydatiformes a montré qu'il y avait une différence dans le développement de ces tissus et que cette différence était corrélée à l'origine parentale du génome qui en a été à l'origine (Kajii and Ohama, 1977; Linder et al., 1975; Wake et al., 1978). Le tératome ovarien est gynogénote et est composé de structures fœtales bien différenciées, alors que la mole hydatiforme est androgénote et ne contient que des structures extra-embryonnaires.

Les premières expériences mettant en évidence le phénomène d’empreinte parentale datent des années 1980s. Avec d’une part les expériences de parthénogenèse murine médiées par le **transfert nucléaire de pronuclei**, et d’autre part les expériences de **complémentation chromosomique** dans des modèles murins de translocations chromosomiques.

### *1.2.1. Expériences de complémentation chromosomique*

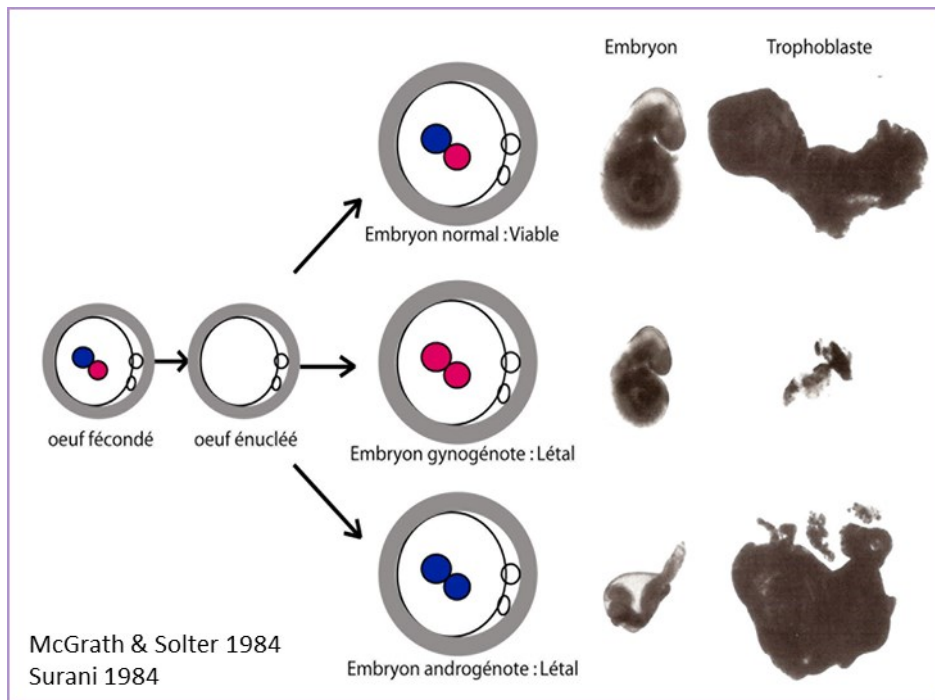
Searle & Beechey et Cattanach & Kirk ont manipulé des gamètes de souris porteuses de translocations robertsoniennes équilibrées impliquant différents chromosomes (Cattanach and Kirk, 1985; Searle and Beechey, 1978). Ils ont fécondé des ovules portant un déséquilibre (nullisomie ou disomie des chromosomes transloqués) avec des spermatozoïdes portant l’anomalie réciproque produisant ainsi des conceptus portant des disomies uniparentales (Figure 2). L’idée était d’évaluer si la létalité embryonnaire observée avec les gamètes déséquilibrés pouvait être corrigée en rééquilibrant l’anomalie chromosomique. Ces expériences ont démontré que le fait de restituer l’état équilibré de la translocation restaurait en général le « phénotype létal » sauf pour certaines régions génomiques qui restaient associées à des anomalies de comportement, de croissance et/ou de viabilité malgré l’équilibre chromosomique. Au total, 13 régions pour lesquelles l’origine parentale semblait avoir son importance ont été identifiées, qualifiées de régions à effet parental. Ces expériences murines corroboraient des observations, précédemment rapportées par des cliniciens, de maladies génétiques humaines dont la transmission ne répondait pas aux règles d’hérédité mendélienne telles que définies à l’époque. Le phénotype anormal dépendait de l’identité du parent transmetteur. La première pathologie ainsi décrite est le Syndrome de Beckwith-Wiedmann (Lubinsky et al., 1974).



**Figure 2 :** Représentation schématique des expériences de complémentation chromosomique. Des souris portant des translocations équilibrées robertsoniennes (a) ou réciproques (b) sont croisées. Les rares gamètes déséquilibrés issus d'une non disjonction méiotique portant des anomalies réciproques peuvent se compléter au moment de la fécondation pour générer un zygote équilibré mais qui portera deux copies d'un chromosome (a) ou segment chromosomique (b) provenant d'un même parent. Dans cette figure, en (a) le zygote a une disomie uniparentale paternelle du chromosome « vert » pour lequel les deux copies proviennent du père sans que la copie maternelle ne soit représentée, et en (b) le zygote porte deux copies d'une partie du chromosome « bleu » provenant du père et deux copies d'une partie du chromosome rouge provenant uniquement de la mère. (D'après Oakey & Beechey, Trends in Genet 2002 (Oakey and Beechey, 2002)).

### 1.2.2. Expériences de transfert de pronoyaux

Des expériences de génération d'embryons murins parthénogénétiques par transfert nucléaire de deux pronuclei du même parent dans un cytoplasme d'ovocyte énucléé ont montré que les génomes parentaux n'étaient pas interchangeables (McGrath and Solter, 1984; Surani and Barton, 1983). Les gestations n'arrivaient pas à leur terme ; les œufs gynogénètes aboutissaient à une hypoplasie placentaire avec un développement embryonnaire a priori normal alors que les œufs androgénètes montraient un développement placentaire normal avec un défaut de croissance de l'embryon (Figure 3). Ces premières observations ont démontré la non-équivalence fonctionnelle des deux génomes parentaux. Le développement embryonnaire et placentaire nécessite non seulement un œuf diploïde mais surtout un lot diploïde provenant de la fusion d'un pronoyau femelle et d'un pronoyau mâle.



**Figure 3 :** Expériences princeps de parthénogenèse par transfert nucléaire de pronoyaux et aspect macroscopique des trophoblastes et des embryons générés selon l'origine parentale des pronoyaux transférés (D'après : (McGrath and Solter, 1984; Surani et al., 1984)).

## 2. Mécanismes moléculaires

L'empreinte parentale est régie par un processus épigénétique complexe qui consiste en l'apposition d'une **marque épigénétique ou empreinte** sur l'un des deux allèles parentaux permettant son identification et la régulation de sa transcription. Grâce à la génétique humaine, ces mécanismes ont pu commencer à être mieux appréhendés même s'il reste beaucoup à découvrir.

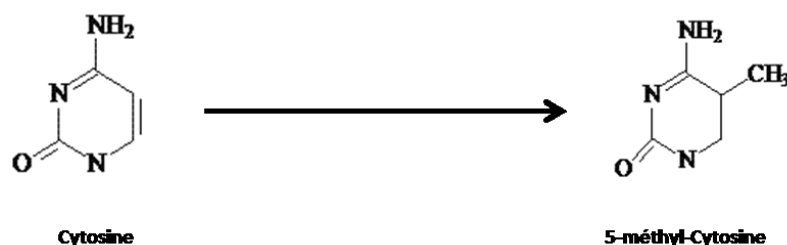
Cette régulation épigénétique de l'empreinte est médiée par tous les mécanismes épigénétiques connus, comme la **méthylation de l'ADN**, les **modifications d'histones**, les **ARN non codants** et les protéines de **remodelage de la chromatine**. L'ensemble de ces processus agit en général de façon combinatoire. Il n'y a pas un processus épigénétique universel spécifique à l'empreinte. Le mécanisme majeur serait la méthylation de l'ADN même si son implication en tant que cause ou conséquence n'est pas totalement élucidée. L'étude de différents loci soumis à empreinte parentale a montré qu'il existe différents mécanismes moléculaires qui varient d'un locus à l'autre. Par ailleurs, les mécanismes

impliqués dépendent également du tissu. Dans le placenta, par exemple, le maintien de l’empreinte serait plus largement assuré par les modifications d’histones et les ARN non codants que par la méthylation dont le taux est par ailleurs plus bas dans le placenta par rapport au reste des tissus somatiques (Lewis et al., 2004; Umlauf et al., 2004; Wagschal and Feil, 2006; Wagschal et al., 2008).

Les marques épigénétiques à l’origine de l’empreinte doivent être (1) établies dans les gamètes en fonction du sexe ; puis (2) maintenues au cours du développement en résistant à la vague de déméthylation que connaît l’embryon avant l’implantation ainsi qu’aux nombreuses divisions successives ; et enfin (3) effacées dans la nouvelle lignée germinale où une nouvelle marque sera à nouveau apposée en fonction du sexe de l’embryon. C’est ce qu’on appelle le **cycle de l’empreinte**.

## 2. 1. Méthylation de l’ADN

La méthylation de l’ADN est une marque épigénétique essentielle à la régulation de l’expression génique. Contrairement aux bactéries, les Eucaryotes ne peuvent méthyler que le résidu Cytosine en position C5 donnant naissance à la 5-méthyl-Cytosine (5mC) (Figure 4). Cette 5ème base de l’ADN a été mise en évidence en 1948 (Weissbach, 1993). Elle ne perturbe pas l’appariement avec la base complémentaire (Guanine) et ne modifie pas la structure tridimensionnelle de la double hélice d’ADN. La méthylation de la Cytosine n’est pas répartie de façon homogène dans le génome, preuve que sa mise en place n’est pas stochastique mais bel et bien régulée par des mécanismes sous-jacents. L’environnement chromatinien (histones et séquence nucléotidique) est déterminant. Cette méthylation est généralement, mais pas exclusivement, associée à une répression de la transcription.



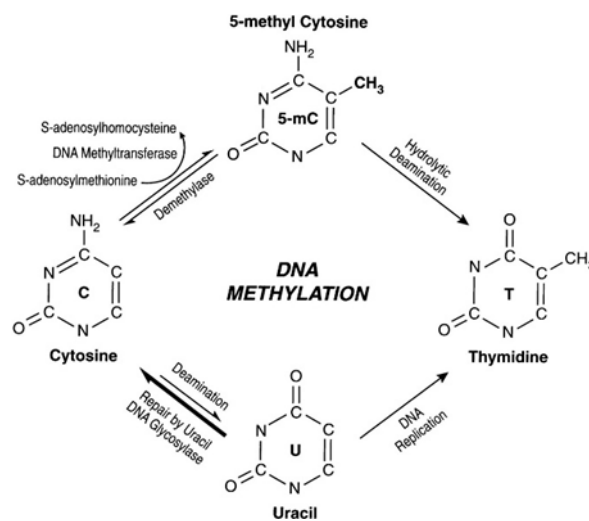
**Figure 4** : Structures chimiques de la Cytosine et de la 5-méthyl-Cytosine.



La méthylation de l'ADN est observée essentiellement au niveau de régions riches en dinucléotides CpG, c'est-à-dire 5'CG3', et est transmise lors des divisions cellulaires. La symétrie de ces dinucléotides aide à la propagation de la méthylation lors de la réplication afin de préserver la mémoire du profil transcriptionnel de la cellule (Holliday and Pugh, 1975; Riggs, 1975; Schubeler, 2015; Song et al., 2011).

### 2.1.1. Îlots CpG

Les résidus Cytosine et méthyl-Cytosine peuvent se désaminer spontanément pour donner respectivement l'Uracile ou la Thymine (Figure 5). Contrairement au produit de désamination de la Cytosine, qui lui peut être réparé par les systèmes de réparation des mésappariements de l'ADN, le produit de désamination de la méthyl-Cytosine n'est pas réparé puisqu'il est l'une des 4 bases canoniques de l'ADN. La désamination des 5mC dans la lignée germinale serait à l'origine de l'appauvrissement du génome humain en CpG au cours de l'évolution, moins de 20 % de ce qui est prédit (Illingworth and Bird, 2009).



**Figure 5 :** Transformations chimiques possibles de la Cytosine (Singal and Ginder, 1999).

Il existe néanmoins des régions qui semblent avoir été « protégées » de la désamination, dites îlots CpG. Différentes définitions de ce qu'est un îlot CpG ont été proposées mais elles s'accordent toutes sur le fait que c'est une région riche en résidus Guanine et Cytosine, et riches également en dinucléotides CpG. Ce sont Gardiner-Garden et Frommer qui ont proposé la première définition de l'îlot CpG comme étant une séquence d'au moins 200 pb

contenant plus de 50% de C+G (contre 40% dans le reste du génome) et un ratio CpG observé sur attendu supérieur à 0,6 (contre 0,25 dans le reste du génome) (Gardiner-Garden and Frommer, 1987). Cette définition a été revue en 2002 par Takai et Jones (Takai and Jones, 2002). Après analyse de données de séquençage de génomes de mammifères, ces auteurs observent l'existence de séquences « parasites » pouvant répondre aux critères de définition de Gardiner-Garden et Frommer sans pour autant être des îlots CpG. L'exemple étant les séquences Alu qui sont des répétitions d'environ 280 pb pouvant contenir une forte proportion de C+G et un ratio CpG observé sur attendu élevé. Takai et Jones ont développé un algorithme de prédiction des CpG qui tient compte de leur localisation en privilégiant les régions 5' des gènes, et exclut la majorité des séquences Alu (Takai and Jones, 2002). Ils sont arrivés ainsi à proposer une définition plus discriminante de l'îlot CpG : séquence d'au moins 500 pb, 55 % de C+G et 0,65 de ratio CpG observé / CpG attendu. Les différentes définitions de l'îlot CpG ont donné naissance à différents algorithmes de prédiction informatique de ces îlots mais aucun d'eux n'est certainement parfait.

Ces îlots sont généralement sous-méthylés dans la lignée germinale, à l'exception des régions soumises à l'empreinte, et certains d'entre eux acquièrent une méthylation tissu-spécifique au cours de la différenciation (Bird, 2002; Strichman-Almashanu et al., 2002). On les retrouve essentiellement au niveau des promoteurs d'environ 70 % des gènes chez l'Homme, à tel point qu'ils ont été proposés comme marqueurs géniques dans la quête de nouveaux gènes (Larsen et al., 1992), mais aussi dans certaines régions intergéniques (séquences répétées, transposons) et parfois à l'intérieur même du gène (Illingworth and Bird, 2009). La répartition des CpG diffère selon la fonction des gènes ; tous les gènes de ménage en contiennent contre seulement 40 % des gènes exprimés de façon tissu-spécifique (Larsen et al., 1992). Il a été montré que la chromatine associée aux îlots CpG diffère par certains points de la chromatine du reste du génome, avec (1) une perte de l'histone H1, (2) une hyper-acétylation des histones H3 et H4 et (3) la présence de régions sans nucléosome. Ce qui confère un état ouvert à la chromatine dans ces régions (Tazi and Bird, 1990).

### *2.1.2. Régions différenciellement méthylées*

Dans les situations particulières où l'expression des gènes se fait de façon mono-allélique (l'empreinte parentale, l'inactivation de l'X, l'expression mono-allélique somatique), l'un des

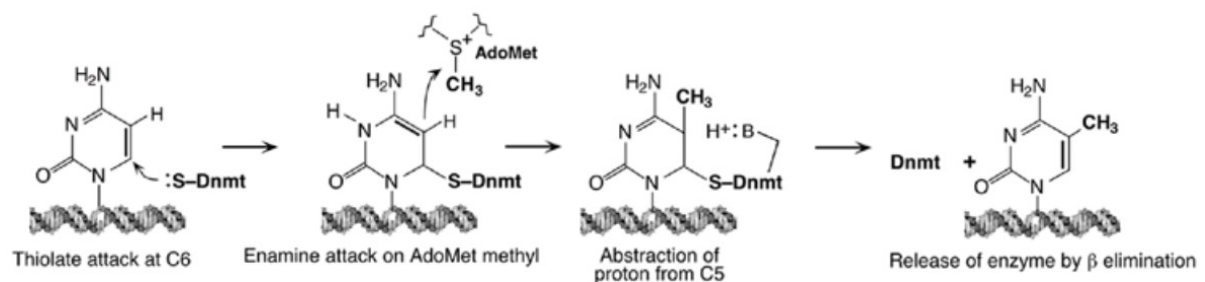
deux allèles subit une méthylation de ses îlots CpG. On parle alors de régions à méthylation différentielle ou DMR pour *Differentially Methylated Regions*. On distingue les DMR primaires et les DMR secondaires. Les premières sont mises en place dans les lignées germinales par méthylation *de novo* et résistent à la vague de déméthylation qui a lieu après la fécondation pendant la phase pré-implantatoire (cf Section « Le cycle de l’empreinte », p.36). De nombreuses DMR primaires sont des régions de contrôle de l’empreinte d’un cluster de GSE (*Imprinting Control Region* ou ICR). Leur délétion perturbe l’empreinte de tout le cluster dans la région (Bartolomei, 2009; Fitzpatrick et al., 2002; Lin et al., 2003; Thorvaldsen et al., 1998; Williamson et al., 2006; Wutz et al., 1997; Yang et al., 1998). Les DMR secondaires sont, quant à elles, mises en place dans les cellules somatiques après la vague de déméthylation massive qui a lieu dans l’embryon, et résultent des mécanismes de répression induits par les ICR dans la région.

Les DMR sont retrouvées au niveau des régions promotrices de gènes mais également dans les régions intergéniques notamment dans le cas des ICR. L’allèle méthylé est généralement l’allèle réprimé mais il existe des situations inverses. Par exemple, dans la région *IGF2/H19* la méthylation de l’allèle paternel réprime l’expression de *H19pat* mais autorise l’expression d’*IGF2pat*. Parmi les DMR identifiées à ce jour, une grande majorité est caractérisée par une méthylation de l’allèle maternel, signe peut être que la méthylation *de novo* est plus active dans l’ovocyte que dans le spermatozoïde.

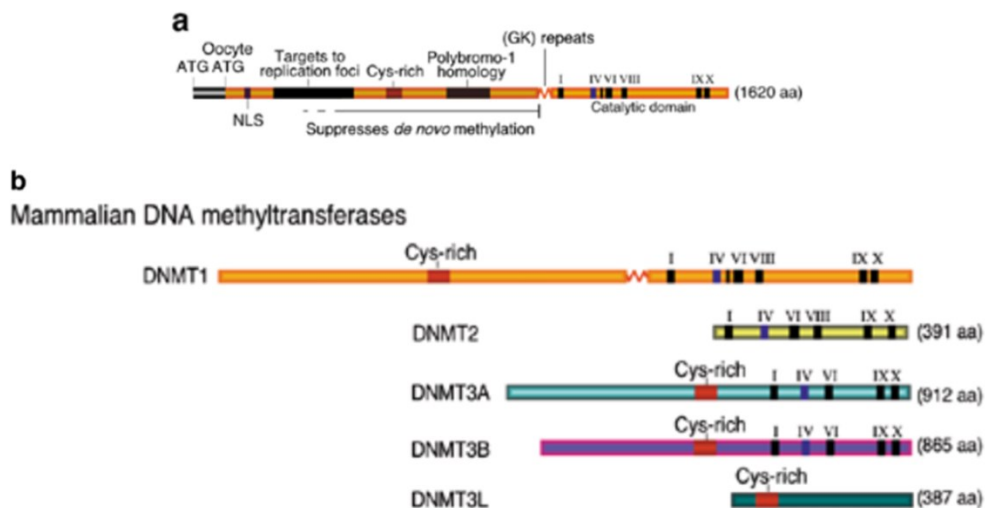
Les DMR peuvent également jouer un rôle dans l’empreinte tissu-spécifique ou spécifique d’un stade du développement. Dans le cas de *GNAS*, la régulation de l’empreinte est toute particulière et complexe. La transcription de ce GSE est sous le contrôle de 4 promoteurs alternatifs donnant naissance à différents transcrits à expression maternelle, paternelle ou biallélique, ainsi qu’à un transcrit antisens non-codant. Trois DMR ont été identifiées dans le cluster *GNAS* dont deux sont germinales. Des expériences de délétions ciblées au sein de ces régions ont pu démontrer que les deux DMR germinales sont des ICR. L’une d’entre elles contrôle l’empreinte tissu-spécifique d’un seul transcrit et interagit avec la deuxième région de contrôle principale qui régule l’expression de tous les transcrits y compris l’ARN non-codant qui pourrait avoir un rôle majeur dans le mécanisme de l’empreinte dans ce locus (Peters and Williamson, 2008).

### 2.1.3. Enzymes de méthylation de l'ADN

Le rajout du groupement Méthyl en position 5 de la Cytosine est une réaction enzymatique médiée par les 5-méthyl-transférases dites DNMT. Ce groupement Méthyl- est apporté par un donneur de Méthyl- qui est la S-Adénosyl-Méthionine (Figure 6). Ces enzymes vont cibler les motifs CpG. Chez les mammifères, on trouve les méthyl-transférases *de novo* (DNMT3A et DNMT3B), leur cofacteur DNMT3L et la méthyl-transférase de maintenance DNMT1 (Figure 7).



**Figure 6 :** Les différentes réactions chimiques aboutissant à la méthylation de la Cytosine. La Dnmt se lie au Carbone 6 par son résidu thiolate. Ensuite, le groupement méthyl est apporté par un donneur de méthyl, la S-Adénosyl-Méthionine qui réagit avec le Carbone 5. Le proton en C5 est capté par la Dnmt qui est ensuite libérée (d'après Bestor et al., Hum Mol Genet 2000 (Bestor, 2000)).



**Figure 7 :** Représentation graphique des différents domaines des DNMT. (a) exemple de la DNMT1. Le domaine N terminal (extrémité gauche) contient des éléments qui guident l'action de l'enzyme : la localisation nucléaire (NLS), le ciblage des lieux de réplication et la suppression partielle de l'activité de méthylation de novo. Le domaine C terminal (extrémité droite) comporte 10 motifs participant à l'activité catalytique de l'enzyme (b) comparaison des structures linéaires des différentes DNMT des mammifères. Contrairement à la DNMT1, les DNMT3A et DNMT3B ont perdu une partie du domaine N terminal. DNMT3L ne possède pas le domaine C terminal catalytique, d'où l'absence d'activité enzymatique propre (d'après Bestor et al., Hum Mol Genet 2000 (Bestor, 2000)).

### 2.1.3.a. DNMT3A et DNMT3B

Ces enzymes sont à l'origine de la méthylation *de novo* de l'ADN. Okano et al ont montré que ces deux méthyl-transférases sont fortement exprimées dans les cellules ES et qu'elles sont capables de méthyler à la fois les séquences hémi-méthylées et les séquences non-méthylées, *in vitro*, avec la même efficacité (Okano et al., 1998). Ce qui était en faveur d'une activité de méthylation *de novo*. Cette même équipe a pu ensuite démontrer cette hypothèse grâce à des expériences sur des souris transgéniques déficientes en Dnmt3a et/ou Dnmt3b (Okano et al., 1999). Dans ce travail, les auteurs ont prouvé (1) que l'inactivation de ces deux enzymes empêche la méthylation *de novo* dans les cellules ES et dans l'embryon murin au stade précoce avec une létalité embryonnaire précoce, (2) que cette inactivation n'affecte pas la maintenance de la méthylation et (3) que ces deux enzymes peuvent avoir des actions différentes plus spécifiques, la Dnmt3b étant spécifiquement nécessaire à la méthylation de certaines régions centromériques. Cette dernière observation est compatible avec le phénotype associé aux mutations homozygotes et hétérozygotes composites de *DNMT3B* qui correspond au syndrome ICF (Immunodéficience, instabilité Centromérique et dysmorphie Faciale). Ce syndrome est associé à un défaut de méthylation des satellites II et III au niveau de l'hétérochromatine constitutive des régions juxtacentromériques des chromosomes 1, 9 et 16. Par ailleurs, dans les travaux de Okano et al., l'état de méthylation de certains GSE (*Igf2* et *H19*) dans les souris *Dnmt3a*<sup>-/-</sup> et *Dnmt3b*<sup>-/-</sup> avait également été étudié et a montré une déméthylation des DMR étudiées en l'absence des enzymes Dnmt3. L'étude a conclu au rôle majeur de Dnmt3a et Dnmt3b dans la méthylation *de novo* et le développement embryonnaire.

Il existe différentes isoformes pour DNMT3A, parmi lesquelles DNMT3A2 qui est une petite isoforme dépourvue des 5 premiers exons (N-terminal) mais ayant conservé l'activité méthyl-transférase, avec une répartition subcellulaire et tissulaire spécifique. DNMT3A2 est préférentiellement localisée au niveau de la chromatine active, euchromatine, et est l'isoforme prédominante dans les cellules souches embryonnaires, les cellules tumorales d'origine embryonnaire, le testicule, les ovaires, le thymus et la rate (Chen et al., 2002).

Même si ces deux enzymes interviennent essentiellement dans la méthylation *de novo*, certaines données sont en faveur de leur participation dans le maintien de la méthylation également (Jones and Liang, 2009).

#### 2.1.3.b. DNMT3L

Cette protéine est un cofacteur des méthyl-transférases *de novo* dénué d'activité méthyl-transférase propre. Elle possède des homologies de séquence avec DNMT3A et DNMT3B, d'où son nom (*DNMT3-Like protein*) mais elle est dépourvue des domaines catalytiques à l'origine de l'activité enzymatique ainsi que des domaines de reconnaissance et de liaison à la séquence nucléotidique (Figure 7).

Les souris transgéniques déficientes en Dnmt3l montrent un défaut de méthylation spécifique des régions soumises à l'empreinte, alors que le niveau de méthylation du reste du génome n'est pas impacté (Bourc'his et al., 2001). Dnmt3l serait donc impliquée dans l'établissement de l'empreinte via la méthylation pendant la gamétogenèse.

Les études protéiques ont montré que DNMT3L interagit avec DNMT3A2, DNMT3B et les histones du corps nucléosomal notamment le domaine N-terminal de l'histone H3 à condition qu'elle ne soit pas méthylée en position Lysine-4. Cette dernière modification de l'histone est la seule modification qui inhibe l'interaction (Ooi et al., 2007). Par ailleurs, Jia et al. ont montré par des études de cristallographie que le domaine C-terminal de Dnmt3l interagit avec le domaine catalytique de Dnmt3a (Jia et al., 2007).

L'expression de Dnmt3l est régulée, entre autres, par Dnmt3a et Dnmt3b. La transcription de Dnmt3l est réprimée dans les souris *Dnmt3a*<sup>-/-</sup> et *Dnmt3b*<sup>-/-</sup> et son promoteur est hypométhylé (Hu et al., 2008).

#### 2.1.3.c. DNMT1

C'est la première méthyl-transférase qui a été découverte et clonée. Des expériences *in vitro* ont démontré que cette enzyme possède à la fois une activité méthyltransférase *de novo* sur des séquences non méthylées et de maintenance sur des séquences hémi-méthylées avec une préférence 5 à 30 fois supérieure pour les séquences hémi-méthylées, et que son activité de maintenance était plus efficace que son activité *de novo* (Bestor, 1992). En outre, DNMT1 a été localisée au niveau des fourches de réplication suggérant un rôle de

méthylation au moment de la réplication dans le but d'assurer la transmission de cette marque épigénétique au cours des divisions cellulaires (Leonhardt et al., 1992). Des études protéiques ont mis en évidence une interaction entre certains domaines de DNMT1 et le facteur PCNA (Proliferating Cell Nuclear Factor) qui fait partie du complexe de réplication ou réplisome, et interagit avec l'ADN polymérase (Chuang et al., 1997). DNMT1 est essentiellement transcrite en phase S (réplicative) du cycle cellulaire. Enfin, des expériences sur des souris transgéniques déficientes en *Dnmt1* ont montré que l'absence de *Dnmt1* entraînait une perte globale de la méthylation et une expression biallélique de certains GSE (Lei et al., 1996; Li et al., 1992).

Concernant une éventuelle activité de méthylation *de novo* de DNMT1, aujourd'hui il n'y pas d'arguments en faveur de sa capacité à initier la méthylation chez les mammifères au cours du développement.

Il faut noter que la forme classique de la protéine *Dnmt1* qui est largement présente dans les noyaux des cellules somatiques est absente des cellules germinales et de l'embryon préimplantatoire. Une forme tronquée ayant toujours l'activité méthyl-transférase, apparaît spécifiquement dans l'ovocyte (*Dnmt1o*) (Bestor, 2000).

Cette séparation nette des actions des ADN méthyl-transférases *de novo* et de maintenance dans le temps est aujourd'hui discutée. Dans une revue de Jones et Liang (Jones and Liang, 2009), les auteurs présentent les arguments permettant de suggérer un nouveau modèle de fonctionnement de la méthylation de l'ADN. Ils constatent les lacunes dans l'ancien modèle et présentent les données scientifiques en faveur du nouveau modèle. En effet, l'un des points essentiels restant sans explication est le manque d'un système de réparation du maintien de la méthylation. Ils postulent que DNMT1 ne peut être la seule enzyme assurant le maintien d'une marque épigénétique aussi importante pour le développement. Il semble très peu probable qu'un tel système soit dépourvu d'un mécanisme de réparation des erreurs possibles de la DNMT1 dans le maintien de la méthylation. En outre, chez les souris déficientes en *Dnmt3a* et *Dnmt3b* une partie des CpG au niveau des séquences répétées se retrouve hémi-méthylée. De même, dans les cellules souches embryonnaires DNMT3A<sup>-/-</sup> et DNMT3B<sup>-/-</sup> il a été observé une perte progressive de la méthylation au fur et à mesure des divisions cellulaires. Ces données soutiennent l'hypothèse que DNMT1 ne suffit pas à elle

seule au maintien de la méthylation. DNMT3A et DNMT3B pourraient également intervenir. Une autre donnée allant également dans ce sens est le fait que même si le niveau d'expression de DNMT3A et DNMT3B diminue dans les cellules différenciées par rapport aux cellules souches embryonnaires, elles continuent néanmoins à être exprimées. Par ailleurs, une équipe a mené une étude visant à déterminer la fidélité de maintenance de la méthylation au cours de la division cellulaire. Ceci a été réalisé au niveau d'un îlot CpG hyperméthylé du gène *FMR1* à partir de lymphocytes humains. Les résultats sont en faveur de la survenue d'un mécanisme *de novo* à minima en plus du mécanisme de maintenance (Laird et al., 2004). Enfin, une étude récente de Wu et al. a montré que Dnmt3a est exprimée dans les cellules souches neuronales postnatales où elle méthyle des régions intergéniques et intragéniques en dehors des régions promotrices entraînant une régulation de l'expression de certains gènes importants pour le développement (Wu et al., 2010).

#### 2.1.4. Signification biologique de la méthylation de l'ADN

Globalement, la méthylation de l'ADN induit une répression de la transcription. Elle est ainsi impliquée dans le maintien de l'intégrité globale du génome en réprimant les séquences répétées (ex : rétrotransposons, séquences *Long Interspersed Nuclear Elements* ou LINE) et en les stabilisant (ex : séquences satellites centromériques). Dans un travail de Bourc'his et Bestor (Bourc'his and Bestor, 2004), les auteurs montrent que les souris *Dnmt3l*<sup>-/-</sup> ont une perte de la méthylation *de novo* des séquences répétées type LTR (Long Terminal Repeat) et rétrotransposons, dans les cellules germinales mâles où *Dnmt3l* est normalement exprimée, entraînant leur surexpression dans ce type cellulaire. Dans une autre étude, Webster et al. (Webster et al., 2005) montrent que la délétion de *Dnmt3l* a également pour conséquence un défaut de compaction de la chromatine dans les cellules germinales mâles en méiose avec également un défaut d'appariement des chromosomes homologues entraînant un arrêt méiotique. Ces résultats démontrent le lien entre la méthylation et la conformation chromatinienne. Concernant son rôle dans la stabilisation des régions centromériques, les données sur les patients atteints du syndrome ICF qui ont une mutation de *DNMT3B* et un défaut de méthylation des satellites II et III péri-centromériques des chromosomes 1, 9 et 16 démontrent le rôle de la méthylation dans l'organisation de l'hétérochromatine péri-centromérique.



Même si la fonction de répression transcriptionnelle est la plus communément observée pour la méthylation de l'ADN, il est aujourd'hui admis que cette marque épigénétique peut avoir d'autres fonctions selon sa localisation dans le génome. En effet, lorsque la méthylation a lieu au niveau du site d'initiation de la transcription cela induit une inhibition de l'expression génique en bloquant l'initiation et non l'élongation. En revanche, la méthylation au sein du corps du gène n'éteint pas la transcription mais pourrait même la stimuler et peut-être jouer un rôle dans l'épissage (Jones and Liang, 2009). Ceci est également illustré par les résultats de l'étude de Wu et al. dans laquelle les auteurs montrent que la méthylation de sites en dehors de la région promotrice active la transcription des gènes en contrant l'effet de protéines répressives qui sont les protéines du complexe Polycomb. La méthylation maintiendrait dans ce cas la chromatine dans un état ouvert transcriptionnellement actif (Wu et al., 2010).

D'un point de vue mécanistique, les 5mC vont promouvoir ou, au contraire, empêcher la fixation d'éléments régulateurs de la transcription. Dans le cas de la répression, deux cas de figures ont été observés : (1) La méthylation des CpG empêche la reconnaissance de séquences particulières par des protéines capables de promouvoir la transcription telles que CTCF dans le cas de l'empreinte du locus *Igf2/H19* par exemple. (2) La 5mC est reconnue par des protéines qui se lient aux CpG méthylés via un domaine dit MBD (Methyl-CpG-Binding Domain). Ces protéines vont ensuite induire une répression de la transcription par interaction avec des histones désacétylases (Bird, 2002). La méthylation de l'ADN interagit également avec les modificateurs de la chromatine. L'analyse des modifications d'histones au niveau des ICR dans un contexte de déficience en *Dnmt3l* et donc d'absence de méthylation des Cytosines de ces régions génomiques, montre une diminution des marques répressives des histones (H3K9me3, H4K20me3 et H2A/H4R3me2) (Henckel et al., 2009). Un lien existe entre la méthylation des Cytosines et la méthylation des histones. Ces deux marques épigénétiques semblent dépendre l'une de l'autre via une interaction directe entre les méthyl-transférases des histones et celles des Cytosines (Cedar and Bergman, 2009). Une autre interaction directe existe entre les Cytosine-méthyltransférases et les protéines répressives du groupe Polycomb. La protéine EZH2 des complexes PRC2/3 recrute les différentes méthyl-transférases (1, 3A et 3B) (Vire et al., 2006).

La méthylation de l'ADN est indispensable à certains processus du développement tels que l'inactivation de l'X et l'empreinte. Elle est également à la base de la différenciation cellulaire. La perte du caractère totipotent d'une cellule est médiée par la méthylation d'un certain nombre de ses gènes. Chaque type cellulaire aura ainsi son profil de méthylation qui lui sera bien spécifique. Dans d'autres situations, néanmoins, elle joue un rôle pathologique notamment dans la physiopathologie cancéreuse en réprimant l'expression de gènes suppresseurs de tumeurs, ou à l'inverse la déméthylation de certains gènes du développement ce qui conduit à la réactivation de leur expression et à la transformation cellulaire (Feinberg and Tycko, 2004).

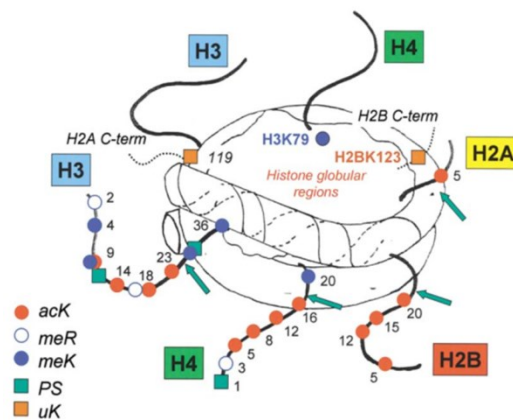
Dans le cas de l'empreinte, la méthylation joue un rôle essentiel en particulier la méthylation différentielle. Cette dernière est à l'origine de l'expression mono-allélique en induisant des mécanismes répressifs ou activateurs de l'expression de l'allèle méthylé. Le défaut de méthylation se traduit par un défaut d'expression des GSE et une perte de l'empreinte (Li et al., 1993).

## **2.2. Modifications d'histones**

### *2.2.1. Les marques des histones*

Les histones sont des protéines associées à l'ADN qui permettent la compaction de l'ADN et l'organisation tridimensionnelle de la chromatine. La chromatine est formée par la répétition d'une unité structurale qui est le nucléosome. Ce dernier est composé d'un octamère d'histones sur lequel s'enroulent 147 pb d'ADN. Les nucléosomes sont reliés entre eux par un fragment d'ADN de 20 – 60 pb dit « Linker ». L'octamère d'histones est composé de 4 histones en deux exemplaires chacune, H2A, H2B, H3 et H4. Les extrémités N-terminales des histones ou queues d'histones sont exposées à la face extérieure du nucléosome les rendant accessibles à plus d'une centaine de modifications enzymatiques post-traductionnelles. Ces modifications peuvent survenir sur différents résidus des différents composants de l'octamère et impliquer différentes enzymes (Figure 8). Chaque modification est associée à un état de compaction particulier de la chromatine et constitue, à elle seule ou en combinaison avec d'autres modifications enzymatiques, ce qu'on appelle le code des histones. Ce dernier est « lu » par des protéines qui ne sont pas des histones et se traduit soit par un état chromatinien « ouvert » transcriptionnellement actif soit par un état

« fermé » transcriptionnellement silencieux (Jenuwein and Allis, 2001; Richards and Elgin, 2002; Spotswood and Turner, 2002; Turner, 2002).



**Figure 8 :** Représentation graphique des différentes modifications post-traductionnelles des histones constituant le nucléosome, essentiellement dans les extrémités N terminales et rarement dans les extrémités C terminales (Turner, 2002).

Les modifications enzymatiques les plus connues sont l'acétylation, la méthylation, la phosphorylation et l'ubiquitination (Kouzarides, 2007). Récemment, de nouvelles marques d'acylation des Lysines des histones ont été découvertes, à savoir la propionylation ( $K_{Pr}$ ), la butyrylation ( $K_{Bu}$ ), la crotonylation ( $K_{Cr}$ ), la 2-hydroxyisobutyrylation ( $K_{Hib}$ ), la succinylation ( $K_{Su}$ ), la malonylation ( $K_{Ma}$ ) et la glutarylation ( $K_{Glu}$ ) (Dai et al., 2014; Rousseaux and Khochbin, 2015; Tan et al., 2011; Tan et al., 2014; Xie et al., 2012). Ces modifications sont mises en place par différentes enzymes spécifiques : des Histone-Methyl-Transférases (HMT), des Histone-Acetyl-Transférases (HAT), des kinases, ubiquitylases... Ces marques épigénétiques sont dynamiques et réversibles par une série d'enzymes telles que les Histone-Déméthylases, les phosphatases, les désubiquitylases et les Histone-DéACétylases (HDAC).

### 2.2.2. Héritabilité des marques d'histones

Pour répondre à la définition du terme « Epigénétique », les modifications des histones doivent être transmissibles lors de la réplication de l'ADN. Les mécanismes conférant une mémoire cellulaire des modifications des histones et donc leur conservation à travers les

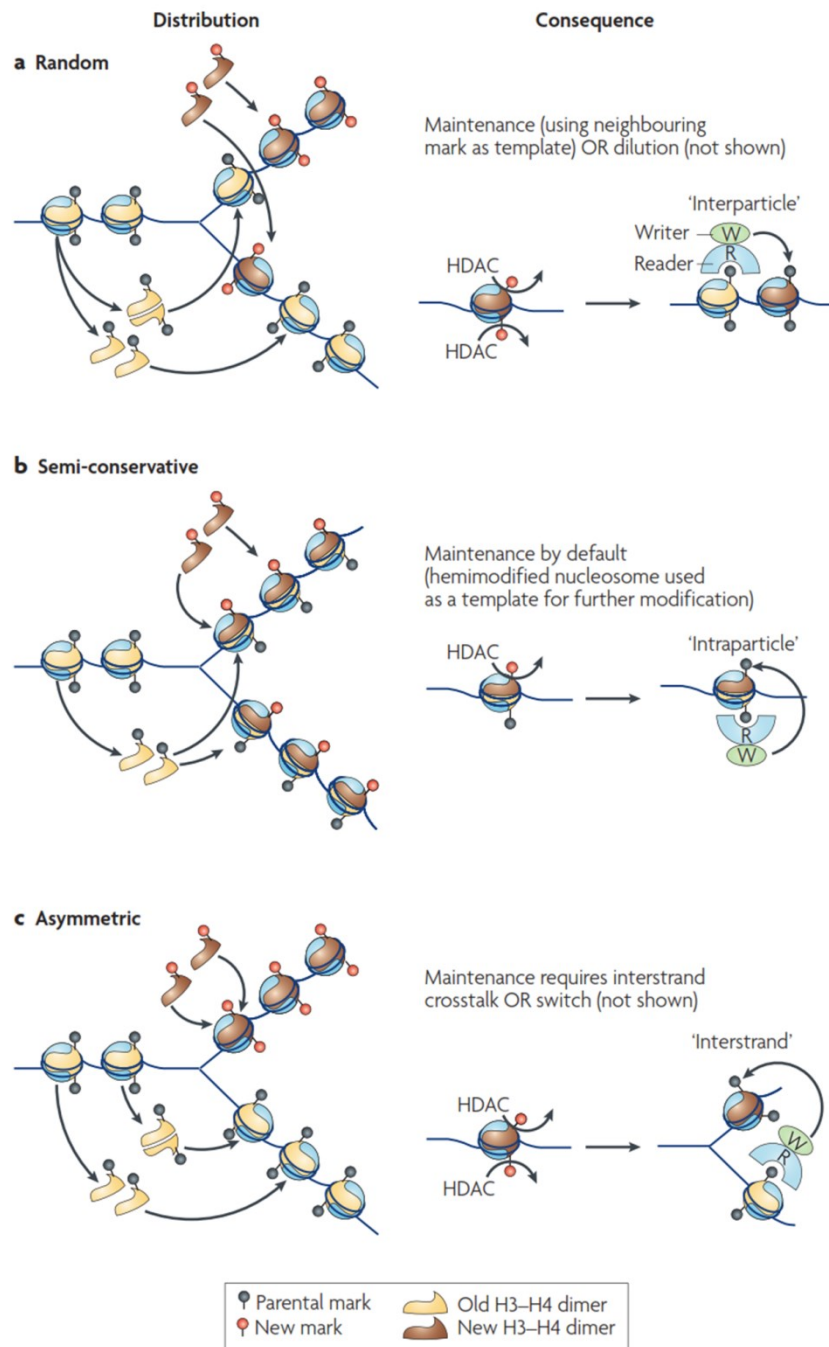
divisions cellulaires sont mal connus à ce jour. L'héritabilité des marques des histones lors des divisions cellulaires a été observée chez les microorganismes unicellulaires. Elle passerait par des interactions entre ces marques et les protéines qu'elles recrutent. Ces dernières assureraient à leur tour le recrutement des enzymes de modification des histones pour transmettre les marques à la chromatine naissante (Grewal and Moazed, 2003). Chez les mammifères, les marques d'histones pourraient être transmises grâce à la répartition aléatoire des histones entre les brins nouvellement synthétisés puis ces marques se propageraient dans les histones nouvellement incorporées par un mécanisme d'auto-entretien de la marque enzymatique via les différentes interactions protéiques qu'elle induit (Bernstein et al., 2007; Groth et al., 2007; Tagami et al., 2004).

La transmission de la méthylation de la lysine H3K9 au niveau de l'hétérochromatine serait médiée par la protéine HP1 alors que la transmission des marques H3K4me et H3K27me se ferait au travers des complexes Polycomb et Trithorax dont certains composants possèdent des activités enzymatiques modificatrices des histones (exemple : E(Z) qui a une activité histone-méthyl-transférase) (Bernstein et al., 2007; Ringrose and Paro, 2004; Schuettengruber et al., 2007). En revanche, il n'y a pas de données en faveur de la transmission d'autres modifications. Des mécanismes médiés par des ARN non codants ont été proposés (Kouzarides, 2007).

La vitesse d'échange très court des histones H2A et H2B n'est pas compatible avec une conservation de leurs modifications post-traductionnelles en vue de leur transmission, contrairement aux histones H3 et H4. Ces dernières sont, à l'inverse, ré-intégrées dans les brins nouvellement synthétisés par des chaperonnes d'histones (Anti-Silencing Function 1, ASF1) constituant ainsi un modèle pour reproduire les mêmes modifications sur les histones nouvellement incorporées par d'autres chaperonnes (Chromatin Assembly Factor 1, CAF1) (De Koning et al., 2007; English et al., 2006; Natsume et al., 2007; Probst et al., 2009; Verreault et al., 1996).

Trois modèles de transmission ont été proposés (Figure 9) : (1) le modèle de distribution « aléatoire » des tétramères H3-H4 du double brin initial entre les deux simples brins à répliquer. Dans ce cas, les marques des histones sont maintenues grâce au recrutement de

protéines par les marques existantes. Ces mêmes protéines qui reconnaissent ces marques peuvent soit posséder une activité enzymatique modificatrice des histones comme c'est le cas du complexe Polycomb répressif (PRC2) qui reconnaît et instaure la méthylation de H3K27, soit recruter à leur tour des protéines modificatrices des histones pour retranscrire les marques sur les histones avoisinantes. Ce dernier modèle semble être compatible avec la transmission de l'hétérochromatine de grandes régions répétées telles que les régions péricentromériques où la protéine HP1 aiderait à la propagation des marques répressives des histones. (2) le modèle de distribution « semi-conservatrice », selon lequel chaque dimère H3-H4 est attribué à l'un des simples brins. Ceci résulte en une distribution symétrique de ces histones et donc des modifications qu'elles portent. La cellule n'aurait donc plus qu'à recopier la modification initiale sur le nouveau dimère incorporé au sein du nucléosome. (3) le modèle de distribution « asymétrique », où les tétramères H3-H4 seraient ré-intégrés unilatéralement sur un seul des deux brins en réplication. Dans ce modèle, une communication entre les deux brins est nécessaire et pourrait se produire dans les trois plans de l'espace afin de retranscrire les marques d'histones d'origine. Ces trois modèles pourraient coexister et être mis en place dans des conditions spatio-temporelles spécifiques à chacun (Probst et al., 2009). Les modifications des histones peuvent également être mises en place de façon concomitante à la réplication car des protéines modificatrices des histones sont greffées au complexe PCNA.



**Figure 9** : Représentation graphique des trois hypothèses de transmission des marques d'histones H3 et H4 pendant la réplication (Probst et al., 2009).

### 2.2.3. Signification biologique des modifications des histones

La présence de résidus modifiés aux queues des histones va impacter l'état de compaction de la chromatine soit en modifiant les interactions physiques entre l'ADN et les histones et entre les nucléosomes, soit en recrutant ou en refoulant des protéines, dont certaines

possèdent une activité enzymatique telle que la méthylation de l'ADN, qui participent au remodelage et à la modification de la chromatine (Kouzarides, 2007). L'hypothèse des interactions physiques n'a pas encore pu être démontrée *in vivo* mais des expériences *in vitro* ont montré que l'acétylation de la Lysine H4K16 perturbe la formation de la structure chromatinienne (Kouzarides, 2007; Shogren-Knaak and Peterson, 2006). Ceci serait lié à la modification de la charge positive de la Lysine qui se retrouve neutralisée par la charge négative du groupement Acétyl. La phosphorylation pourrait également exercer le même effet. D'un autre côté, l'impact de chaque modification des histones sur les interactions chromatiniennes reste encore mal connu sauf pour les plus étudiées d'entre elles à savoir la méthylation et l'acétylation. Cette dernière est reconnue par des protéines à bromodomaine et est quasi toujours associée à une répression transcriptionnelle. Quant à la méthylation, elle peut être reconnue par des protéines à chromodomaine ou à domaine PHD (Plant Homeo-Domain) et est corrélée soit à une activation soit à une répression de la transcription en fonction de sa localisation. La méthylation en H3K4 et H3K36 est associée à une chromatine transcriptionnellement active alors que la méthylation des Lysines H3K9, H3K27 et H4K20 sont plutôt des marques répressives (Bernstein et al., 2007). Ces marques épigénétiques vont participer au recrutement de facteurs pouvant modifier la structure chromatinienne. Les Lysines acétylées ainsi que les marques actives de méthylation des histones (H3K4me et H3K36me) vont fixer des protéines à Bromo-domaines qui à leur tour recrutent des facteurs activateurs (Pray-Grant et al., 2005; Sims et al., 2005), alors que les méthylation H3K9 et H3K27 répressives fixent l'Heterochromatin Protein 1 (HP1) et le complexe Polycomb qui vont induire la compaction de la chromatine (Margueron et al., 2005).

Il existe une coordination entre les différentes modifications des histones. Elles peuvent s'influencer mutuellement négativement ou positivement. Ainsi, lorsque les fonctions des deux modifications d'histones sont opposées l'une d'entre elles peut empêcher la fixation des protéines normalement recrutées par la deuxième. A l'inverse, lorsque deux modifications induisent le même effet sur la structure chromatinienne elles peuvent agir en synergie pour renforcer la fixation de protéines recrutées. L'ensemble des modifications va concourir à créer un environnement chromatinien adéquat selon la séquence nucléotidique sous-jacente, le stade du développement, le timing du cycle cellulaire, etc.

D'un point de vue fonctionnel, les modifications des histones vont ainsi permettre la structuration de la chromatine d'un côté et organiser l'état de la chromatine en fonction des besoins cellulaires du moment. La première fonction de structuration se traduit par l'existence de l'hétérochromatine compacte, essentiellement au niveau des régions péri-centromériques, et de l'euchromatine relâchée. L'état hétérochromatinien est associé à une diminution de l'acétylation des histones et un taux élevé de H3K9, H3K27 et H3K20 méthylées. A l'inverse, l'euchromatine est associée à des niveaux élevés d'histones acétylées et aux marques H3K4, H3K36 et H3K79 triméthylées. Il faut néanmoins noter que dans la cellule vivante le système est plus complexe avec parfois la colocalisation des marques répressives et activatrices notamment au niveau de gènes du développement (Bernstein et al., 2005; Kouzarides, 2007). Au sein de l'euchromatine, les modifications des histones vont participer (1) à la régulation de l'expression des gènes où chaque modification sera associée à un état transcriptionnel actif ou réprimé (Tableau 1), (2) à la réparation de l'ADN et (3) à sa réplication.

**Tableau 1** : Liste des principales modifications post-traductionnelles des histones avec les principales enzymes qui en sont responsables et les fonctions connues.

Modification enzymatique	Enzymes modificatrices	Acides aminés modifiés	Fonction
Acétylation	Acétyl-transférases (GNAT, MYT et CBP/p300)	Lysine	Transcription (activation) Réparation et réplication de l'ADN Condensation des chromosomes
Désacétylation	Désacétylases (Classes I, II et III (Sir))	Lysine	Transcription (répression)
Méthylation	Lysine-Méthyl-transférases (SET, MLL, NSD1, EZH2, SUV39H1/2, etc)	Lysine	Transcription (activation ou répression en fonction de la localisation des résidus méthylés) Réparation de l'ADN
	Arginine-Méthyl-transférases (PRMT, CARM)	Arginine	Transcription (activation ou répression)
Déméthylation	Déméthylases (LSD1, JHDM, JMJD...)	Lysine	Transcription (activation ou répression). Fonctions encore mal connues
Phosphorylation	MSK1/2, RSK2...	Sérine et Thréonine	Transcription (fonctions encore mal connues), Réparation de l'ADN Condensation des chromosomes



**Tableau 1 (suite) : Liste des principales modifications post-traductionnelles des histones avec les principales enzymes qui en sont responsables et les fonctions connues.**

Modification enzymatique	Enzymes modificatrices	Acides aminés modifiés	Fonction
Ubiquitylation	Bmi/Ring1A, RNF20/RNF40	Lysine	Transcription (activation. Fonctions encore mal connues) Réparation de l'ADN
Désubiquitylation	Ubp	Lysine	Transcription (activation ou répression).
Sumoylation		Lysine	Transcription (répression).

Par ailleurs, il existe une communication entre les marques des histones et la méthylation des Cytosines. Il a été observé que la répression de certains gènes ne passe que par la méthylation des histones et jamais par la méthylation de l'ADN. Ainsi, des îlots CpG au niveau de certaines régions promotrices ne sont jamais méthylés au niveau des Cytosines mais sont marqués par la méthylation H3K27me3. De même, la méthylation H3K4 est présente uniquement en absence de méthylation de l'ADN. Cette marque d'histone inhiberait la méthylation *de novo* des Cytosines (Ooi et al., 2007).

Dans un contexte de déficience en *Dnmt3l* et donc d'absence de méthylation des Cytosines des ICR, les marques répressives des histones (H3K9me3, H4K20me3 et H2A/H4R3me2) sont diminuées (Henckel et al., 2009). Ces deux marques épigénétiques, la méthylation des Cytosines et la méthylation des histones, semblent dépendre l'une de l'autre via une interaction directe entre les méthyl-transférases des histones et celles des Cytosines (Cedar and Bergman, 2009). Une autre interaction directe existe entre les Cytosine-méthyl-transférases et les protéines répressives du groupe Polycomb. La protéine EZH2 des complexes PRC2/3 recrute les différentes méthyl-transférases (1, 3A et 3B) (Vire et al., 2006). Or, certaines marques de méthylation des histones (H3K9 et H3K27) recrutent des protéines du complexe Polycomb qui peut donc à son tour recruter les DNMT.

Les marques des histones seraient plus dynamiques que la méthylation de l'ADN. Elles assureraient une régulation plus flexible alors que la méthylation de l'ADN serait plus stable dans le temps (Lee et al., 2006; Nelissen et al., 2011; Reik, 2007).

## 2.3. ARN non codants

### 2.3.1. Définitions et fonctions

La grande majorité des transcrits chez les mammifères ne codent pas pour des protéines. Ces transcrits sont dits ARN non codants (ARNnc) (Carninci et al., 2005; Katayama et al., 2005). Il existe différentes classes d'ARNnc en fonction de leur taille et leur fonction. On distingue les ARNnc longs qui font quelques centaines à plusieurs centaines de milliers de nucléotides, et les petits ARNnc dont la taille ne dépasse pas quelques centaines de nucléotides. Dans ces derniers, on différencie les siRNA (short interfering RNA) de 21 nucléotides, les miRNA (micro RNA) d'environ 22 nucléotides, les piRNA (piwi interacting RNA) de 26 à 31 nucléotides et les snoRNA (short nucleolar RNA) qui comportent 60 à 300 nucléotides.

Ces ARNnc ont des fonctions diverses de maintien de l'intégrité génomique. Ils interviennent également dans l'architecture nucléaire et la régulation de l'expression des gènes, et ce dans des processus cellulaires importants tels que l'empreinte génomique, l'inactivation de l'X et la différenciation cellulaire.

Globalement, les ARNnc longs exercent leurs fonctions en *cis* c'est-à-dire sur les séquences environnantes du même chromosome mais certains ont également des fonctions en *trans*. *Xist* et *Tsix* sont deux exemples d'ARNnc longs qui agissent en *cis* sur le chromosome X dont ils proviennent et qui sont à l'origine de l'inactivation fonctionnelle de l'X chez les femelles des mammifères. Les petits ARNnc, quant à eux, agissent en *trans* sur d'autres chromosomes que celui dont ils proviennent et peuvent également interagir avec les ARNs cytoplasmiques (Nelissen et al., 2011).

Les miRNA se lient à la région 3'UTR des ARNm pour empêcher leur traduction ou les guider vers les voies de dégradation. Les piARN sont spécifiques des cellules germinales où ils auraient une fonction de protection contre les transposons. Les snoRNA agissent comme des chaperonnes pour les ARN ribosomiaux.

Dans le cas de l'empreinte génomique, les régions soumises à l'empreinte contiennent en général une ou plusieurs séquences donnant des transcrits non codants qui peuvent être des ARNnc longs ou des petits ARNnc qui sont donc eux même soumis à l'empreinte. A ce jour, il

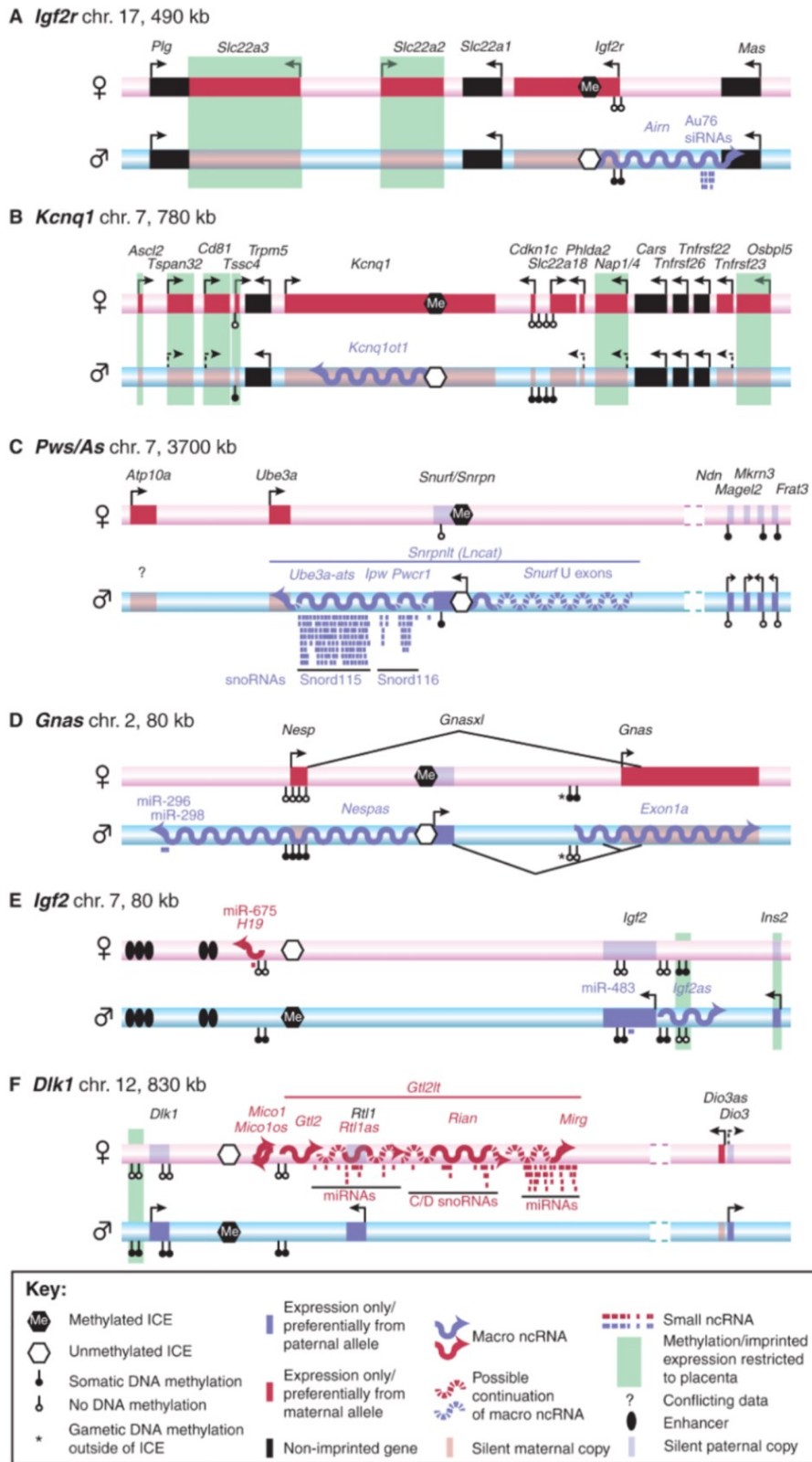
Il y a peu d'arguments en faveur d'une régulation de l'empreinte par les petits ARNnc. En revanche, les ARNnc longs peuvent participer au mécanisme de l'empreinte par une action en *cis* au niveau du cluster de GSE.

### 2.3.2. Les ARNnc longs (lncRNA)

Ces ARNnc contiennent plus de 200 nucléotides et résulteraient de deux phénomènes évolutifs : la duplication des gènes et l'insertion d'éléments génétiques mobiles. On peut distinguer différents types d'ARNnc longs mais leur classification n'est pas encore clairement établie. Globalement, il y a les lncRNA intergéniques qui comprennent les pseudogènes, et les lncRNA à l'intérieur de gènes codants (Khorkova et al., 2015).

Les lncRNA participent à divers processus cellulaires tels que la stabilité télomérique, l'établissement des boucles chromatinienne pour les interactions intra- et inter-chromosomiques qui régulent la transcription de certains clusters de gènes, la stabilité des ARNm, la régulation de la traduction ainsi que l'inactivation de l'X et l'empreinte.

Concernant les lncRNA appartenant aux GSE et participant à la régulation de l'empreinte, les plus connus sont *Airn* (108 kb) dans le locus *Igf2r*, *Kcnq1ot1* (91 kb) dans le locus *Kcnq1*, *Nespas* (31 kb) dans le locus *Gnas*, et *Lncat*, *H19* (2,7kb) et *Gtl2lt* dans les clusters *Snrpn*, *Igf2/H19* et *Dlk1* respectivement (Figure 10). Ces lncRNA possèdent des caractéristiques communes à savoir que : (1) ils s'expriment tous à partir de l'allèle à proximité de l'ICR hypométhylé qui se chevauche avec la région promotrice de certains d'entre eux (*Airn*, *Kcnq1ot1* et *Nespas*), et (2) ils répriment majoritairement les allèles des GSE en *cis* (Koerner et al., 2009).



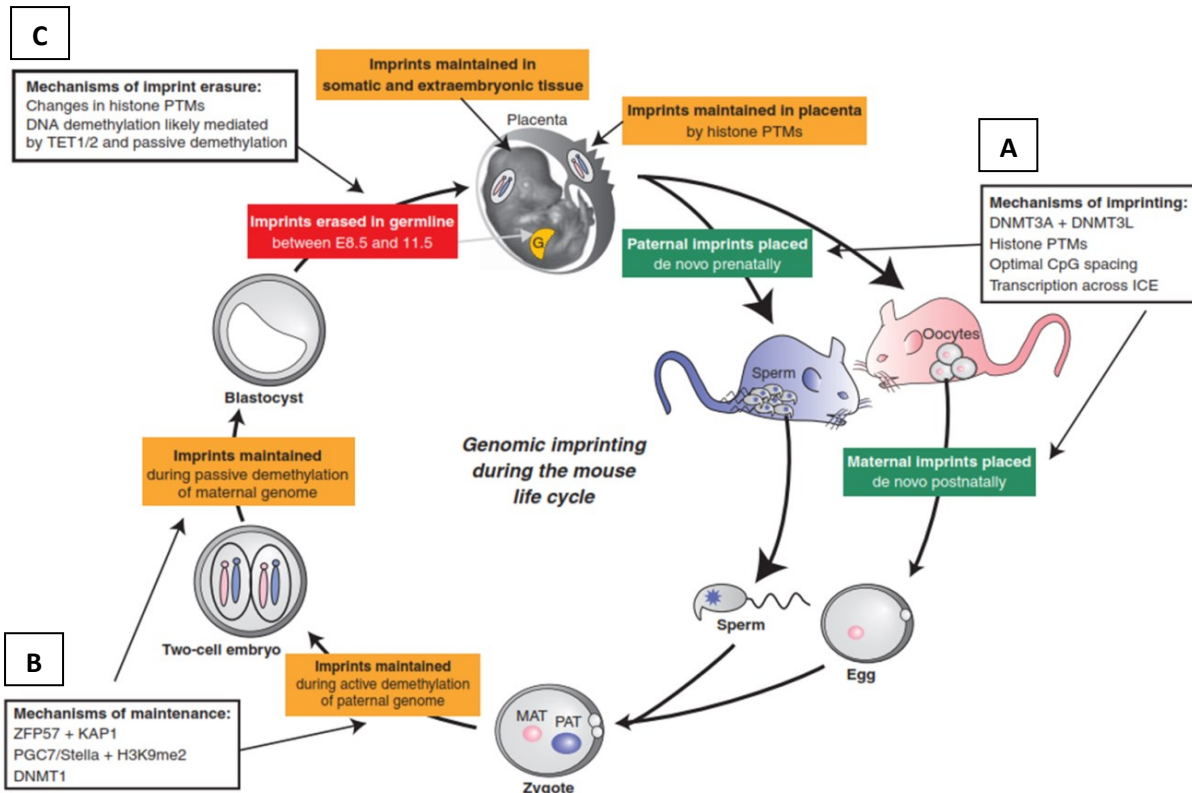
**Figure 10** : Exemples d'organisation génomique de clusters de GSE contenant des ARNnc. (Chr : Chromosome, ICE : Imprinting Control Element) (Koerner et al., 2009).

Le rôle des lncRNA en tant qu'élément régulateur de l'empreinte est soutenu par les résultats d'expériences de mutagenèse. En effet, des mutations introduites dans les ARNnc *Airn* et *Kcnq1ot1* conduisant à des ARNnc tronqués non fonctionnels, perturbent l'empreinte pour l'ensemble des gènes du cluster qui perdent leur expression mono-allélique (Mancini-Dinardo et al., 2006; Sleutels et al., 2002). Les lncRNA guideraient le dépôt des marques épigénétiques. Ils s'expriment en général à partir de l'allèle opposé à celui à partir duquel s'exprime le GSE codant pour une protéine. Ils recrutent des protéines du groupe Polycomb (PRC1 et PRC2) modifiant la chromatine et permettant le dépôt des marques répressives en *cis* pour assurer la répression des allèles des gènes voisins (Khorkova et al., 2015; Koerner et al., 2009; Pandey et al., 2008). Ces ARN pourraient également avoir un rôle dans le maintien même des marques épigénétiques de l'ICR puisque la délétion du promoteur *d'Airn*, contenu dans l'ICR hypométhylé, dans des cellules souches embryonnaires conduit à la méthylation de l'ICR (Stricker et al., 2008). Enfin, des questions se posent sur l'implication des lncRNA dans la spécificité tissulaire de l'empreinte car certains GSE ne sont exprimés de façon mono-allélique que lorsque l'ARNnc long du cluster est exprimé sinon l'expression du GSE est bi-allélique (Koerner et al., 2009). D'autres fonctions restent certainement à découvrir.

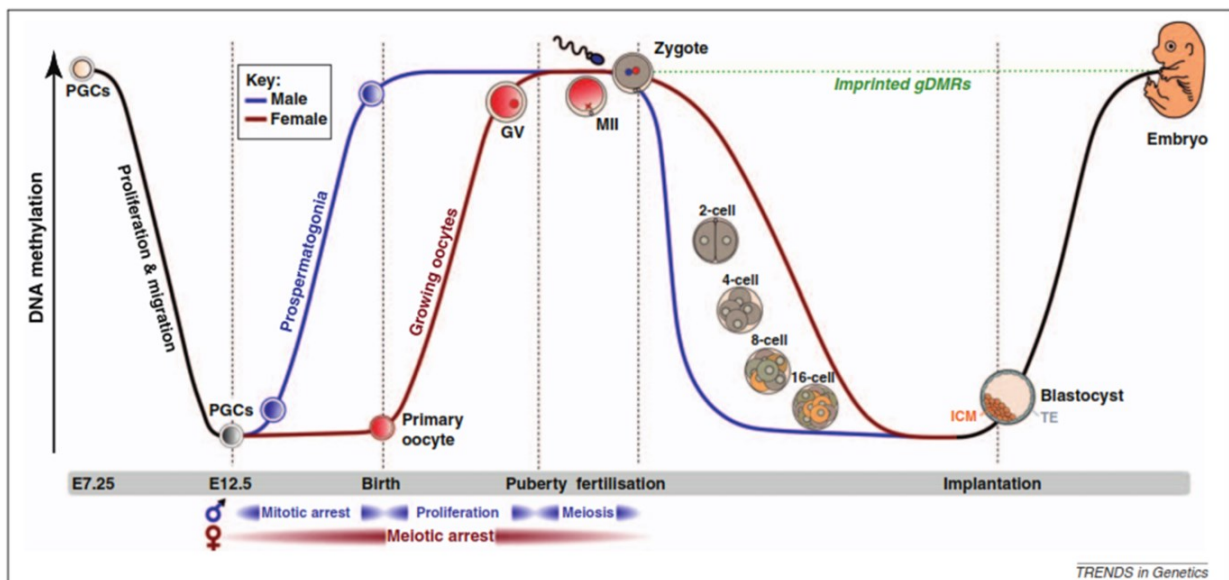
### **3. Le cycle de l'empreinte**

L'empreinte est maintenue d'une cellule mère aux cellules filles dans les cellules somatiques. En revanche, elle doit être effacée dans la lignée germinale pour y être apposée à nouveau selon le sexe de l'individu, c'est ce qu'on appelle le cycle de l'empreinte. On distingue trois phases dans ce cycle : établissement, maintenance et effacement de l'empreinte. L'établissement a lieu dans les pro-spermatogonies et les ovocytes. Après la fécondation, l'embryon des mammifères subit une vague de reprogrammation épigénétique qui consiste en une déméthylation globale de tout le génome, pour conférer le caractère totipotent à la cellule, suivie d'une reméthylation au stade blastocyste lors de son implantation (Figure 11). Pendant cette reprogrammation épigénétique, les marques de l'empreinte doivent être maintenues dans toutes les cellules somatiques et résister à cette déméthylation pour que les allèles puissent toujours être distingués selon leur origine parentale. En revanche, dans les cellules germinales ces marques doivent être effacées afin de rétablir une nouvelle marque en fonction du sexe de l'embryon. Ceci a lieu grâce à une autre vague de déméthylation de tout le génome qui s'opère dans les cellules germinales primordiales après

la détermination du sexe foetal. Dans ces cellules uniquement, les DMR des GSE vont être déméthylées (Figure 12).



**Figure 11 :** Cycle de l'empreinte chez la souris d'après (Barlow and Bartolomei, 2014). A. Phase d'établissement (DNMT3A et DNMT3L), B. Phase de maintenance (ZFP57, KAP1, Stella et DNMT1), C. Phase d'effacement (TET1/2 ?).



**Figure 12 :** variations des niveaux de méthylation pangénomique durant le développement murin. (PGC : Primordial Germ Cells, GV : Germinal vesicle, MII : Metaphase II oocyte, ICM : Inner Cell Mass, TE : Trophectoderm) (Smallwood and Kelsey, 2012).

### 3.1. Etablissement de l’empreinte

Cette phase consiste en une méthylation progressive du génome des cellules germinales après détermination du sexe fœtal. Cela concerne les ICR qui sont des DMR principales ou germinales. Les DMR secondaires, ou somatiques, sont elles mises en place après la fécondation et dépendent des ICR (Adalsteinsson and Ferguson-Smith, 2014).

La méthylation *de novo* des allèles maternels a lieu dans l’ovocyte après la naissance alors que celle des allèles paternels se met en place pendant la vie fœtale dans les pro-spermatogonies (Lucifero et al., 2002).

Les mécanismes par lesquels les ICR sont reconnus dans les gamètes afin de les méthyler ou non en fonction du sexe de l’individu restent encore très peu connus, alors que les mécanismes d’établissement des marques épigénétiques sont de mieux en mieux documentés. Des études ont tenté d’identifier les caractéristiques inhérentes à la séquence nucléotidique qui la rendent reconnaissable des systèmes de méthylation. La comparaison des séquences des DMR maternelles ainsi que l’analyse des séquences cibles des mécanismes de méthylation *de novo* (notamment DNMT3A et DNMT3L) ont permis de déduire des caractéristiques communes telles que l’enrichissement en CpG, la présence de répétitions en tandem direct et un espacement de 8 à 10 nucléotides entre les CpG (Barlow and Bartolomei, 2014). Ces caractéristiques n’étant pas parfaitement spécifiques, d’autres éléments sont nécessaires. Les modifications post-traductionnelles des histones pourraient également permettre l’identification de ces régions. La non-méthylation de H3K4 par exemple permet une interaction avec DNMT3L et une méthylation *de novo* de la séquence d’ADN sous-jacente (Ooi et al., 2007). Par ailleurs, l’activité transcriptionnelle semble être également nécessaire à l’accessibilité de la séquence d’ADN à la méthylation (Chotalia et al., 2009; Henckel et al., 2012).

Une fois que les séquences des ICR sont reconnues dans les cellules germinales, la méthylation *de novo* s’opère par DNMT3A et DNMT3L (Bourc’his and Bestor, 2004; Kaneda et al., 2004; Kato et al., 2007). La grande majorité des ICR sont méthylées dans les ovocytes contre quatre seulement identifiés à ce jour dans les cellules germinales mâles (les locus *Igf2-H19*, *Dlk1-Gtl2*, *Rasgrf1*, et *Gpr1-Zdbf2* (Tomizawa et al., 2011)). Les ICR paternels se distinguent des maternels par leur localisation. Les premiers sont intergéniques alors que les

seconds se situent dans les régions promotrices. Un des éléments pouvant expliquer la méthylation des ICR essentiellement dans les ovocytes est que la méthyl-transférase DNMT3A est recrutée vers la séquence cible par son cofacteur DNMT3L qui est lui-même attiré par la Lysine H3K4 non-méthylée. Or, les ICR non méthylés dans la lignée germinale mâle sont enrichies en H3K4 méthylée (Delaval et al., 2007; Hammoud et al., 2009), alors que dans la lignée germinale femelle l'activité de la H3K4-déméthylase KDM1B est corrélée à l'établissement de l'empreinte maternelle (Ciccone et al., 2009). D'autres mécanismes épigénétiques restent à mettre en lumière pour expliquer l'ensemble du phénomène de mise en place de l'empreinte dans les différents ICR.

### **3.2. Maintien de l'empreinte**

Les gamètes sont des cellules très différenciées qui ont une grande proportion de régions promotrices et îlots CpG qui sont méthylés. Après la fécondation, le zygote doit devenir totipotent afin de pouvoir donner naissance à différents types cellulaires nécessaires au développement. En période pré-implantatoire, le génome paternel subit une déméthylation globale active alors que le génome maternel se déméthyle progressivement de façon passive par effet de dilution des marques épigénétiques par la réplication et l'apport de nouvelles histones (Reik, 2007). Cette déméthylation progressive de la totalité du génome est suivie d'une re-méthylation globale. Or, les ICR doivent eux conserver leur méthylation différentielle qui permet la distinction entre les allèles paternels et les allèles maternels nécessaire au développement ultérieur. Pour garder ce statut de DMR, l'allèle méthylé doit être protégé contre la vague de déméthylation et l'allèle non méthylé doit être protégé contre la vague de re-méthylation. D'autres DMR non soumises à l'empreinte parentale échappent également à cette vague de déméthylation mais l'existence ou non d'un lien avec les ICR reste encore inconnu (Smallwood et al., 2011).

Diverses protéines participent au maintien de la méthylation différentielle en protégeant les allèles méthylés de la déméthylation. Parmi ces protéines : (1) la méthyl-transférase DNMT1 qui est la protéine majeure de maintien de la méthylation (Hirasawa et al., 2008) ; (2) STELLA (connue également sous les noms DPPA3 ou PGC7) jouerait aussi un rôle protecteur contre la déméthylation des ICR (Nakamura et al., 2007). Cette protéine est fortement exprimée pendant l'ovogenèse et dans l'embryon préimplantatoire. Des études récentes ont montré



que STELLA inhibe l'action de TET3 qui oxyde la 5mC en 5-hydroxyméthylCytosine (Bian and Yu, 2014; Nakashima et al., 2013). L'action protectrice de STELLA pourrait également passer par des modifications d'histones conduisant à une condensation de la chromatine (Liu et al., 2012). Elle serait recrutée au niveau des séquences méthylées par son interaction avec la marque H3K9me2 qui est enrichie dans les allèles méthylés des DMR (Nakamura et al., 2012) ; (3) ZFP57 qui a une action spécifique au niveau des ICR. Des mutations homozygotes de *ZFP57* dans sept familles de diabète néonatal transitoire, une pathologie liée à l'empreinte, ont été associées à une hypométhylation de plusieurs DMR (Mackay et al., 2008). Des données chez la souris et dans les cellules souches embryonnaires sont également en faveur d'un rôle de ZFP57 dans le maintien de la méthylation au niveau des DMR (Li et al., 2008; Quenneville et al., 2011). Cette protéine à doigt de Zinc possède un domaine KRAB et reconnaît un motif de séquence quasi spécifique des DMRs (TGCCGC). Le domaine KRAB permet la liaison d'un cofacteur qui est la protéine KRAB-Associated Protein 1 ou KAP1 également appelée TRIM28. Cette protéine est capable de recruter des modificateurs répressifs des histones tels que des protéines à domaine SET capables de méthyler la Lysine H3K9 (Quenneville et al., 2011). KAP1 est apportée par l'ovocyte pour participer au maintien de la méthylation pendant les premières divisions, puis rapidement produite par le zygote (Messerschmidt et al., 2012). ZFP57 est capable d'autres interactions impliquées dans le maintien de la méthylation telles que l'interaction du complexe ZFP57/KAP1 avec les méthyl-transférases et la protéine NP95 (ou UHRF1) qui se lie aux séquences hémiméthylées pour maintenir la méthylation (Bostick et al., 2007; Sharif et al., 2007) ; (4) Enfin, certaines modifications post-traductionnelles des histones telles que la méthylation des Lysines H3K9, H3K64 et H4K20, et de l'Arginine H4R3, ainsi que l'enrichissement en protéine HP1, sont spécifiquement associées à l'allèle méthylé mais leur rôle éventuel dans le maintien de la méthylation reste inconnu (Delaval et al., 2007).

D'autres facteurs assurent la protection de l'allèle non méthylé contre la méthylation tels que ZFP42 et CTCF. Le déficit en ces deux protéines est associé à une hyperméthylation de certains ICR (Engel et al., 2006; Kim et al., 2011; Schoenherr et al., 2003). De même, la méthylation de la Lysine H3K4 prévient l'allèle non-méthylé de la méthylation *de novo* en empêchant la fixation de DNMT3A et DNMT3B/DNMT3L. De plus, cette modification post-traductionnelle de l'histone H3 est spécifiquement enrichie au niveau de l'allèle non-

méthylé des ICR (Ooi et al., 2007; Zhang et al., 2010). Les protéines de la famille des Ten-Eleven Translocation (TET) sont à l'origine de la transformation enzymatique de la 5mC en 5hmC. Ce mécanisme de déméthylation pourrait être le garant du maintien de l'absence de méthylation d'un des deux allèles des DMR en déméthylant les Cytosines anormalement méthylées. Bien d'autres protéines seraient impliquées (Sanli et Feil, 2015). Probablement que la protection de chaque DMR met en œuvre un ensemble de protéines selon des règles encore mal connues.

### **3.3. Effacement de l'empreinte**

Cette phase a lieu dans les précurseurs des cellules germinales. Ces derniers migrent du mésoderme pour aller coloniser les crêtes génitales au stade E11,5 chez la souris et pendant la 4<sup>ème</sup> semaine du développement de l'embryon humain. Dès leur arrivée dans les crêtes génitales, elles subissent une vague de déméthylation globale de leur génome y compris des ICR (Hajkova et al., 2002; Lee et al., 2002). Cette reprogrammation épigénétique implique à la fois la déméthylation de l'ADN ainsi que des modifications des histones avec perte des marques H3K9me3, H3K27me3 et H2A/H4Rme2, et une perte des protéines HP1 (Hajkova, 2010; Hajkova et al., 2008). La déméthylation de l'ADN se fait très probablement de façon active étant donnée la rapidité avec laquelle elle s'opère (1 cycle seulement) mais les mécanismes sous-jacents ne sont pas encore élucidés. Les systèmes de réparation de l'ADN pourraient être impliqués dans ce processus (Popp et al., 2010).

## **Chapitre II : Les gènes soumis à l'empreinte parentale**

À ce jour, 151 GSE ont été identifiés chez la souris (Figure 13) (<http://www.mousebook.org>, et 96 à 106, selon les sources, chez l'homme (Tableau 2) (<http://www.geneimprint.com>, <http://igc.otago.ac.nz/home.html>). Ce groupe de gènes représente moins de 1% des gènes chez les mammifères.

## Mouse Imprinted Genes, Regions and Phenotypes

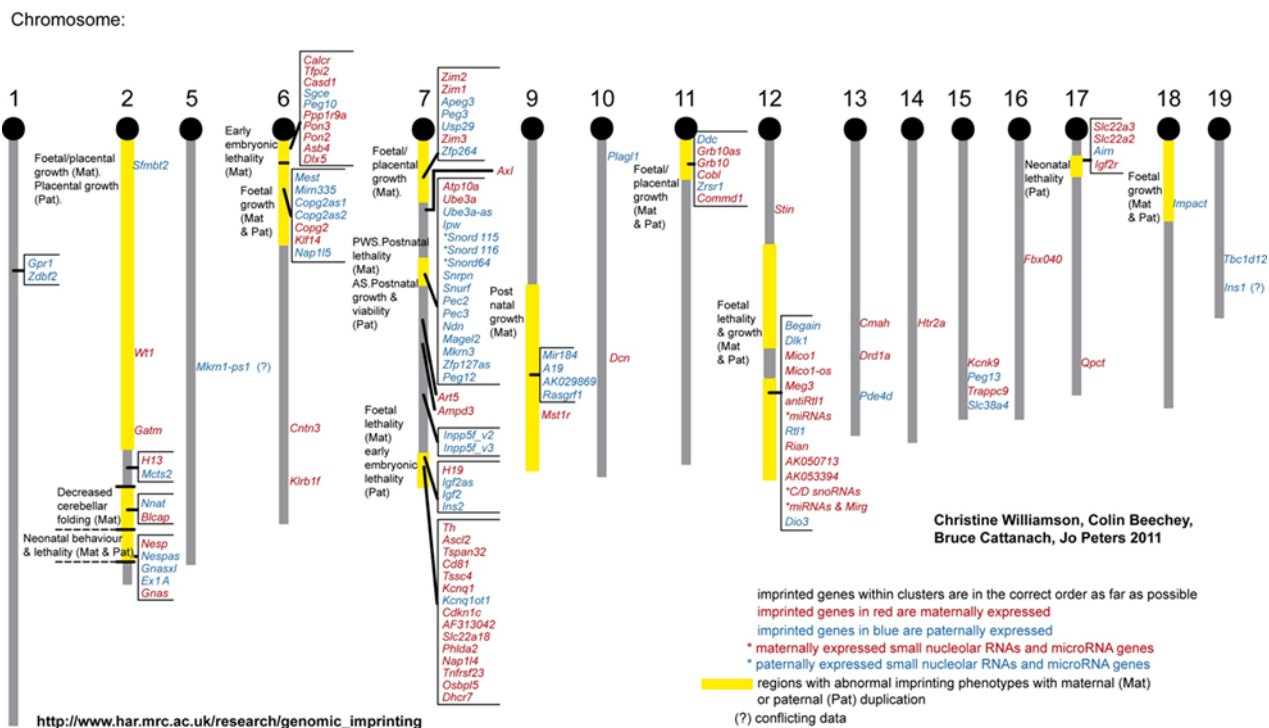


Figure 13 : Cartographie des GSE murins issue de la base de données publique [www.mousebook.org](http://www.mousebook.org).

Tableau 2 : Liste actuelle des GSE identifiés chez l'Homme. Source : [www.geneimprint.com](http://www.geneimprint.com).

Gene	Aliases	Location	Status	Expressed Allele
<a href="#">DIRAS3</a>	<a href="#">ARHI</a> , <a href="#">NOEY2</a>	1p31 AS	Imprinted	Paternal
<a href="#">RNU5D-1</a>	<a href="#">U5DL</a> , <a href="#">U5DS</a> , <a href="#">RNU5D</a>	1p34.1 AS	Imprinted	Paternal
<a href="#">TP73</a>	<a href="#">P73</a>	1p36.3	Imprinted	Maternal
<a href="#">LRRTM1</a>		2p12 AS	Imprinted	Paternal
<a href="#">GPR1</a>		2q33.3 AS	Imprinted	Paternal
<a href="#">ZDBF2</a>		2q33.3	Imprinted	Paternal
<a href="#">NAP1L5</a>	<a href="#">DRLM</a>	4q22.1 AS	Imprinted	Paternal
<a href="#">FAM50B</a>	<a href="#">X5L</a> , <a href="#">D6S2654E</a>	6p25.2	Imprinted	Paternal
<a href="#">AIM1</a>	<a href="#">ST4</a> , <a href="#">CRYBG1</a>	6q21	Imprinted	Paternal
<a href="#">LIN28B</a>	<a href="#">CSDD2</a>	6q21	Imprinted	Paternal
<a href="#">PLAGL1</a>	<a href="#">ZAC</a> , <a href="#">LOT1</a> , <a href="#">ZAC1</a> , <a href="#">MGC126275</a> , <a href="#">MGC126276</a> , <a href="#">DKFZp781P1017</a>	6q24-q25 AS	Imprinted	Paternal
<a href="#">HYMAI</a>	<a href="#">NCRNA00020</a>	6q24.2 AS	Imprinted	Paternal
<a href="#">SLC22A2*</a>	<a href="#">OCT2</a> , <a href="#">MGC32628</a>	6q26 AS	Imprinted	Maternal
<a href="#">SLC22A3*</a>	<a href="#">EMT</a> , <a href="#">EMTH</a> , <a href="#">OCT3</a>	6q26-q27	Imprinted	Maternal
<a href="#">GRB10</a>	<a href="#">RSS</a> , <a href="#">IRBP</a> , <a href="#">MEG1</a> , <a href="#">GRB-IR</a> , <a href="#">Grb-10</a> , <a href="#">KIAA0207</a>	7p12-p11.2 AS	Imprinted	Isoform Dependent
<a href="#">DDC</a>	<a href="#">AADC</a>	7p12.2 AS	Imprinted	Isoform Dependent

**Tableau 2 (suite)** : Liste actuelle des GSE identifiés chez l'Homme. Source : [www.geneimprint.com](http://www.geneimprint.com).

Gene	Aliases	Location	Status	Expressed Allele
<a href="#">MAGI2</a>	<a href="#">AIP1</a> , <a href="#">AIP-1</a> , <a href="#">ARIP1</a> , <a href="#">SSCAM</a> , <a href="#">MAGI-2</a> , <a href="#">ACVRIP1</a>	7q21 AS	Imprinted	Maternal
<a href="#">PEG10</a>	<a href="#">EDR</a> , <a href="#">HB-1</a> , <a href="#">Mar2</a> , <a href="#">MEF3L</a> , <a href="#">Mart2</a> , <a href="#">RGAG3</a>	7q21	Imprinted	Paternal
<a href="#">SGCE</a>	<a href="#">ESG</a> , <a href="#">DYT11</a>	7q21-q22 AS	Imprinted	Paternal
<a href="#">PPP1R9A</a>	<a href="#">NRB1</a> , <a href="#">NRBI</a> , <a href="#">FLJ20068</a> , <a href="#">KIAA1222</a> , <a href="#">Neurabin-1</a>	7q21.3	Imprinted	Maternal
<a href="#">DLX5</a>		7q22 AS	Imprinted	Maternal
<a href="#">TFPI2</a>	<a href="#">PP5</a> , <a href="#">REF1</a> , <a href="#">TFPI-2</a> , <a href="#">FLJ21164</a>	7q22 AS	Imprinted	Maternal
<a href="#">COPG2IT1</a>	<a href="#">CIT1</a> , <a href="#">COPG2AS</a> , <a href="#">FLJ41646</a> , <a href="#">NCRNA00170</a> , <a href="#">DKFZP761N09121</a>	7q32	Imprinted	Paternal
<a href="#">MEST</a>	<a href="#">PEG1</a> , <a href="#">MGC8703</a> , <a href="#">MGC111102</a> , <a href="#">DKFZp686L18234</a>	7q32	Imprinted	Paternal
<a href="#">CPA4</a>	<a href="#">CPA3</a>	7q32	Imprinted	Maternal
<a href="#">MESTIT1</a>	<a href="#">MEST-IT</a> , <a href="#">PEG1-AS</a> , <a href="#">MEST-</a> <a href="#">AS1</a> , <a href="#">MEST-IT1</a> , <a href="#">NCRNA00040</a>	7q32.2 AS	Imprinted	Paternal
<a href="#">KLF14</a>	<a href="#">BTEB5</a>	7q32.3 AS	Imprinted	Maternal
<a href="#">DLGAP2</a>	<a href="#">DAP2</a> , <a href="#">SAPAP2</a>	8p23	Imprinted	Paternal
<a href="#">KCNK9</a>	<a href="#">KT3.2</a> , <a href="#">TASK3</a> , <a href="#">K2p9.1</a> , <a href="#">TASK-3</a> , <a href="#">MGC138268</a> , <a href="#">MGC138270</a>	8q24.3 AS	Imprinted	Maternal
<a href="#">ZFAT-AS1</a>	<a href="#">ZFATAS</a> , <a href="#">ZFAT-AS</a> , <a href="#">SAS-</a> <a href="#">ZFAT</a> , <a href="#">NCRNA00070</a>	8q24.22	Imprinted	Paternal
<a href="#">PEG13</a>		8q24.22	Imprinted	Paternal
<a href="#">ZFAT</a>	<a href="#">AITD3</a> , <a href="#">ZFAT1</a> , <a href="#">ZNF406</a>	8q24.22 AS	Imprinted	Paternal
<a href="#">GLIS3</a>	<a href="#">ZNF515</a>	9p24.2 AS	Imprinted	Paternal
<a href="#">INPP5F V2</a>	<a href="#">SAC2</a> , <a href="#">hSAC2</a> , <a href="#">MSTP007</a> , <a href="#">MSTPO47</a> , <a href="#">FLJ13081</a> , <a href="#">KIAA0966</a> , <a href="#">MGC59773</a> , <a href="#">MGC131851</a>	10q26.11	Imprinted	Paternal
<a href="#">WT1-Alt trans</a>	<a href="#">WT1</a> , <a href="#">GUD</a> , <a href="#">WAGR</a> , <a href="#">WT33</a> , <a href="#">WIT-2</a>	11p13 AS	Imprinted	Paternal
<a href="#">KCNQ10T1</a>	<a href="#">LIT1</a> , <a href="#">KvDMR1</a> , <a href="#">KCNQ10T1</a> , <a href="#">KvLQT1-AS</a> , long QT intronic transcript 1	11p15	Imprinted	Paternal
<a href="#">KCNQ1DN</a>	<a href="#">BWRT</a> , <a href="#">HSA404617</a>	11p15.4	Imprinted	Maternal
<a href="#">OSBPL5</a>	<a href="#">ORP5</a> , <a href="#">OBPH1</a> , <a href="#">FLJ42929</a>	11p15.4 AS	Imprinted	Maternal
<a href="#">IGF2AS</a>	<a href="#">PEG8</a> , <a href="#">MGC168198</a>	11p15.5	Imprinted	Paternal
<a href="#">IGF2</a>	<a href="#">INSIGF</a> , <a href="#">pp9974</a> , <a href="#">C11orf43</a> , <a href="#">FLJ22066</a> , <a href="#">FLJ44734</a>	11p15.5 AS	Imprinted	Paternal
<a href="#">PHLDA2</a>	<a href="#">IPL</a> , <a href="#">BRW1C</a> , <a href="#">BWR1C</a> , <a href="#">HLDA2</a> , <a href="#">TSSC3</a>	11p15.5 AS	Imprinted	Maternal
<a href="#">CDKN1C</a>	<a href="#">BWS</a> , <a href="#">WBS</a> , <a href="#">p57</a> , <a href="#">BWCR</a> , <a href="#">KIP2</a>	11p15.5 AS	Imprinted	Maternal
<a href="#">H19</a>	<a href="#">ASM</a> , <a href="#">BWS</a> , <a href="#">ASM1</a> , <a href="#">MGC4485</a> , <a href="#">PRO2605</a> , <a href="#">D11S813E</a>	11p15.5 AS	Imprinted	Maternal

**Tableau 2 (suite) : Liste actuelle des GSE identifiés chez l'Homme. Source : [www.geneimprint.com](http://www.geneimprint.com).**

Gene	Aliases	Location	Status	Expressed Allele
<a href="#">KCNQ1</a>	<a href="#">LQT</a> , <a href="#">RWS</a> , <a href="#">WRS</a> , <a href="#">LQT1</a> , <a href="#">SQT2</a> , <a href="#">ATFB1</a> , <a href="#">ATFB3</a> , <a href="#">JLNS1</a> , <a href="#">KCNA8</a> , <a href="#">KCNA9</a> , <a href="#">Kv1.9</a> , <a href="#">Kv7.1</a> , <a href="#">KVLQT1</a> , <a href="#">FLJ26167</a>	11p15.5	Imprinted	Maternal
<a href="#">SLC22A18</a>	<a href="#">HET</a> , <a href="#">ITM</a> , <a href="#">BWR1A</a> , <a href="#">IMPT1</a> , <a href="#">TSSC5</a> , <a href="#">ORCTL2</a> , <a href="#">BWSCR1A</a> , <a href="#">SLC22A1L</a> , <a href="#">p45-BWR1A</a> , <a href="#">DKFZp667A184</a>	11p15.5	Imprinted	Maternal
<a href="#">INS</a>	<a href="#">ILPR</a> , <a href="#">IRDN</a>	11p15.5 AS	Imprinted	Paternal
<a href="#">ANO1</a>	<a href="#">DOG1</a> , <a href="#">TAOS2</a> , <a href="#">ORAOV2</a> , <a href="#">TMEM16A</a>	11q13.3	Imprinted	Maternal
<a href="#">ZC3H12C</a>	<a href="#">MCP1P3</a>	11q22.3	Imprinted	Paternal
<a href="#">NTM</a>	<a href="#">HNT</a> , <a href="#">NTRI</a> , <a href="#">IGLON2</a>	11q25	Imprinted	Maternal
<a href="#">RBP5</a>	<a href="#">CRBP3</a> , <a href="#">CRBP3III</a> , <a href="#">CRBP-III</a>	12p13.31 AS	Imprinted	Maternal
<a href="#">RB1</a>	<a href="#">RB</a> , <a href="#">pRb</a> , <a href="#">OSRC</a> , <a href="#">pp110</a> , <a href="#">p105-Rb</a>	13q14.2	Imprinted	Maternal
<a href="#">DLK1</a>	<a href="#">DLK</a> , <a href="#">FA1</a> , <a href="#">ZOG</a> , <a href="#">pG2</a> , <a href="#">PREF1</a> , <a href="#">Pref-1</a>	14q32	Imprinted	Paternal
<a href="#">MEG3</a>	<a href="#">GTL2</a> , <a href="#">FP504</a> , <a href="#">prebp1</a> , <a href="#">PRO0518</a> , <a href="#">PRO2160</a> , <a href="#">FLJ31163</a> , <a href="#">FLJ42589</a>	14q32	Imprinted	Maternal
<a href="#">RTL1</a>	<a href="#">MART1</a> , <a href="#">PEG11</a> , <a href="#">LOC388015</a>	14q32.31 AS	Imprinted	Paternal
<a href="#">MEG8</a>	<a href="#">Rian</a>	14q32.31	Imprinted	Maternal
<a href="#">MAGEL2</a>	<a href="#">nM15</a> , <a href="#">NDNL1</a>	15q11-q12 AS	Imprinted	Paternal
<a href="#">UBE3A</a>	<a href="#">AS</a> , <a href="#">ANCR</a> , <a href="#">E6-AP</a> , <a href="#">HPVE6A</a> , <a href="#">EPVE6AP</a> , <a href="#">FLJ26981</a>	15q11-q13 AS	Imprinted	Maternal
<a href="#">NPAP1</a>	<a href="#">C15orf2</a>	15q11-q13	Imprinted	Unknown
<a href="#">MKRN3</a>	<a href="#">D15S9</a> , <a href="#">RNF63</a> , <a href="#">ZFP127</a> , <a href="#">ZNF127</a> , <a href="#">MGC88288</a>	15q11-q13	Imprinted	Paternal
<a href="#">ATP10A</a>	<a href="#">ATPVA</a> , <a href="#">ATPVC</a> , <a href="#">ATP10C</a> , <a href="#">KIAA0566</a>	15q11.2 AS	Imprinted	Maternal
<a href="#">PWCR1</a>	<a href="#">PET1</a> , non-coding RNA in the Prader-Willi critical region	15q11.2	Imprinted	Paternal
<a href="#">SNORD107</a>	<a href="#">HBII-436</a> , <a href="#">HBII-436 C/D box snoRNA</a>	15q11.2	Imprinted	Paternal
<a href="#">SNORD108</a>	<a href="#">HBII-437</a> , <a href="#">HBII-437 C/D box snoRNA</a>	15q11.2	Imprinted	Paternal
<a href="#">SNORD115@</a>	<a href="#">HBII-52</a>	15q11.2	Imprinted	Paternal
<a href="#">SNORD115-48</a>	<a href="#">HBII-52-48</a>	15q11.2	Imprinted	Paternal
<a href="#">SNORD116@</a>	<a href="#">PET1</a> , <a href="#">PWCR1</a> , <a href="#">HBII-85</a>	15q11.2	Imprinted	Paternal
<a href="#">SNORD109B</a>	<a href="#">HBII-438B</a> , <a href="#">HBII-438B C/D box snoRNA</a>	15q11.2	Imprinted	Paternal
<a href="#">SNORD109A</a>	<a href="#">HBII-438A</a>	15q11.2	Imprinted	Paternal

**Tableau 2 (suite) : Liste actuelle des GSE identifiés chez l'Homme. Source : [www.geneimprint.com](http://www.geneimprint.com).**

Gene	Aliases	Location	Status	Expressed Allele
<a href="#">SNRPN</a>	<a href="#">SMN</a> , <a href="#">PWCR</a> , <a href="#">SM-D</a> , <a href="#">RT-LI</a> , <a href="#">HCERN3</a> , <a href="#">SNRNP-N</a> , <a href="#">FLJ33569</a> , <a href="#">FLJ36996</a> , <a href="#">FLJ39265</a> , <a href="#">MGC29886</a> , <a href="#">SNURF-SNRPN</a> , <a href="#">DKFZp762N022</a> , <a href="#">DKFZp686C0927</a> , <a href="#">DKFZp76111912</a> , <a href="#">DKFZp686M12165</a>	15q11.2	Imprinted	Paternal
<a href="#">NDN</a>	<a href="#">HsT16328</a>	15q11.2-q12 AS	Imprinted	Paternal
<a href="#">SNORD64</a>	<a href="#">HBII-13</a> , <a href="#">HBII-13 snoRNA</a>	15q12	Imprinted	Paternal
<a href="#">SNURF</a>		15q12	Imprinted	Paternal
<a href="#">IRAIN</a>	<a href="#">IGF1R-AS</a>	15q26.3 AS	Imprinted	Paternal
<a href="#">ZNF597</a>		16p13.3 AS	Imprinted	Maternal
<a href="#">NAA60</a>	<a href="#">HAT4</a> , <a href="#">NAT15</a>	16p13.3	Imprinted	Maternal
<a href="#">TCEB3C</a>	<a href="#">HsT829</a> , <a href="#">TCEB3L2</a> , <a href="#">Elongin A3</a>	18q21.1 AS	Imprinted	Maternal
<a href="#">DNMT1</a>	<a href="#">AIM</a> , <a href="#">DNMT</a> , <a href="#">MCMT</a> , <a href="#">CXXC9</a> , <a href="#">HSN1E</a> , <a href="#">ADCADN</a>	19p13.2 AS	Imprinted	Paternal
<a href="#">MIMT1</a>	<a href="#">MIM1</a> , <a href="#">LINC00067</a> , <a href="#">NCRNA00067</a>	19q13.4	Imprinted	Paternal
<a href="#">PEG3</a>	<a href="#">PW1</a> , <a href="#">ZNF904</a> , <a href="#">ZSCAN24</a>	19q13.4 AS	Imprinted	Paternal
<a href="#">ZIM2</a>	<a href="#">ZNF656</a>	19q13.4 AS	Imprinted	Paternal
<a href="#">NLRP2</a>	<a href="#">NBS1</a> , <a href="#">PAN1</a> , <a href="#">NALP2</a> , <a href="#">PYPAF2</a> , <a href="#">CLR19.9</a>	19q13.42	Imprinted	Maternal
<a href="#">MIR371A</a>	<a href="#">C19MC</a> , <a href="#">MIR371</a> , <a href="#">MIRN371</a> , <a href="#">hsa-mir-371</a> , <a href="#">hsa-mir-371a</a>	19q13.42	Imprinted	Paternal
<a href="#">PSIMCT-1</a>		20q11.2	Imprinted	Paternal
<a href="#">BLCAP</a>	<a href="#">BC10</a>	20q11.2-q12 AS	Imprinted	Isoform Dependent
<a href="#">NNAT</a>	<a href="#">Peg5</a>	20q11.2-q12	Imprinted	Paternal
<a href="#">MCTS2</a>		20q11.21	Imprinted	Paternal
<a href="#">GDAP1L1</a>	<a href="#">dj881L22.1</a> , <a href="#">dj995J12.1.1</a>	20q12	Imprinted	Paternal
<a href="#">SGK2</a>	<a href="#">H-SGK2</a> , <a href="#">dj138B7.2</a>	20q13.2	Imprinted	Paternal
<a href="#">GNAS</a>	<a href="#">AHO</a> , <a href="#">GSA</a> , <a href="#">GSP</a> , <a href="#">POH</a> , <a href="#">GPSA</a> , <a href="#">NESP</a> , <a href="#">GNAS1</a> , <a href="#">PHP1A</a> , <a href="#">PHP1B</a> , <a href="#">C20orf45</a> , <a href="#">MGC33735</a> , <a href="#">dj309F20.1.1</a> , <a href="#">dj806M20.3.3</a>	20q13.3	Imprinted	Isoform Dependent
<a href="#">L3MBTL</a>	<a href="#">L3MBTL1</a> , <a href="#">FLJ41181</a> , <a href="#">KIAA0681</a> , <a href="#">H-L(3)MBT</a> , <a href="#">dj138B7.3</a> , <a href="#">DKFZp586P1522</a>	20q13.12	Imprinted	Paternal
<a href="#">MIR296</a>	<a href="#">MIRN296</a> , <a href="#">miRNA296</a>	20q13.32 AS	Imprinted	Paternal
<a href="#">GNASAS</a>	<a href="#">SANG</a> , <a href="#">NESPAS</a> , <a href="#">GNAS1AS</a> , <a href="#">NCRNA00075</a>	20q13.32 AS	Imprinted	Paternal
<a href="#">MIR298</a>	<a href="#">MIRN298</a> , <a href="#">hsa-mir-298</a>	20q13.32 AS	Imprinted	Paternal
<a href="#">SANG</a>	<a href="#">SANG</a> , <a href="#">Nespas</a>	20q13.32	Imprinted	Paternal
<a href="#">DGCR6L</a>		22q11.21 AS	Imprinted	Random
<a href="#">DGCR6</a>		22q11.21	Imprinted	Random

## 1. Caractéristiques des GSE

Les GSE sont répartis dans tout le génome y compris sur le chromosome X. La grande majorité d'entre eux est organisée en clusters d'environ 1 Mb qui sont sous le contrôle d'éléments en *cis* à savoir les ICRs. Néanmoins, des GSE isolés ont également été décrits.

La régulation épigénétique de l'empreinte varie d'un cluster/GSE à l'autre mais des caractéristiques communes sont observées. Ce groupe de gènes doit posséder en sa séquence des éléments permettant de le reconnaître parmi d'autres pour y mettre en place les mécanismes épigénétiques de l'expression mono-allélique selon l'origine parentale de l'allèle. L'étude comparative des séquences de différents GSE a montré qu'il y a un enrichissement en sites CpG qui peuvent donner lieu à une méthylation différentielle des deux allèles mais aussi à une acquisition différentielle de modifications post-traductionnelles des histones. D'autres caractéristiques sont retrouvées de façon inconstante telles que la présence de sites de fixation du facteur CTCF ou du facteur YY1, ou la présence d'ARN non codants longs (Ishida and Moore, 2013). La composition en séquences répétées au niveau des GSE diffère du reste du génome. Les Long Interspersed Nuclear Elements (LINE) y sont sous-représentés en faveur des répétitions en tandem qui elles y sont enrichies. Cette répartition non aléatoire participerait à l'identité des GSE sans pour autant être un élément principal (Abramowitz and Bartolomei, 2012; Walter et al., 2006).

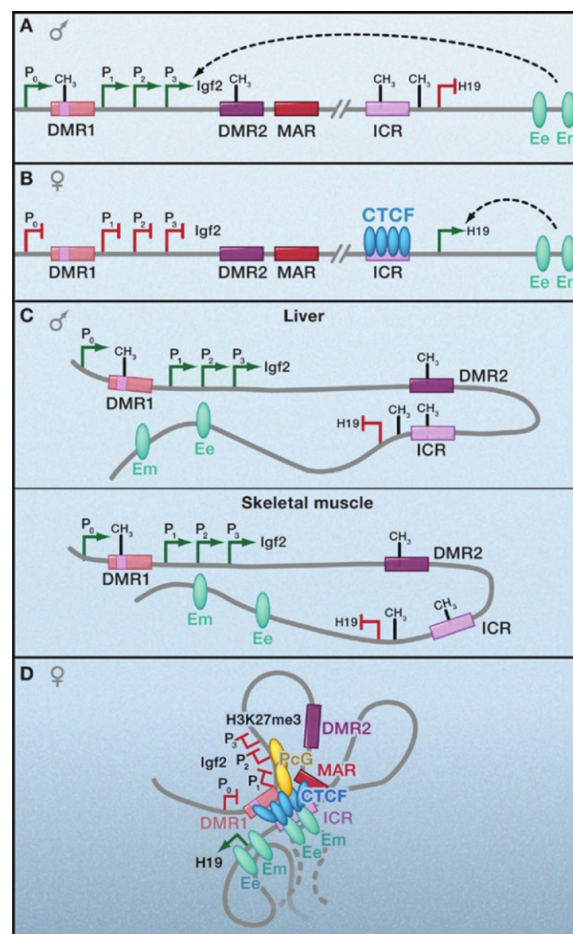
## 2. Régulation de l'empreinte

### 2.1. Deux grands modèles de régulation de l'empreinte

Deux modèles de contrôle de l'expression des GSE ont été identifiés à partir de deux clusters de GSE : le modèle de l'ARNnc long des locus *IGF2R* et *KCNQ1* et le modèle de l'insulateur comme dans le locus *IGF2/H19*.

Dans le cas d'*IGF2/H19*, l'ICR se situe entre *H19* et *IGF2*. Il est méthylé sur l'allèle paternel. Un enhancer capable de favoriser l'expression de ces deux gènes est localisé en leur aval (Figure 14). L'ICR de l'allèle maternel n'est pas méthylé, ce qui le rend accessible au facteur CTCF pour CCCTC-binding Factor. Ce facteur va avoir un effet insulateur en se fixant à la fois à l'ICR et à une DMR secondaire en amont d'*IGF2* ce qui va créer une boucle chromatinienne qui va éloigner l'enhancer d'*IGF2* mais le rapprocher de *H19*. Ces boucles impliquent

également les cohésines (Wendt et al., 2008). Il en résulte l'expression de *H19* et l'absence d'expression d'*IGF2* à partir de l'allèle maternel. Du côté de l'allèle paternel, l'ICR est méthylé ce qui empêche la fixation du facteur CTCF. Une boucle chromatinienne peut alors se former pour rapprocher l'enhancer de *IGF2* qui va pouvoir donc s'exprimer (Hark et al., 2000; Kanduri et al., 2000; Phillips and Corces, 2009). En revanche, la méthylation de l'ICR paternel induit, selon un mécanisme qui reste à identifier, une méthylation secondaire du promoteur de *H19* empêchant sa transcription. D'autres GSE possédant des sites de fixation pour CTCF, *Grb10*, *Rasgrf1* et *Kcnq1ot1*, pourraient être régulés de la même façon.



**Figure 14 :** Modèle de régulation de l’empreinte mettant en œuvre le facteur CTCF au niveau du locus *H19/Igf2* (Phillips and Corces, 2009).

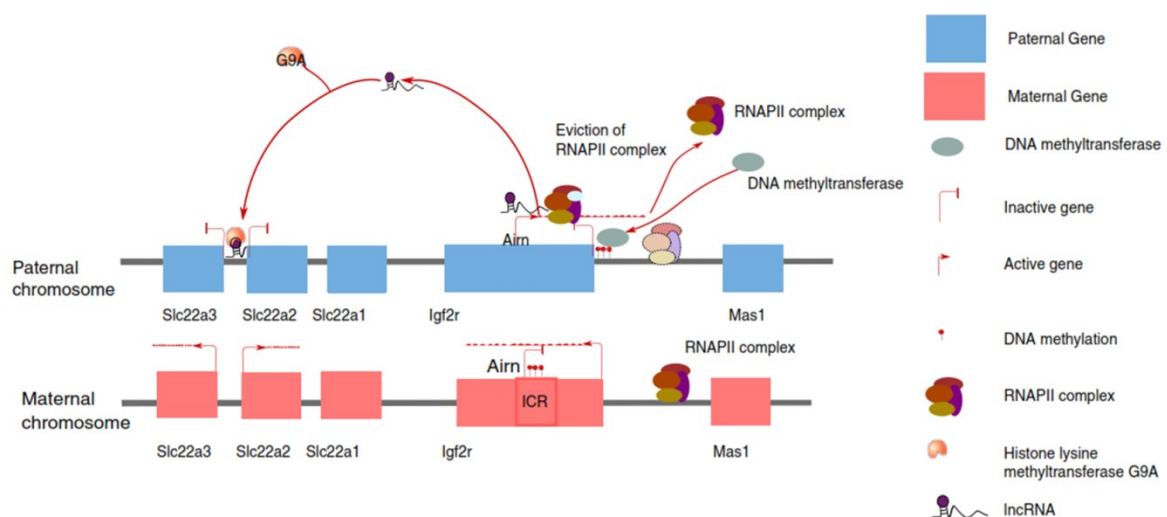
(A et B) Représentation linéaire du locus au niveau de l’allèle paternel et de l’allèle maternel respectivement. L’allèle paternel est méthylé au niveau de l’ICR (A) alors que l’allèle maternel ne l’est pas (B). Les éléments verts correspondent à des Enhancers.

(C) Représentation de la conformation de la chromatine dans l’allèle paternel méthylé. L’enhancer se retrouve proche d’*Igf2* dont il active la transcription mais éloigné de *H19* sur lequel il ne peut plus exercer son action.

(D) Représentation de la conformation chromatinienne au niveau de l’allèle maternel non méthylé. CTCF qui s’est fixé à l’ICR crée une boucle chromatinienne reliant l’ICR à une DMR secondaire en amont d’*Igf2*, ce qui cette fois rapproche l’enhancer de *H19* et active sa transcription et l’éloigne de *Igf2*.



L'autre modèle de régulation de l'empreinte est basé sur l'action en *cis* d'ARNnc longs. Un des exemples les mieux décrits est celui du locus *Igf2r*. L'ICR est localisé dans le deuxième intron du gène *Igf2r*. Il est méthylé uniquement sur l'allèle maternel. Un lncRNA *Airn* est transcrit à partir d'une séquence chevauchant celle d'*Igf2r* et en antisens. Sa région promotrice est au sein de l'ICR donc il n'est transcrit que lorsque l'ICR n'est pas méthylé c'est-à-dire à partir de l'allèle paternel. Lorsqu'il est transcrit, *Airn* réprime l'expression des GSE en *cis*, *Igf2r*, *Slc22a2* et *Slc22a3*. Ces derniers s'expriment à partir de l'allèle maternel où l'ICR est méthylé empêchant la transcription de *Airn* (Figure 15). *Airn* exercerait son action répressive *via* différents mécanismes selon les gènes cibles : (1) la transcription de *Airn* interfère avec le recrutement de l'ARN polymérase II au promoteur d'*Igf2r* et donc sa transcription (Latos et al., 2012). (2) *Airn* pourrait créer une enveloppe autour des GSE en *cis* comme c'est le cas pour *Xist* dans l'inactivation de l'X et empêcher ainsi l'accès de la machinerie transcriptionnelle aux GSE *Slc22a2* et *Slc22a3*. Un argument en faveur de ce mode d'action est la localisation de l'ARN de *Airn*, montrée par FISH, créant un « nuage » autour des gènes cibles (Nagano et al., 2008). (3) *Airn* est également capable de recruter des enzymes de modifications répressives des histones telles que EHMT2 (ou G9a) qui méthyle la Lysine H3K9 au niveau des promoteurs des GSE voisins (Nagano et al., 2008). Ce modèle s'applique également à d'autres ARNnc longs dans d'autres loci soumis à l'empreinte tels que *Kcnq1ot1*, *Nespas* et *Ube3a-AS*.



**Figure 15 :** Modèle de régulation de l'empreinte mettant en œuvre un lncRNA au locus *Airn/Igf2r*. (Kanduri, 2015). L'ICR est déméthylée sur l'allèle paternel à partir duquel *Airn* peut donc s'exprimer. *Airn* empêche la transcription d'*Igf2r* en empêchant la fixation de la machinerie transcriptionnelle au niveau de son promoteur (RNA polymérase II). La répression de *Slc22a2* et *Slc22a3* par *Airn* est quant à elle médiée par le recrutement de l'histone méthyl-transférase G9a qui méthyle la Lysine H3K9.

## 2.2. Dynamique de l’empreinte parentale

En général, les GSE ne s’expriment qu’à partir d’un seul des deux allèles. Néanmoins, pour certains GSE une grande variabilité de l’empreinte a été mise en évidence avec des situations d’empreinte partielle où l’un des deux allèles n’est pas totalement réprimé, ou encore des cas où l’empreinte présente une variabilité inter-espèces, inter-tissus, inter-cellules, inter-isoformes et parfois même inter-individus. Cette empreinte peut également être dynamique dans le temps. Ces notions sont de plus en plus étudiées grâce au développement des techniques à haut débit (Babak et al., 2015; Baran et al., 2015; Stelzer et al., 2015).

Même si la grande majorité des GSE humains sont également soumis à l’empreinte chez la souris et vice-versa, il y a des GSE spécifiques de l’une ou l’autre espèce tels que *LIN28B* qui est soumis à l’empreinte chez l’humain mais pas chez la souris, et *Commd1* et *Gab1* qui, à l’inverse, sont soumis à l’empreinte chez la souris mais pas chez l’humain. D’autres GSE ont une empreinte tissu-spécifique. Prickett et Oakey ont mené un travail d’analyse de données précédemment publiées pour évaluer la variabilité intertissulaire de l’empreinte pour des GSE chez la souris (Prickett and Oakey, 2012). Parmi les GSE pour lesquels l’empreinte a été étudiée dans au moins deux tissus, ils ont distingué deux groupes de GSE à empreinte tissu-spécifique : le premier contient des gènes soumis à l’empreinte spécifiquement dans un seul tissu, et le deuxième regroupe des GSE ayant une expression monoallélique dans plusieurs tissus mais aussi sporadiquement bi-allélique. Cette étude a montré que 28 % des gènes sont soumis à l’empreinte dans un seul des tissus analysés. Pour ces GSE, 48 % ne sont soumis à l’empreinte que dans le placenta (ex : *Ampd3*, *Dcn*, *Ascl2*, *Slc22a2*, *Slc22a3*, etc), 9 % uniquement dans le sac vitellin (*Ins1* et *Ins2*) et pour 39 % d’entre eux l’empreinte n’est détectée que dans le cerveau (ex : *Cacr*, *Copg2*, *Ube3a*, *Zfp264*, etc). D’un point de vue évolutif, cette spécificité tissulaire de l’empreinte permettrait une protection contre des mutations délétères sur l’allèle actif. L’expression bi-allélique est un moyen de contre-balancer la mutation par l’expression du deuxième allèle sauvage. Par ailleurs, la spécificité de l’empreinte dépend aussi de la fonction des gènes. Pour un GSE ayant une fonction universelle telle que la croissance pour *Igf2* et *Igf2r*, l’empreinte est ubiquitaire. Alors que des GSE agissant sur l’apport des ressources par le placenta ou par le comportement alimentaire guidé par le cerveau, ont développé une empreinte spécifique au placenta ou au

cerveau respectivement (Prickett and Oakey, 2012). Par exemple, *Gnas* qui a une empreinte complexe, a un transcrit à expression paternelle spécifiquement dans différentes régions du cerveau contrôlant les muscles oro-faciaux. Les souris déficientes en *Gnas* présentent une mortalité péri-natale en raison d'un manque d'alimentation car elles sont incapables de téter (Cattanach, 1986; Plagge et al., 2004). Le deuxième groupe de gènes montre une empreinte ubiquitaire sauf pour un tissu où l'expression est bi-allélique. Ceci est le cas pour des GSE tels que *Igf2*, *Igf2r*, *Dlk1*, *Plagl1*, etc. L'empreinte serait perdue dans un tissu où elle n'aurait plus d'intérêt ou bien par besoin essentiel de dosage génique comme c'est le cas de *Dlk1* dans des cellules nerveuses où son expression bi-allélique est nécessaire au développement et s'opère par acquisition post-natale d'une méthylation de l'ADN au niveau de l'ICR entraînant la réactivation de l'allèle maternel (Ferron et al., 2011). Dans une étude publiée récemment, les auteurs ont analysé le statut d'empreinte dans 37 tissus/types cellulaires différents issus de prélèvements humains (Baran et al., 2015). Les résultats ont mis en évidence une importante variabilité inter-tissulaire de l'empreinte. La variabilité existe même au niveau cellulaire comme c'est le cas pour *CTNNA3* qui est soumis à l'empreinte dans le placenta avec une expression mono-allélique maternelle dans le cytotrophoblaste villositaire et une expression bi-allélique dans le trophoblaste extra-villositaire. Ceci est également le cas pour *CDKN1C*, ce qui soutiendrait un rôle de l'empreinte dans la différenciation trophoblastique (van Dijk et al., 2004). Les mécanismes moléculaires sous-jacents restent à élucider mais une étude récente met l'accent sur le rôle central des promoteurs alternatifs (Stelzer et al., 2015). La spécificité tissulaire de l'empreinte peut également se manifester par un switch de l'origine parentale de l'allèle exprimé entre un tissu et un autre. *Grb10*, par exemple, est exprimé à partir de l'allèle paternel dans le cerveau et de l'allèle maternel dans d'autres tissus tels que le muscle, la thyroïde ou le rein, et ceci est en plus isoforme-dépendant (McCann et al., 2001). Ceci est régulé par une DMR secondaire spécifique du cerveau qui est hyperméthylée sur les deux allèles dans les autres tissus, et par une perte de la marque répressive H3K27 tri-méthylée de l'allèle paternel au niveau des cellules nerveuses (Arnaud et al., 2003; Sanz et al., 2008).

Ce dynamisme spatial de l'empreinte peut être relié à un dynamisme temporel. *Kcnq1* perd son statut de GSE à la mi-gestation dans le cerveau et le rein, et entre E13,5 et E14,5 dans le cœur. Des études montrent que cette réversion de l'expression mono-allélique à

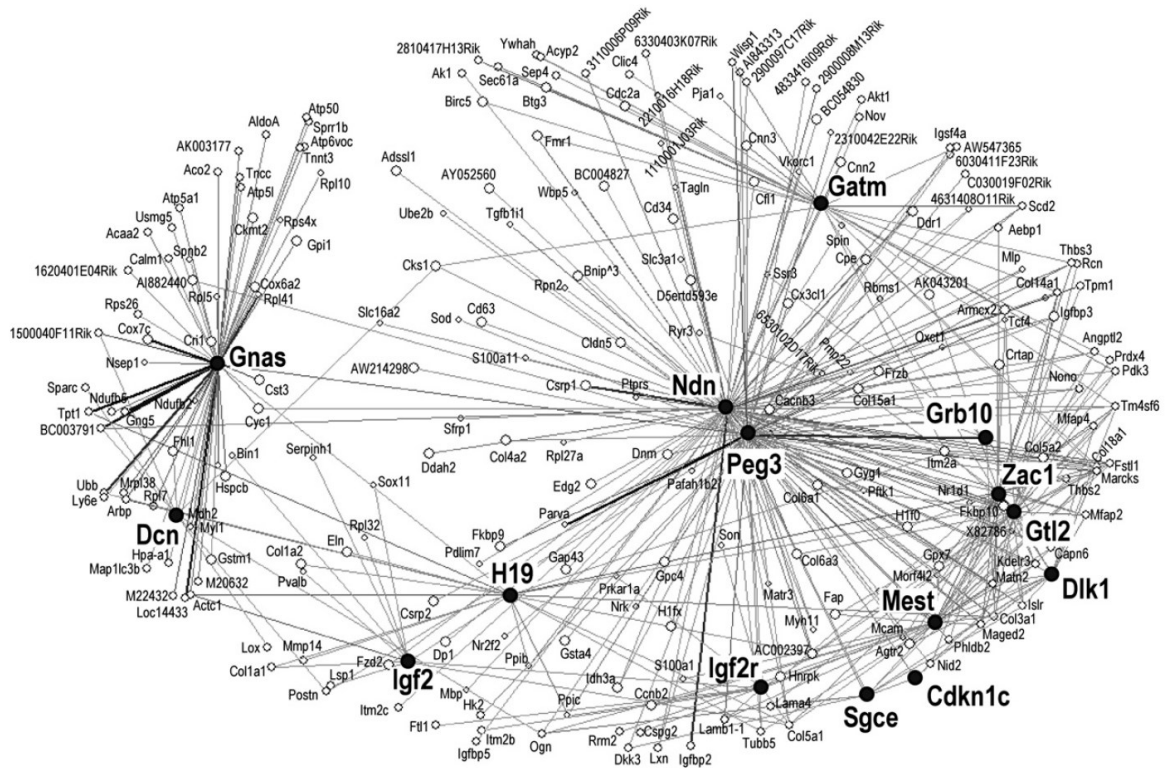
l'expression biallélique ne s'accompagne pas d'une déméthylation du promoteur de *Kcnq1* au niveau de l'allèle réprimé mais d'un changement de la conformation locale de la chromatine (Korostowski et al., 2011). Dans le placenta humain, l'empreinte parentale évolue également dans le temps. Pozharny et al. ont mis en évidence une expression biallélique, avec un ratio allélique variable, pour 14 GSE placentaires analysés (*H19*, *IGF2*, *MEST*, *PLAGL1*, etc) dans des placentas humains du premier trimestre et à terme. Ils ont qualifié ce phénomène de perte d'empreinte et ceci semble évoluer dans le temps puisque la proportion d'échantillons où l'expression était biallélique est significativement plus importante au premier trimestre qu'à terme (Pozharny et al., 2010). Ce changement de statut s'effectue sans changement du niveau d'expression des gènes et jouerait un rôle dans le développement et la maturation placentaire. A l'inverse, des GSE comme *CTNNA3* ne sont soumis à l'empreinte que pendant le premier trimestre (van Dijk et al., 2004). Par ailleurs, Marcho et al. ont étudié les variations épigénétiques (Méthylation de l'ADN et modifications des histones) temporelles et tissulaires au niveau du locus *Igf2r/Airn* pendant le développement embryonnaire. Ce travail a mis en évidence la mise en place de l'empreinte après la gastrulation de façon concomitante avec des modifications épigénétiques différentielles entre les deux allèles, en accord avec une expression monoallélique selon l'origine parentale (Marcho et al., 2015).

Un gène peut également être soumis à une empreinte isoforme-spécifique. Ce phénomène serait sous le contrôle d'une DMR spécifique d'une isoforme (Stelzer et al., 2015) et caractérise quelques GSE tels que *MEST*, *GRB10* et *GNAS* (Blagitko et al., 2000; Kamei et al., 2007; Peters et al., 2006). Cette spécificité d'isoforme peut se traduire par une mise en place de l'empreinte pour un transcrite et pas un autre, ou même par une différence de l'origine parentale de l'allèle exprimé entre différentes isoformes dans différents tissus comme c'est le cas pour *GNAS*.

Enfin, l'empreinte se complexifie encore pour certains GSE qui montrent une variabilité inter-individuelle. L'analyse de placentas humains a permis de mettre en évidence des situations fréquentes de perte de l'empreinte avec une expression de l'allèle réprimé qui peut être partielle et minoritaire par rapport au niveau d'expression de l'allèle opposé (*PLAGL1* ou *IGF2*), voire complète (*PEG3* ou *MEST*-isoforme 2) (Lambertini et al., 2008; Tycko, 2006).

### 2.3. Réseau de GSE

La notion de réseau de GSE a émergé suite à deux travaux. Le premier, de Varrault et al., a montré que l'expression de *Zac1* (ou *Plagl1*), qui est un GSE exprimé à partir de l'allèle paternel et qui code pour un facteur de transcription, est corrélée à celle d'un ensemble de GSE répartis dans tout le génome (Varrault et al., 2006) (Figure 16). Une analyse transcriptomique globale a ainsi permis de montrer que l'expression de *Plagl1* est co-régulée avec celle de 353 autres gènes dont 13 GSE. Ce qui représente un enrichissement significatif en GSE. Ensuite, les auteurs ont ré-analysé les données des transcriptomes en centrant la recherche sur les liaisons les plus significatives autour de 60 GSE, ce qui a abouti à l'identification d'un réseau de 246 gènes co-régulés comprenant des GSE et des gènes non soumis à l'empreinte et qui participent à la croissance fœtale. Ce réseau transcriptionnel aurait pour intérêt d'équilibrer les forces afin de conférer une meilleure adaptabilité de l'organisme aux variations génétiques et environnementales. Cette même équipe a complété ces résultats quelques années après et affirmé que la quasi totalité des GSE appartiennent à un seul et même réseau dont la fonction principale est le contrôle du cycle cellulaire et la différenciation, probablement via les protéines de la matrice extra-cellulaire (Al Adhami et al., 2015). En plus de cette étude, une deuxième équipe a mené un travail d'analyse des interactions chromatinienne en utilisant une technique de capture de la chromatine 4C (*Capturing Chromatin Conformation in a fourth dimension*) (Zhao et al., 2006). Les séquences liées physiquement seraient potentiellement co-régulées. Les auteurs se sont intéressés aux séquences physiquement associées à l'ICR du locus *Igf2-H19* et ont trouvé 114 sites d'interaction dont 21 correspondaient à 15 loci soumis à l'empreinte dont *Plagl1*. Là encore, il y avait un enrichissement significatif en GSE suggérant un réseau d'interactions chromatinienne de GSE. Le rôle de ce réseau chromatinien dans l'orchestration de l'expression de GSE reste encore à définir. Ce réseau pourrait être tissu- et espèce-spécifique (Smits and Kelsey, 2006).



**Figure 16 :** Réseau des Gènes Soumis à l’Empreinte parentale, comportant 246 gènes coréglés dont 15 GSE (ici en gras) (Varrault et al., 2006).

### 3. Méthodes d’identification des GSE

Depuis la découverte de l’empreinte parentale, des recherches actives de nouveaux GSE ont été menées afin d’enrichir ce panel de gènes et permettre une meilleure analyse de leurs fonctions, interactions et modes de régulation communs. Différentes approches ont été mises en œuvre : (1) l’approche clinique et expérimentale, (2) l’approche globale, ou encore (3) l’approche *in silico*.

#### 3.1. L’approche clinique et expérimentale

Les observations de certains cliniciens quant à la transmission parent dépendante de certaines maladies génétiques humaines, associées à l’identification d’anomalies chromosomiques communes chez ces patients a permis de définir des régions contenant des GSE. Ces études de corrélations génotype-phénotype ont notamment été menées pour les syndromes de Prader-Willi, Beckwith-Wiedemann, Silver-Russel, etc.

Chez la souris, l'étude de modèles murins portant des disomies uniparentales de chromosomes entiers ou de segments chromosomiques ont également abouti à l'identification de nouveaux GSE. La disomie uniparentale correspond à la présence de chromosomes/segments chromosomiques, d'une même paire de chromosomes, provenant du même parent au lieu d'avoir un exemplaire de chaque parent (Cattanach and Kirk, 1985). L'étude consistait à définir les régions pour lesquelles la disomie uniparentale qu'elle soit maternelle ou paternelle était associée à un phénotype anormal notamment de la croissance et du développement foeto-placentaire. Ces expériences ont été à la base de la première cartographie des régions contenant potentiellement des clusters de gènes soumis à l'empreinte parentale .

Cette approche a été très efficace mais les résultats n'étaient pas exhaustifs car seuls les défauts d'empreinte entraînant d'importantes anomalies développementales ont pu être mis en évidence. Les anomalies de l'empreinte associées à des phénotypes plus modérés ne pouvaient être suspectées. D'où l'intérêt d'autres approches globales qui ont été possibles grâce au développement des technologies génomiques à haut débit ainsi que des outils bio-informatiques.

### **3.2. L'approche globale**

Cette approche est basée sur l'analyse globale de l'expression allélique des gènes, ou des caractéristiques épigénétiques notamment le profil de méthylation et les modifications des histones. Les modèles d'étude adoptés varient des lignées cellulaires aux souris parthénogénétiques en passant par des organes spécifiques (notamment le placenta et le cerveau).

#### *3.2.1. Analyse du transcriptome*

Le but de ces analyses est de mettre en évidence des gènes à expression monoallélique donc pouvant être soumis à empreinte parentale. Ensuite, l'étude de l'origine parentale de l'allèle exprimé est nécessaire pour écarter les expressions monoalléliques aléatoires qui sont fréquentes mais sans lien avec l'empreinte parentale (Gimelbrant et al., 2007). Les analyses étant effectuées sur un tissu ou type cellulaire donné correspondant à un stade donné du développement ou à un état de différenciation donné, les résultats ne reflètent que l'état de

ce matériel à un temps donné. L’empreinte parentale étant très dynamique pour un grand nombre de GSE (cf Section « Dynamique de l’empreinte parentale », p. 49), les résultats de ce type d’analyses ne sont pas exhaustifs.

De nombreuses études ont mis en œuvre des puces de génotypage pour analyser la transmission allélique de variants nucléotidiques (*Single Nucleotide Polymorphism*, SNP) (Bjornsson et al., 2008; Pollard et al., 2008). La limite des puces de génotypage réside dans le fait qu’elles ne permettent d’analyser que des variations connues et qui y sont représentées ce qui ne représente pas l’ensemble des variations du génome et ne cible pas la totalité des gènes du génome. En outre, la quantification par fluorescence du ratio allélique n’est pas suffisamment précise pour permettre la mise en évidence des empreintes partielles où il existe une perte partielle de l’empreinte sur l’allèle normalement réprimé ni les empreintes spécifiques d’un type cellulaire dans un échantillon complexe (cf Section « Dynamique de l’empreinte parentale », p. 49). Enfin, elle ne permet pas de distinguer les différents transcrits d’un gène donc ne détecte pas les empreintes isoforme-spécifiques lorsque plusieurs isoformes sont représentées dans l’échantillon analysé. Récemment, l’évolution des techniques de séquençage en particulier le séquençage massivement parallèle de l’ARN (RNA-Seq) ont offert une alternative intéressante. Parmi les équipes ayant mis à profit cette technologie, Metsalu et al. ont effectué un criblage de placentas humains et ont pu identifier 4 nouveaux gènes potentiellement soumis à l’empreinte dans le placenta (*LGALS8*, *LGALS14*, *PAPPA2* et *SPTLC3*) (Metsalu et al., 2014). Une autre étude menée chez des souris issues de croisements entre des souches distinctes afin de générer des polymorphismes, a permis l’identification de 5 nouveaux GSE dans le placenta murin (*Pde10*, *Phf17*, *Phactr2*, *Zfp64*, et *Htra3*) (Wang et al., 2011). Même si le RNA-seq a été efficace pour mettre en évidence de nouveaux GSE, cette efficacité est inférieure à ce qu’il en était attendu. Ceci est essentiellement dû à une moins bonne précision de quantification allélique et aux erreurs de séquençage générant respectivement des faux négatifs et faux positifs. De plus, des disparités existent entre les différentes technologies de séquençage massivement parallèle qui existent sur le marché. Une publication de Wang et al. récapitule les différents aspects à prendre en compte lors d’une analyse globale utilisant le RNA-seq, notamment lors de la construction de la librairie à séquencer, de l’alignement sur



le génome de référence et de l'analyse bioinformatique des résultats (Wang and Clark, 2014).

### 3.2.2. Analyse de l'épigénome

Le but de cette approche est d'identifier des régions portant des marques épigénétiques pathognomoniques de l'empreinte, essentiellement la présence de DMR et de modifications particulières des histones (Fang et al., 2012; Henckel and Arnaud, 2010; Hiura et al., 2010; Maeda and Hayashizaki, 2006). L'avantage de ces marques est qu'elles sont présentes indépendamment du stade du développement ou du profil transcriptionnel du tissu étudié, mais l'inconvénient est qu'elles peuvent identifier aussi bien un GSE isolé qu'une région chromosomique contrôlant un cluster de GSE pouvant s'étendre sur quelques mégabases, ou même encore une marque épigénétique de la différenciation sans lien avec l'empreinte.

Cette approche pourra également bénéficier des évolutions technologiques récentes avec la possibilité de combiner les moyens d'étude de la méthylation de l'ADN, enzymes méthyl-sensibles ou traitement au bisulfite, et le séquençage massivement parallèle (BS-seq), ainsi que la combinaison des techniques d'étude de modifications des histones (*Chromatin ImmunoPrecipitation*, ChIP) et le séquençage (ChIP-seq).

### 3.3. L'approche *in silico*

L'enrichissement du groupe de GSE avec les nombreuses découvertes a permis une analyse comparative de leurs séquences pour identifier des caractéristiques communes grâce à un apprentissage par le système informatique. Luedi et al. ont ainsi pu mettre en évidence des éléments discriminant des GSE tels que : (1) la sous-représentation de séquences répétées type SINE (*Short Interspersed Nuclear Elements*), et leur orientation positive par rapport à celle du gène d'intérêt, (2) la présence d'éléments répétés type ERV (*Endogenous Retrovirus Elements*) et LINE (*Long Interspersed Nuclear Elements*) essentiellement dans les introns, et (3) la présence de sites de fixation de certains facteurs de transcription plus que d'autres (Luedi et al., 2007; Luedi et al., 2005).

Bien que cette approche présente l'intérêt d'être plus facile à mettre en œuvre que l'expérimentation, elle est très peu efficace puisque très peu de gènes parmi ceux ayant été prédits (quelques centaines chez l'homme et la souris) ont pu être validés comme étant des

GSE (Luedi et al., 2007; Ruf et al., 2007). Par ailleurs, ces études étaient restreintes aux gènes codants alors qu'on sait aujourd'hui que de nombreux GSE ne sont pas codants (Tableau 3).

**Tableau 3 :** Résumé des avantages et des inconvénients des principales approches d'identification de nouveaux GSE.

Type d'approche	Avantages	Inconvénients
<b>Approche clinique et expérimentale</b>	Très efficace	- N'identifie que les régions associées à d'importantes anomalies développementales - Identifie de grandes régions et non des GSE
<b>Approche globale</b>	Techniques relativement exhaustives et résolutive	
- <b>Transcriptome</b>	- Identifie des GSE potentiels et non des régions	- Le résultat dépend du type et stade du tissu analysé car basé sur le profil d'expression, donc peut ne pas détecter les empreintes dynamiques - Détecte tous les gènes à expression monoallélique y compris aléatoire et non liée à l'empreinte
- <b>Epigénome</b>	- Ne dépend pas du type ou stade du tissu analysé car cible des marques épigénétiques stables et non un profil d'expression	- Identifie soit des GSE potentiels soit de régions correspondant à des clusters de GSE (donc moins résolutive)
<b>Approche <i>in silico</i></b>	- Exhaustive - Facile - Peu coûteuse	- Nombreux faux positifs - Basée sur des caractéristiques définies à partir de GSE connus, donc ne détectera pas des GSE avec des caractéristiques nouvelles

### Chapitre III : Fonctions et altérations de l'empreinte parentale

Dès la découverte des GSE, des évidences ont émergé sur leur implication dans la croissance foeto-placentaire et les échanges nutritionnels foeto-maternels. Depuis, d'autres fonctions ont été mises en évidence dans le métabolisme via la régulation de l'homéostasie énergétique et la croissance en période post-natale précoce, ainsi que le développement d'organes dont ceux du système endocrinien tels que le pancréas, le foie, l'hypophyse, l'hypothalamus, le tissu adipeux et le placenta (Lambertini et al., 2012). L'empreinte

parentale a également été associée au contrôle du comportement de l'individu. Sa perturbation est associée à diverses pathologies neurodéveloppementales et endocrines.

Des auteurs ont proposé une classification fonctionnelle des GSE en trois classes (Charalambous et al., 2007) :

- (1) Les gènes soumis à empreinte placentaire, exprimés dans le placenta et ayant une action au niveau de cet organe
- (2) Les GSE non associés à la croissance fœtale mais qui régulent le métabolisme en post-natal
- (3) Les GSE exprimés dans l'embryon et le placenta, impliqués à la fois dans la croissance fœtale et le développement de certains organes

## **1. L'empreinte parentale et le développement placentaire**

### **1.1. Le placenta**

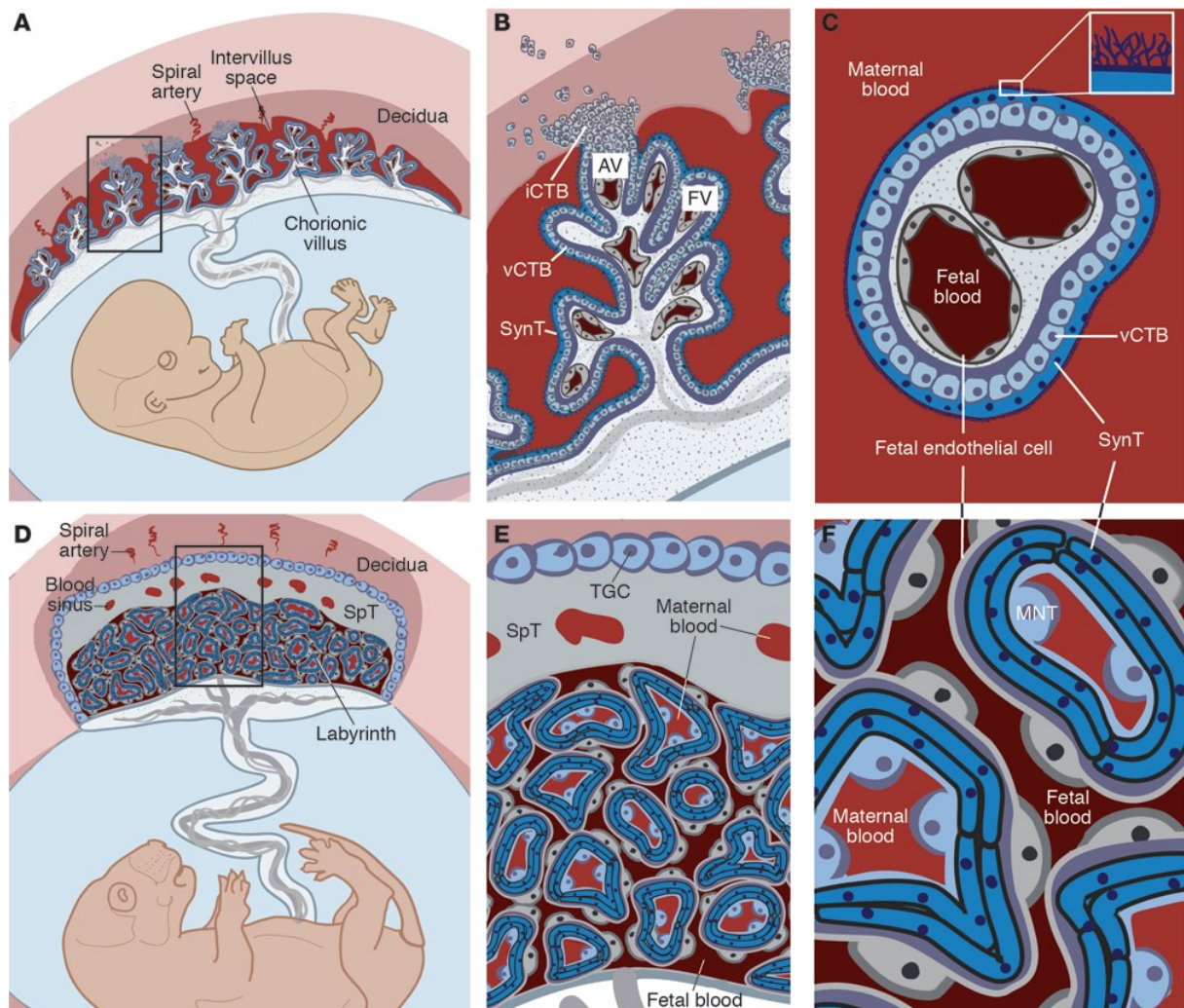
Le placenta est un organe transitoire indispensable au bon déroulement de la gestation. Il fait partie des annexes extra-embryonnaires et est largement composé de tissu fœtal assurant une intime liaison entre le sang fœtal et le sang maternel sans que ces derniers ne soient jamais en contact direct. Chez l'humain, il est de forme discoïde, décidual car il a une composante maternelle, la décidue (qui se forme pendant la gestation et qui est expulsée à la délivrance) et hémochorial (c'est-à-dire que les structures fœtales baignent dans le sang maternel). La structure et la taille placentaires évoluent au cours de la grossesse. Initialement, le trophoblaste embryonnaire pénètre dans l'endomètre maternel puis se différencie en deux couches superposées : la couche interne formée de cytotrophoblastes et la couche externe de syncytiotrophoblastes composée d'un syncytium issu de la fusion des cytotrophoblastes sous-jacents. La structure du placenta évolue pour former un réseau de villosités choriales de plus en plus arborisées, qui résultent de l'invasion du syncytiotrophoblaste par le cytotrophoblaste ce qui produit des prolongements qui s'enfoncent dans l'endomètre maternel. Dans l'axe mésenchymateux des villosités choriales se développe ensuite la vascularisation fœtale artérielle et veineuse issues des artères et veines ombilicales.

Le placenta a plusieurs missions :

- (1) assurer les échanges foeto-maternels gazeux. L'O<sub>2</sub> passe de la circulation maternelle à la circulation fœtale et le CO<sub>2</sub> fait le chemin inverse
- (2) assurer l'apport nutritif au fœtus (eau, électrolytes, glucides, protéines (peptides et acides aminés après dégradation), lipides et vitamines), pour coordonner sa croissance
- (3) assurer la protection immunologique du fœtus par l'apport d'Immunoglobulines G maternelles et l'effet de barrière contre certains agents infectieux (Bacille de Koch ou certains parasites en fonction de l'âge gestationnel), et une protection contre certains métaux
- (4) le placenta a également des fonctions endocrines puisqu'il synthétise et sécrète des hormones protéiques telles que l'Hormone Chorionique Gonadotrophique et des hormones stéroïdiennes telles que la Progestérone et les Œstrogènes, en plus d'autres protéines et enzymes
- (5) enfin, il est impliqué dans le développement du cerveau en synchronisant l'expression des GSE et en assurant la sécrétion de la sérotonine nécessaire à la différenciation des neurones (Lambertini et al., 2012).

#### *Le placenta humain versus le placenta murin*

Les placentas humain et murin sont dits hémochoriaux c'est-à-dire que les villosités choriales fœtales baignent directement dans le sang maternel. Le placenta humain est composé d'une face fœtale dite la plaque choriale et d'une face maternelle dite la plaque basale (Figure 17 A, B et C). Chez la souris, on distingue trois couches du côté fœtal vers le côté maternel (Figure 17) : (1) le labyrinthe où le sang fœtal est séparé du sang maternel par une couche de cellules endothéliales fœtales adossées à deux couches de syncytiotrophoblastes elles mêmes adjacentes à une couche de cellules trophoblastiques mononuclées (Figure 17 D et F), (2) la zone jonctionnelle ou spongiotrophoblaste, composée de cellules trophoblastiques extraembryonnaires de différents types (cellules géantes, cellules spongiotrophoblastiques et cellules à Glycogène) (Figure 17 D et E), et (3) la décidue maternelle (Figure 17 D). Le labyrinthe assure les fonctions d'échanges foeto-maternels. Il est l'équivalent de la plaque choriale du placenta humain. La fonction de la zone jonctionnelle reste quant à elle inconnue. D'un point de vue structural, elle correspondrait à la plaque basale chez le placenta humain.



**Figure 17 :** Comparaison des anatomies des placentas humain (A, B et C) et murin (D, E et F) (Maltepe et al., 2010). AV : Anchoring villi, FV : Floating villi, ICTB : Invasive cytotrophoblast, vCTB : Villous cytotrophoblast, SynT : Syncytiotrophoblast, TGC : Trophoblastic giant cell, SpT : Spongiotrophoblast, MNT : Mononuclear trophoblast.

## 1.2. L’empreinte parentale et le placenta

La plupart des GSE sont exprimés dans le placenta et nombre d’entre eux sont soumis à l’empreinte uniquement dans le placenta. De plus, en termes d’évolution l’apparition de l’empreinte est concomitante avec la placentation.

L’importance des GSE dans le placenta est reflétée par leur large expression dans ce tissu (Ferguson-Smith et al., 2006; Hemberger, 2007; Reik and Walter, 2001; Tycko, 2006), leur

expression différentielle dans l'embryon et le placenta par rapport aux tissus adultes (Steinhoff et al., 2009) et l'absence de ce phénomène dans les espèces aplacentaires (Reik et al., 2003; Yokomine et al., 2001; Yokomine et al., 2005).

Le rôle de l'empreinte dans le développement placentaire est soutenu par les défauts de développement du trophoblaste constatés dans les expériences de parthénogenèse. De plus, de nombreux GSE présentent une empreinte et/ou une expression spécifique aux tissus extra-embryonnaires.

Les preuves de l'importance des GSE dans le développement placentaire proviennent du phénotype lié au déficit en certains GSE (Tableau 4). Parmi les 31 loci/clusters de GSE identifiés à ce jour chez la souris, au moins 10 contiennent des GSE avec fonction placentaire (Figure 18) (Varmuza and Miri, 2015).

**Tableau 4 :** Synthèse des données connues sur une liste de GSE impliqués dans la croissance placentaire et/ou fœtale. Ni les variabilités inter-individuelles et inter-tissulaires de l'empreinte, ni les empreintes partielles ne sont détaillées ici.

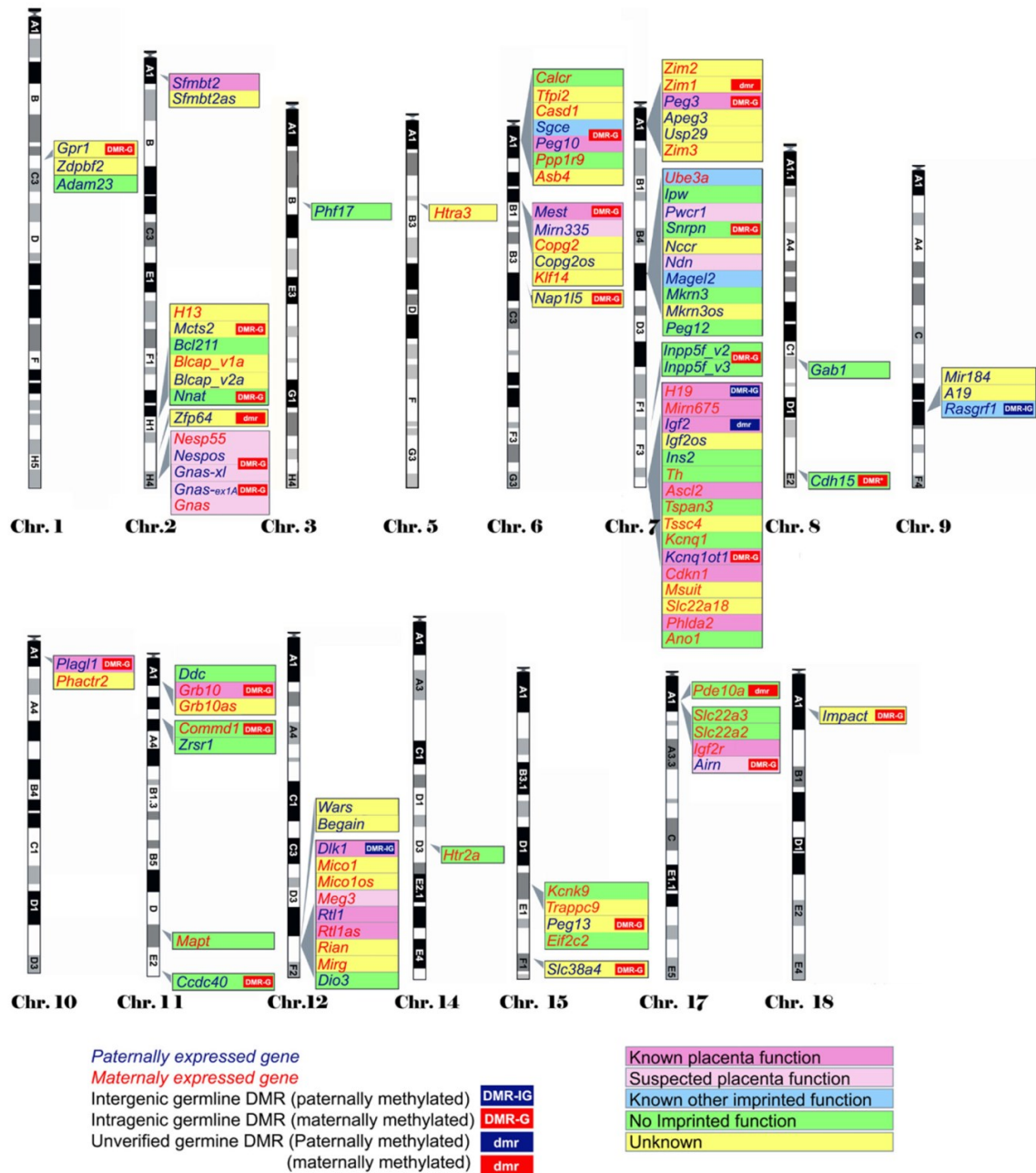
GSE	Expression	Fonction	Modèles murins
<i>ASCL2/Mash2</i>	Maternelle chez la souris Biallélique chez l'Homme. Expression dans le cytotrophoblaste extra-villeux du placenta humain, et le spongiotrophoblaste et le labyrinthe murins	Code pour un facteur de transcription impliqué dans le développement du spongiotrophoblaste placentaire	Knockout : arrêt de développement fœtal léthal en milieu de gestation, <b>défaut placentaire avec absence des cellules du spongiotrophoblaste, un épaissement de la couche de cellules géantes et un défaut de développement du labyrinthe</b>
<i>CDKN1C/p57KIP2</i>	Maternelle. Forte expression dans les placentas humain et murin notamment le labyrinthe	Code pour une protéine du cycle cellulaire, qui régule négativement la prolifération cellulaire	Knockout : létalité embryonnaire, <b>retard de croissance fœtale, placentomegalie</b> et pré-éclampsie maternelle
<i>DLK1</i>	Paternelle	Code pour une protéine Delta-like à domaine EGF, impliquée dans la différenciation cellulaire	Knockout : <b>couche du labyrinthe réduite et anormalement développée sans réduction du poids placentaire, retard de croissance intra-utérin</b>
<i>GRB10/MEG1</i>	Maternelle	Code pour une protéine à domaine homologue Pleckstrin qui interagit avec des facteurs de croissance	Knockout : <b>excès de croissance placentaire et fœtale</b>
<i>H19</i>	Maternelle. Fortement exprimé dans les placentas humain et murin	Code pour un lncRNA impliqué dans la régulation de l'empreinte au locus <i>IGF2/H19</i>	Knockout : <b>placentomegaly et excès de croissance fœtale</b>

**Tableau 4 (suite) :** Synthèse des données connues sur une liste de GSE impliqués dans la croissance placentaire et/ou fœtale. Ni les variabilités inter-individuelles et inter-tissulaires de l’empreinte, ni les empreintes partielles ne sont détaillées ici.

GSE	Expression	Fonction	Modèles murins
<i>IGF2</i>	Paternelle. Forte expression dans le mésenchyme vasculaire, le labyrinthe et le spongiotrophoblaste chez la souris, et les structures équivalentes chez l’Homme	Code pour un facteur de croissance	Knockout : <b>retard de croissance placentaire et fœtale</b>
<i>IGF2R</i>	Maternelle	Code pour un récepteur de facteur de croissance	Knockout : <b>excès de croissance placentaire et fœtale</b>
<i>KCNQ1OT1</i>	Paternelle	Code pour un lncRNA qui régule l’empreinte au locus <i>CDKN1C/KCNQ1OT1</i>	Knockout : <b>retard de croissance placentaire et fœtale</b>
<i>MEST/PEG1</i>	Paternelle. Expression dans le mésoderme embryonnaire et extra-embryonnaire	Code pour une epoxide hydrolase	Knockout : <b>défaut de croissance placentaire et fœtale</b>
<i>PEG3</i>	Paternelle. Fortement exprimé dans les cellules souches trophoblastiques, le cytotrophoblaste villositaire du placenta humain et le trophoblaste murin	Code pour une protéine en doigt de Zn. Rôle dans la prolifération cellulaire ?	Knockout : <b>défaut de croissance placentaire</b> (baisse de 30% du poids), anomalies comportementales
<i>PEG10</i>	Paternelle. Forte expression dans le cytotrophoblaste villositaire du placenta humain et le placenta murin	Protection contre l’apoptose ?	Knockout : <b>défaut majeur de développement du labyrinthe placentaire et quasi absence du spongiotrophoblaste</b> , arrêt du développement fœtal létal en milieu de gestation
<i>PHLDA2</i>	Maternelle. Forte expression dans les placentas humain et murin notamment le labyrinthe	Code pour une protéine à domaine homologue Pleckstrin	Knockout : <b>placentomégalie isolée</b>
<i>PLAGL1/ZAC1</i>	Paternelle	Code pour une protéine à doigt de Zn et au moins 2 ARNnc. Action anti-proliférative ?	Knockout : <b>réduction du poids placentaire et du développement de la couche du labyrinthe, retard de croissance fœtal</b>
<i>RTL1</i>	Paternelle, exprimé dans les cellules souches trophoblastiques et globalement dans le fœtus et le placenta	Contient plusieurs ARNnc. Rôle probable de protection contre l’apoptose	Knockout : <b>infarctus dans la couche du labyrinthe, et à un retard de croissance intra-utérin</b> Surexpression : placentomégalie avec une croissance fœtale normale

**Tableau 4 (suite) :** Synthèse des données connues sur une liste de GSE impliqués dans la croissance placentaire et/ou fœtale. Ni les variabilités inter-individuelles et inter-tissulaires de l’empreinte, ni les empreintes partielles ne sont détaillées ici.

GSE	Expression	Fonction	Modèles murins
<i>SFMBT2</i>	Paternelle chez la souris, dans tissus extra-embryonnaires Biallélique chez l’Homme	Code pour une protéine du groupe Polycomb nécessaire au maintien du pool de cellules souches trophoblastiques	Knockout : défaut de développement du trophoblaste suivi d’un arrêt de développement fœtal létal en milieu de gestation



**Figure 18 :** Carte des GSE chez la souris avec leurs éventuelles fonctions placentaires (Varmuza and Miri, 2015).



Il faut noter que le placenta n'est pas un organe étudié systématiquement dans les modèles animaux. Il existe probablement des défauts placentaires chez des souris déficientes en d'autres GSE que ceux décrits dans le Tableau 4 mais qui ne sont pas rapportés. Une étude systématique du phénotype placentaire avec une standardisation des pratiques est donc nécessaire.

Les GSE à fonctions placentaires peuvent agir à différents niveaux. Certains participent au développement de la structure placentaire, tels que *Ascl2* et *Peg10* dont la déficience chez la souris s'accompagne d'un manque de cellules spongiotrophoblastiques de la zone jonctionnelle (cf Tableau 4) (Guillemot et al., 1994; Ono et al., 2006). D'autres ont une action sur la croissance placentaire comme *Igf2* et *Phlda2* (Salas et al., 2004), ou sur les fonctions physiologiques du placenta notamment le transport de nutriments, tels que le glucose pour *Igf2* ou les solutés pour *Slc22a2* et *Slc22a3* par exemple (Angiolini et al., 2006; Constancia et al., 2002). *Igf2* régule la croissance placentaire et fœtale en régulant le développement et la perméabilité du placenta. Des travaux visant à étudier les effets de la perte d'expression d'*Igf2* sur les fonctions placentaires ont mis en évidence une baisse drastique des capacités de diffusion de certaines molécules, similaires aux nutriments échangés entre la mère et son fœtus pendant la gestation, à travers les placentas déficients en *Igf2* (Sibley et al., 2004).

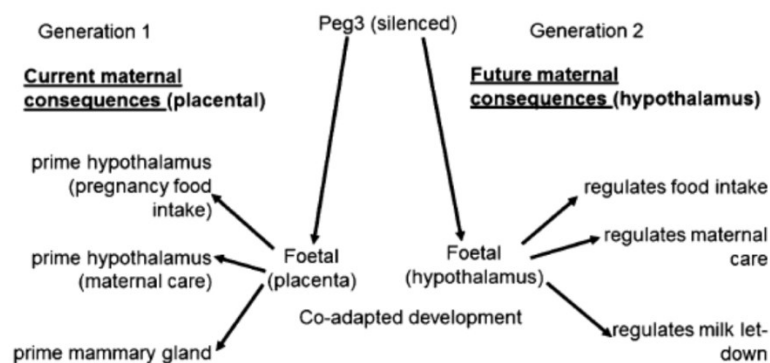
En général, les GSE exprimés dans le placenta auraient un rôle dans le contrôle des apports nutritionnels alors que ceux exprimés dans l'embryon interviendraient dans le contrôle de la demande nutritionnelle en régulant la croissance des tissus fœtaux (Reik et al., 2003). La fonction placentaire des GSE passe par le contrôle de la croissance placentaire globale ou structure-spécifique et la régulation de systèmes de transport des nutriments. Les GSE exprimés dans le placenta uniquement n'interviennent donc pas directement dans la croissance fœtale.

## **2. L'empreinte parentale et le développement embryonnaire**

Plusieurs modèles animaux déficients en différents GSE présentent un défaut de croissance fœtale et néonatale. Parmi eux des GSE qui favorisent la croissance et qui sont essentiellement à expression paternelle tels que *Igf2*, *Prg1*, *Peg3* et *Dlk1*, et des GSE qui

répriment la croissance et à expression maternelle tels que *Igf2r*, *Cdkn1c*, *H19* et *Grb10* (Tableau 4). La déficience en un GSE de chacun de ces deux groupes entraîne respectivement un retard de croissance ou un excès de croissance fœtale. Ces anomalies fœtales s'accompagnent souvent d'anomalies de croissance placentaire rendant difficile l'interprétation des anomalies de croissance fœtale qui pourraient être plutôt la conséquence des défauts placentaires.

Le rôle des GSE pendant l'embryogenèse ne se limite pas à la croissance. D'autres fonctions ont été décrites sur le développement cérébral. Le développement cérébral et le développement placentaire sont liés essentiellement par une communication entre le placenta et l'hypothalamus. Pendant la gestation, une communication s'installe entre le placenta et l'hypothalamus maternel pour assurer les besoins à venir du fœtus. Ceci influence le comportement maternel et les sécrétions hormonales hypothalamiques pour préparer l'accouchement, l'allaitement, etc (Keverne, 2015). De nombreux GSE sont co-exprimés dans le placenta et l'hypothalamus et impliqués dans le comportement et la physiologie maternels (*Grb10*, *Phlda2*, *P57 kip2*, *Dlk*, *Gnas*, *Magel2*, *Necdin*, *Nnat*, *Mest*, *Peg3*, *Hif1α/Arnt* et *Plagl1*). Un défaut de ces GSE impacterait à la fois la gestation en cours par un défaut de réponse hypothalamique maternelle donc un défaut de « préparation à la maternité » mais également la gestation future pour les fœtus féminins via une action sur l'hypothalamus fœtal (Keverne, 2015) (Figure 19). Par exemple, les mutations délétères de *Peg3* ciblées au niveau de l'hypothalamus maternel entraînent un phénotype similaire à celui qu'entraînent les mutations de *Peg3* ciblées dans le placenta. Les deux types de mutations ont pour conséquence un défaut de prise en charge maternelle du nouveau-né notamment un défaut d'allaitement.



**Figure 19 :** Conséquences du knockout de *Peg3* sur le développement foeto-placentaire et hypothalamique illustrant ses fonctions coadaptatives transgénérationnelles (Keverne, 2015).

Des GSE participent également à d'autres processus dans le développement fœtal, à savoir le développement musculaire, le métabolisme et l'homéostasie énergétique.

### **3. L'empreinte parentale et le développement postnatal**

Le rôle de certains GSE dans quelques phénotypes postnataux notamment la croissance, le métabolisme, l'appétit et le comportement est illustré par le phénotype clinique associé à des défauts de l'empreinte dans des syndromes cliniques humains tels que les syndromes de Prader-Willi ou Angelman (cf Section « Pathologies humaines liées à l'empreinte parentale », p. 69).

*Rasgrf1*, par exemple, régule la croissance postnatale sans un effet sur la croissance prénatale. Il n'est d'ailleurs exprimé qu'après la naissance. Les souris déficientes en *Rasgrf1* ont un défaut de croissance postnatale qui persiste toute leur vie alors qu'en période prénatale et néonatale elles ne montrent pas d'anomalie de la croissance. Ces défauts sont associés à une diminution significative des taux du facteur de croissance Igf1 (Itier et al., 1998).

Des GSE sont également impliqués dans le métabolisme et l'homéostasie énergétique. La déficience en *Grb10* s'accompagne d'une meilleure tolérance du glucose, par une augmentation de la sensibilité à l'insuline, une augmentation de la masse musculaire et une diminution de la masse adipeuse, chez les souris adultes (Smith et al., 2007). En plus de jouer un rôle dans la croissance postnatale, *Rasgrf1* est également impliqué dans la prolifération des cellules bêta pancréatiques productrices de l'insuline ainsi que dans le métabolisme lipidique. Les souris déficientes en ce GSE montrent une intolérance au glucose liée à une hypoinsulinémie et une diminution de la masse grasseuse en lien avec une augmentation du catabolisme lipidique (Font de Mora et al., 2003). La surexpression de *Plagl1* chez la souris entraîne un phénotype similaire à celui du diabète néonatal transitoire (cf Section « Pathologies humaines liées à l'empreinte parentale », p. 69) avec une hyperglycémie néonatale associée à une hypoinsulinémie (Ma et al., 2004).

En outre, des GSE peuvent réguler l'appétit. C'est le cas de *Peg3* dont la perte de fonction chez la souris s'accompagne d'une baisse de l'appétit (Curley et al., 2005), et de *Dlk1* dont la surexpression entraîne une hyperphagie (Charalambous et al., 2007).

Ces diverses fonctions métaboliques impliquent également des gènes non soumis à l’empreinte du réseau de GSE.

L’impact des GSE sur le comportement en société a été suggéré depuis longtemps et de plus en plus d’évidences scientifiques soutiennent ce postulat. Le comportement peut être l’interaction mère-fœtus avec l’exemple de *Peg3* et son rôle dans l’acquisition des fonctions maternelles de prise en charge de l’enfant, mais également le comportement entre individus adultes comme c’est le cas de *Grb10* dont la perte d’expression chez la souris est responsable d’une altération du comportement en société (Garfield et al., 2011).

#### **4. Intérêt évolutif de l’empreinte**

L’empreinte est apparue avec la placentation (mis à part un phénomène particulier chez certaines plantes et des insectes). Seuls les mammifères à placenta et les marsupiaux possèdent des GSE. Les mammifères ovipares tels que les Echidnés n’ont pas de GSE. La différence est que contrairement aux embryons des ovipares, ceux des vivipares peuvent contrôler directement leurs ressources nutritionnelles pendant la gestation en raison du lien qu’ils ont avec la mère qui est le placenta. Le corollaire est que les vivipares ne peuvent se reproduire que par la voie sexuée, la parthénogenèse n’est pas viable.

Du point de vue évolutif, différentes théories ont été suggérées pour expliquer l’intérêt qu’il y a eu à développer ce processus épigénétique chez certaines espèces. Il y en a plus de dix. Aucune n’est parfaite car elles n’expliquent pas toutes les situations mais il se pourrait que pour chaque GSE ou groupe de GSE il y ait eu une raison propre de sélectionner l’empreinte à un moment donné et que finalement toutes les théories se valent. Parmi ces théories évolutives il y a la **théorie du conflit parental** (Moore and Haig, 1991), la **théorie de l’inhibition du trophoblaste** (Varmuza and Mann, 1994), la théorie de la co-adaptation materno-fœtale (Wolf and Hager, 2006) et d’autres (théorie de la défense contre les éléments transposables, la théorie de la complémentation, la théorie de la coordination materno-fœtale, etc (Spencer and Clark, 2014).

##### *La théorie du conflit parental*

Cette théorie est basée sur la différence entre les intérêts paternels et les intérêts maternels pendant la gestation et qui s’opposent d’où le terme de « conflit parental ». Chez les

vivipares, la croissance de l'embryon en développement, conçu à la fois par le génome maternel et paternel, ne dépend plus que des ressources en provenance d'un seul des parents à savoir la mère. Plus les ressources sont importantes, plus importante sera la croissance fœtale. Or, les ressources délivrées au fœtus sont prélevées aux dépens de la mère et si les quantités prélevées sont beaucoup trop importantes cela pourrait devenir très préjudiciable à la mère. Donc, même si les deux parents ont intérêt à ce que le fœtus se développe, la mère a également intérêt à se préserver. L'empreinte serait donc parfaitement adaptée pour équilibrer tous ces intérêts : des GSE s'exprimant uniquement à partir du génome paternel exerceraient un rôle de stimulateurs de la croissance fœtale alors que ceux s'exprimant uniquement à partir du génome maternel tendraient à freiner l'action des premiers pour atteindre un équilibre satisfaisant pour le fœtus et la mère. Ceci est d'autant plus important dans les gestations multiples avec potentiellement différents génomes paternels chez les différents embryons. L'inhibition exercée par la mère permettrait d'équilibrer les ressources pour tous les embryons de la portée.

Cette théorie est soutenue par les fonctions opposées des GSE selon qu'ils soient exprimés à partir de l'allèle maternel ou paternel. Ceci est le cas, par exemple, du gène *Igf2* à expression paternelle et qui stimule la croissance du fœtus et du placenta, alors que *Igf2r* qui est à expression maternelle inhibe la croissance en favorisant la dégradation lysosomale de son ligand Igf2 (Lau et al., 1994). D'autres GSE répondent également à ce schéma.

#### *La théorie de l'inhibition trophoblastique*

Même si la parthénogenèse est impossible chez les vivipares, il est possible d'avoir une activation spontanée d'un ovocyte ou d'un spermatozoïde pour donner une cellule diploïde capable de se diviser et d'amorcer un développement embryonnaire. Le risque pour la mère n'est pas le même que pour le père puisque la première est dotée d'un organe capable d'accueillir l'embryon alors que le second non. La théorie proposée par Varmuza et Mann est que les GSE ont pour mission d'empêcher le développement d'un tel embryon chez la mère (Varmuza and Mann, 1994). Les GSE à expression maternelle vont donc inhiber l'invasion trophoblastique, en activant des gènes qui inhibent le développement placentaire ou en réprimant des gènes qui le stimulent, pour empêcher une invasion trophoblastique pathologique pouvant entraîner une pathologie cancéreuse. L'invasion trophoblastique

serait donc guidée uniquement par les GSE à expression paternelle rendant le développement embryonnaire possible uniquement lorsque les deux génomes parentaux sont représentés après une fécondation.

### *La théorie de la co-adaptation materno-fœtale*

Wolf et Hager ont imaginé un modèle de survenue de l’empreinte suite à une pression de sélection liée à l’aptitude du génome fœtal à s’adapter au génome maternel (Wolf and Hager, 2006). L’allèle exprimé à partir d’un GSE donné dépend donc de sa capacité à interagir au mieux avec le génome maternel et donc à induire une bonne compatibilité entre un trait phénotypique du fœtus et celui de la mère.

## **5. Pathologies humaines liées à l’empreinte parentale**

### **5.1. Pathologies de cause endogène**

#### *5.1.1. Les pathologies de l’empreinte*

Les pathologies de l’empreinte regroupent des maladies congénitales syndromiques associant des troubles de la croissance foetale voire post-natale, des anomalies du développement ainsi que des troubles métaboliques (Tableau 5).

Ces pathologies ont pour cause une altération de l’empreinte au niveau d’un GSE ou d’un cluster de GSE entraînant une déficience de l’allèle fonctionnel ou une expression de l’allèle normalement réprimé. Les causes génétiques sont multiples bien qu’aboutissant au même résultat : (1) une mutation délétère de l’allèle normalement fonctionnel, (2) un réarrangement chromosomique au niveau d’une région soumise à l’empreinte, (3) un défaut de méthylation de l’ICR, ou (4) une disomie uniparentale chromosomique ou segmentaire au niveau d’un chromosome portant une région soumise à l’empreinte.

**Tableau 5 : Principales pathologies de l’empreinte chez l’Homme.**

<b>Pathologie de l’empreinte</b>	<b>Principaux signes cliniques</b>	<b>Locus</b>	<b>Principales causes génétiques et épigénétiques</b>
Syndrome de Silver-Russel OMIM 180860	Retard de croissance pré et post-natal, retard de maturation osseuse, asymétrie corporelle (hémihypotrophie), dysmorphie faciale (face triangulaire, front large et bombant, lèvres fines)	11p15, <i>H19/IGF2</i> Chromosome 7 (7p12, <i>GRB10</i> ; 7q21, <i>PEG10</i> ; 7q31, <i>MEST</i> )	DUP(11)mat dup11p15mat Anomalies de méthylation des ICR DUP(7)mat

**Tableau 5 (suite) : Principales pathologies de l’empreinte chez l’Homme.**

Pathologie de l’empreinte	Principaux signes cliniques	Locus	Principales causes génétiques et épigénétiques
Syndrome de Beckwith-Wiedemann OMIM 130650	Excès de croissance pré et post-natal, organomégalie, macroglossie, prédisposition aux tumeurs, hypoglycémie néonatale, héli-hypertrophie, malformations congénitales (omphalocèle, hernie ombilicale, dysplasie rénale...)	11p15, <i>H19/IGF2</i>	DUP(11)pat dup11(p15)pat Anomalies de méthylation des ICR Mutations <i>CDKN1C</i>
Syndrome de Prader-Willi OMIM 176270	Retard croissance intra-utérin, hypotonie néonatale et difficultés alimentaires, puis hyperphagie et obésité précoce, déficience intellectuelle de sévérité variable, hypogonadisme	15q11-q13, <i>SNRPN/UBE3A</i>	DUP(15)mat, Del15q11q13pat, Anomalies de la méthylation de l’ICR
Syndrome d’Angelman OMIM 105830	Déficience intellectuelle, absence de langage, microcéphalie, ataxie, épilepsie, comportement joyeux, hyperactivité, scoliose	15q11-q13, <i>SNRPN/UBE3A</i>	DUP(15)pat, Del15q11q13mat, Anomalies de la méthylation de l’ICR, Mutations <i>UBE3A</i>
Diabète néonatal transitoire OMIM 601410	RCIU, hyperglycémie néonatale transitoire	6q24, <i>PLAGL1</i>	DUP(6)pat, Dup6q24pat, Anomalies de la méthylation
Pseudohypoparathyroïdie OMIM 103580 (type IA), 603233 (type IB), 612462 (type IC)	Résistance à la PTH, ostéodystrophie, ossifications sous-cutanées, anomalies de la croissance, anomalies du comportement alimentaire	20q13, <i>GNAS</i>	del20(q13)mat Défaut de méthylation DUP(20)pat Mutations <i>GNAS</i>
Syndrome de Kagami-Ogata ou Disomie Uniparentale paternelle du chromosome 14 OMIM 608149	Polyhydramnios, placentomégalie, excès de poids à la naissance (hypertrophie musculaire), malformations congénitales notamment du thorax	14q32, <i>DLK1/GTL2</i>	DUP(14)pat Mutations et délétions des centres de l’empreinte
Syndrome de Temple ou Disomie Uniparentale maternelle du chromosome 14 OMIM 616222	Retard de croissance pré- et post-natal, hypotonie néonatale, difficultés alimentaires pendant l’enfance puis obésité, puberté précoce (Prader-Willi-like)	14q32, <i>DLK1/GTL2</i>	del14(q32)pat Défauts de méthylation DUP(14)mat

### 5.1.2. Les pathologies de la grossesse

Deux types de pathologies de la grossesse ont été liés à des anomalies de l’empreinte : le retard de croissance intra-utérin (RCIU) et la pré-éclampsie (PE). Le RCIU inclut les situations prénatales de poids inférieur à celui attendu, selon les courbes de référence, et les cas

d'hypotrophie pathologique avec un poids de naissance trop faible pour l'âge gestationnel (en France, < 10<sup>ème</sup> percentile). La PE associe une hypertension gestationnelle (pressions artérielles systolique et diastolique  $\geq$  140 mmHg et 90 mmHg respectivement), une protéinurie ( $\geq$  0,30g/24 h) et des oedèmes. Ces pathologies sont associées à un fort taux de morbidité pré et postnatale.

En dehors des RCIU syndromiques qui font partie des pathologies de l'empreinte décrites dans la précédente section, les RCIU non syndromiques sont généralement associés à une insuffisance utéro-placentaire qui complique les apports nutritionnels au fœtus dont les causes restent peu connues. Étant donné le rôle des GSE dans les fonctions placentaires, des équipes ont comparé l'expression des GSE dans des placentas issus de grossesses avec RCIU et des placentas de grossesses normales (Diplas et al., 2009; Guo et al., 2008; Koukoura et al., 2011; Kumar et al., 2012; McMinn et al., 2006). Ces études ont révélé une expression différentielle de certains GSE entre les deux groupes de placentas avec une surexpression de GSE tels que *PHLDA2* et une sous-expression d'autres GSE tels que *PLAGL1*. Ces défauts d'expression s'accompagnent d'une perte de l'empreinte et d'une altération de la méthylation dans certains échantillons et perturbent l'expression de tout un réseau de gènes impliqués dans la croissance et le métabolisme fœtaux (Iglesias-Platas et al., 2014). Les mécanismes moléculaires restent encore à élucider.

La PE a également été associée à des anomalies de GSE. Les souris déficientes en *Cdkn1c* développent un phénotype équivalent à la pré-éclampsie humaine (Kanayama et al., 2002). Le GSE *CTNNA3* a été relié à la pré-éclampsie (van Dijk et al., 2004) mais ce résultat n'a pas été reproduit par d'autres équipes. La sous-expression de *H19* associée à une hyperméthylation de son promoteur est elle positivement corrélée à la pré-éclampsie (Gao et al., 2011; Yu et al., 2009).

## **5.2. Pathologies de cause exogène**

Les marques épigénétiques notamment la méthylation de l'ADN, et a fortiori l'empreinte parentale, sont sensibles aux facteurs environnementaux. Des données scientifiques soutiennent des effets de l'Assistance Médicale à la Procréation (AMP), de l'alcool, de la nutrition maternelle, du tabac et probablement aussi des perturbateurs endocriniens, sur l'empreinte.



### 5.2.1. L'assistance médicale à la procréation

De plus en plus d'études démontrent une augmentation de pathologies de la grossesse, de la croissance fœtale, des dysfonctionnements placentaires et même un risque accru de pathologies de l'empreinte en lien avec l'utilisation de techniques d'AMP (Choux et al., 2015). De façon intéressante, les pathologies de l'empreinte qui ont été décrites dans ce contexte, Syndromes de Beckwith-Wiedemann, Silver-Russel et Angelman, avaient pour cause des défauts de méthylation des DMR au niveau des GSE responsables de chacune de ces pathologies (DeBaun et al., 2003; Gicquel et al., 2003; Kagami et al., 2007; Katari et al., 2009). De multiples facteurs peuvent perturber l'empreinte dans le cas de l'AMP : le contexte d'infertilité lui-même, l'hyperstimulation ovarienne, la manipulation des gamètes et les conditions de culture embryonnaire. Ces facteurs interviennent à des moments clés du cycle de l'empreinte mais la part de chacun de ces facteurs ou de leur combinaison reste encore à déterminer.

### 5.2.2. L'alcool et le tabac

Le syndrome d'alcoolisme fœtal associe un retard de croissance, des anomalies neuro-développementales et une dysmorphie faciale caractéristique. Ces manifestations cliniques ont été rattachées à des altérations de la méthylation au niveau de GSE dans quelques études (Haycock and Ramsay, 2009; Liang et al., 2014; Masemola et al., 2015; Stouder et al., 2011; Tyler and Allan, 2014).

L'exposition *in utero* au tabac serait associée à diverses pathologies en post-natal. Des études ont montré des altérations de la méthylation chez la descendance en cas de tabagisme maternel mais seul un GSE, *MEG3*, a été impliqué pour l'instant (Joubert et al., 2012; Markunas et al., 2014).

### 5.2.3. La nutrition maternelle

Les premières études épidémiologiques ayant établi un lien entre l'alimentation de la mère pendant la grossesse et les anomalies de l'empreinte chez la descendance portaient sur des populations ayant vécu la famine hollandaise. Heijmans et al. ont observé que les individus conçus pendant la période de la famine en 1944-1945 avaient une hypométhylation du GSE *IGF2* par rapport aux individus du même sexe non exposés à cette famine (Heijmans et al.,

2008). Cette même équipe a ensuite comparé les profils de méthylation de 15 loci impliqués dans la croissance et le métabolisme, entre des individus exposés à la famine en péri-conceptionnel et les membres de leur fratrie du même sexe non exposés (Tobi et al., 2009). Les auteurs rapportent une différence significative des niveaux de méthylation de GSE à la baisse pour le locus *INS-IGF* et à la hausse pour *ABCA1* et *GNAS-AS*.

#### 5.2.4. *Les perturbateurs endocriniens*

Les perturbateurs endocriniens sont des molécules capables d'interférer avec le fonctionnement du système hormonal. L'un des plus étudiés, et qui était il y a peu de temps encore très répandu dans les produits d'usage courant, est le Bisphénol A. L'exposition à ce composé a été associée à une altération de la méthylation de l'ADN (Mileva et al., 2014). Des études récentes chez la souris ou le rat montrent qu'il perturberait également l'empreinte (Mao et al., 2015; Susiarjo et al., 2013).

# **LE PROJET DE LA THESE**

## LE PROJET DE LA THESE

### I. Contexte du projet

L'équipe de Daniel Vaiman s'intéresse depuis plusieurs années aux perturbations génétiques et épigénétiques associées aux pathologies de la grossesse (notamment la pré-éclampsie et le retard de croissance intra-utérin) et de la reproduction.

En 2010, l'équipe a mené une étude systématique visant à rechercher de nouveaux GSE en utilisant une stratégie de génotypage pangénomique. L'outil utilisé était des puces de génotypage de l'ADN Affymetrix 250K. La méthode consistait à génotyper parallèlement l'ADN génomique (ADNg) et l'ADN complémentaire (ADNc) issus de placentas humains. Les résultats du génotypage de l'ADNg ont servi à identifier les SNP « hétérozygotes » informatifs, et ceux du génotypage de l'ADNc ont permis d'identifier les gènes à expression monoallélique donc pouvant être soumis à l'empreinte parentale. Un certain nombre de critères ont permis d'établir une liste restreinte de gènes candidats dont l'expression monoallélique était ensuite vérifiée par séquençage Sanger sur un nombre d'échantillons plus important et lorsque cette première vérification était concordante l'origine parentale de l'allèle exprimé était recherchée.

#### *Les échantillons utilisés*

Les placentas humains ont été collectés pour des grossesses normales et pathologiques dans le cadre d'une étude approuvée par le Comité de Protection des Personnes, en collaboration avec la maternité de Port-Royal. Seuls les accouchements par césarienne ont été inclus pour s'affranchir de la variabilité des profils transcriptionnels liés à la variabilité des conditions de travail interindividuel. Les césariennes dans le groupe contrôle étaient réalisées essentiellement pour bassin étroit ou utérus cicatriciel. Des échantillons ont donc été prélevés pour extraire l'ADN, l'ARN et certains pour faire des coupes histologiques. Le recueil était effectué dans la demi heure suivant l'accouchement pour préserver la qualité de l'ARN.

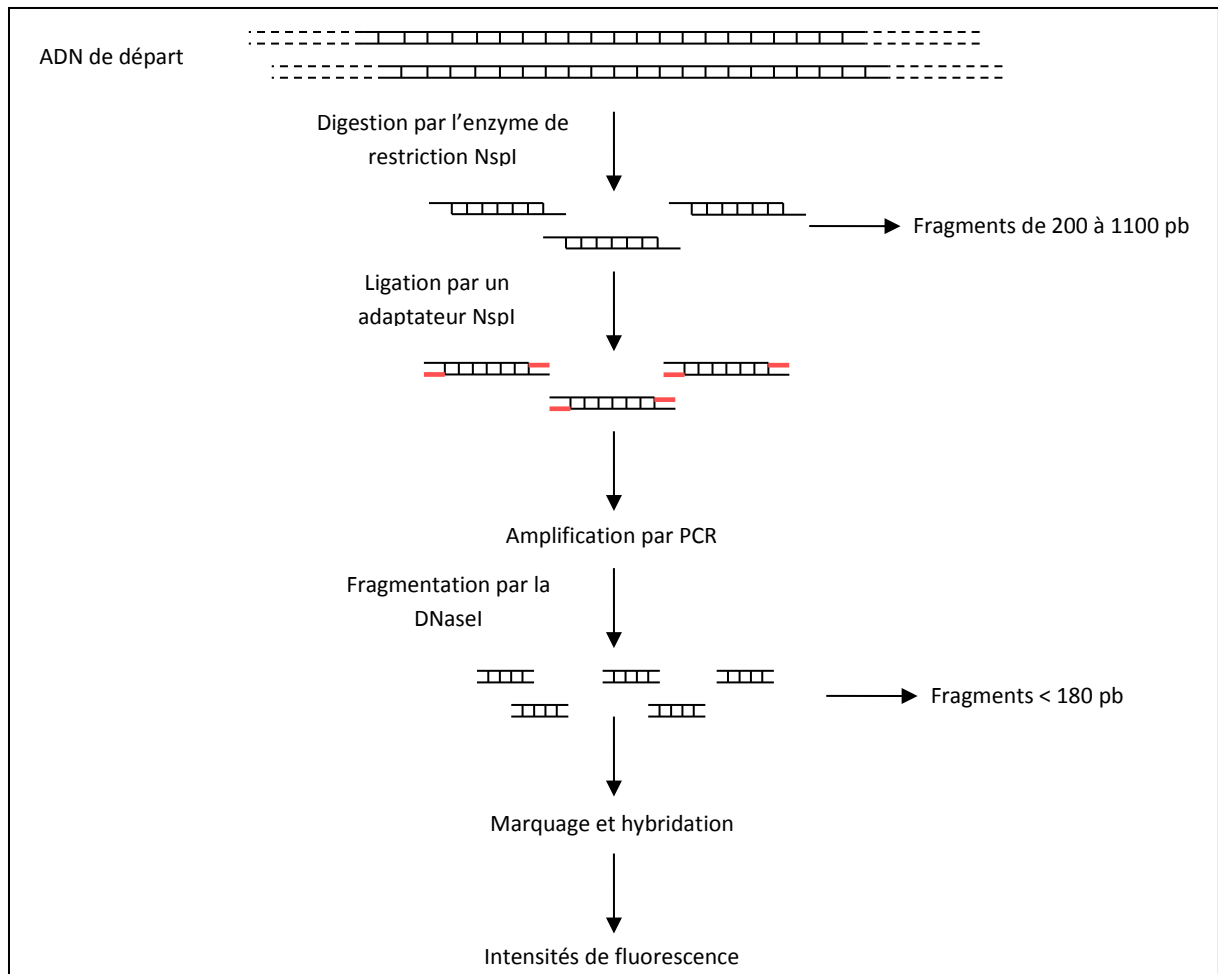
Le sang de la mère a également été prélevé lorsque les patientes l'acceptaient afin d'en extraire l'ADN qui permettait d'étudier l'origine parentale de l'allèle exprimé.

Cette banque d'échantillons est en constante évolution malgré les difficultés de recrutement d'ordre organisationnel (horaire d'accouchement, disponibilité du personnel,

etc) ou religieux ou éthique avec refus des patientes. Néanmoins, nous avons pu inclure plus d'une cinquantaine d'individus dont une trentaine pour lesquelles l'ADN maternel est également disponible. Ce qui permet une analyse de l'origine parentale dans ces derniers.

#### *La technique de génotypage et l'algorithme de sélection des gènes candidats*

Les puces utilisées étaient des Affymetrix SNP 250K Nspl, dont les principales étapes techniques sont représentées dans la Figure 20. Ces puces se présentent sous forme d'une lame en verre d'un peu plus d'un cm<sup>2</sup>, sur laquelle sont greffés des oligonucléotides de 25 bases synthétisés *in situ*. 262000 SNP du génome humain y sont représentés. Chaque SNP est représenté par 56 sondes au maximum ; 28 sondes pour le brin sens et/ou 28 pour le brin anti-sens. Les 28 sondes sont organisées en quartets. Au final, 7 quartets correspondent à chaque brin d'ADN étudié. Ces quartets représentent chacun une sonde qui est centrée soit sur le SNP d'intérêt soit sur un des nucléotides de part ou d'autre de ce SNP. Un quartet est composé de 4 sondes : 1 sonde Perfect Match pour l'allèle A (PMA), 1 sonde MisMatch pour l'allèle A (MMA), 1 sonde Perfect Match pour l'allèle B (PMB) et 1 sonde MisMatch pour l'allèle B (MMB). Un exemple est présenté dans la Figure 21. Un algorithme du fabricant permet de déduire le résultat du génotype pour chaque SNP à partir des données de fluorescence correspondant à chaque sonde et attribue un score qui traduit la qualité et la fiabilité du résultat. Plus le score est faible, plus le résultat est fiable.



**Figure 20 :** Les différentes étapes techniques des puces SNP Affymetrix 250 K NspI.

A

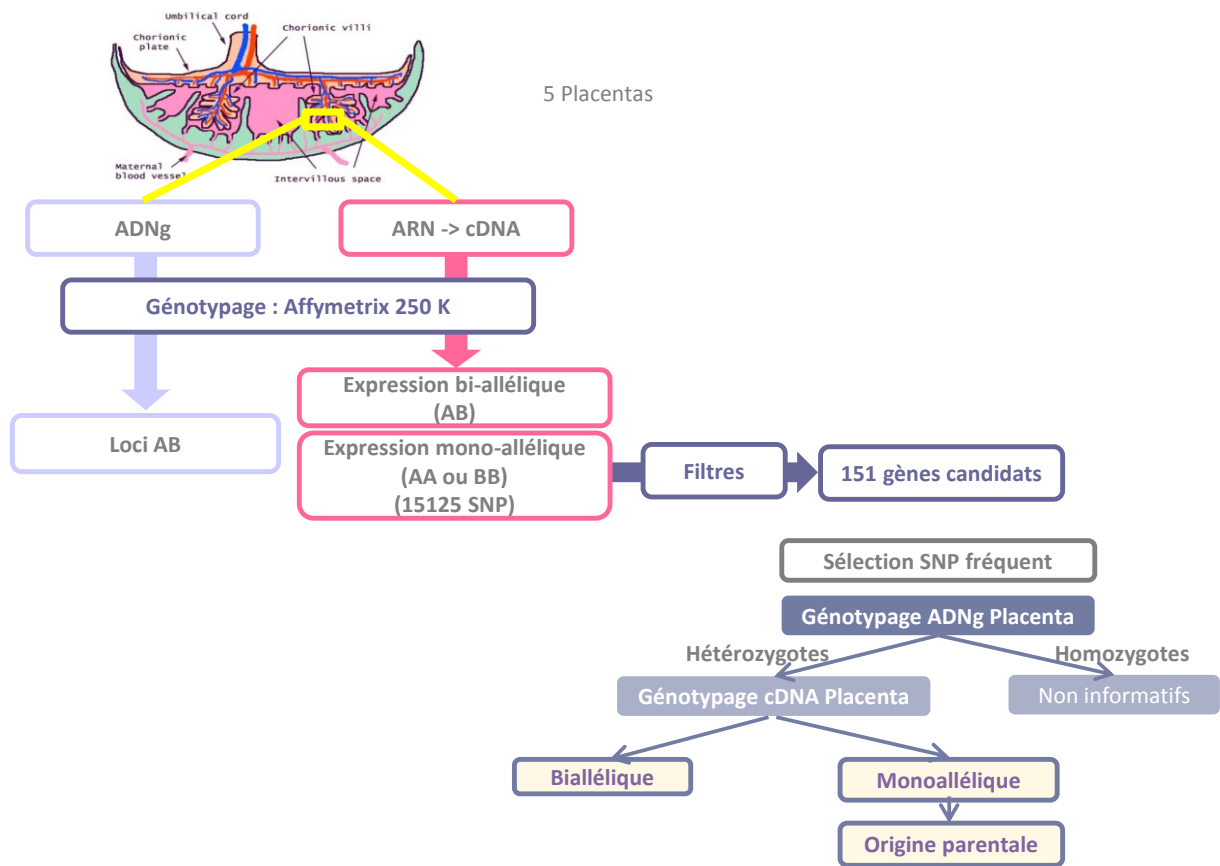


B

-3	-2	-1	0	+1	+2	+3
PMA	PMA	PMA	PMA	PMA	PMA	PMA
<b>MMA</b>	<b>MMA</b>	<b>MMA</b>	<b>MMA</b>	<b>MMA</b>	<b>MMA</b>	<b>MMA</b>
PMB	PMB	PMB	PMB	PMB	PMB	PMB
<b>MMB</b>	<b>MMB</b>	<b>MMB</b>	<b>MMB</b>	<b>MMB</b>	<b>MMB</b>	<b>MMB</b>

**Figure 21 :** Les différents types de sondes utilisées dans les puces Affymetrix SNP. A : Exemples de séquences de sondes ciblant un SNP et composant le quartet. B : Disposition des 28 sondes ciblant un SNP. Il y a 7 quartets de 4 sondes (PMA, MMA, PMB et MMB) chacun, centrées autour d'une position nucléotidique donnée par rapport au SNP d'intérêt (-3, -2, -1, 0, +1, +2, +3).

Ces puces sont destinées au génotypage de l'ADNg. Ici, leur utilisation a été détournée pour servir au génotypage de l'ADNc également (Figure 22). Cinq placentas ont été génotypés pour l'ADNg et l'ADNc. Les résultats du génotypage pour chaque SNP, la position de ce dernier ainsi que le score attribué à chaque résultat de génotypage ont été extraits des résultats bruts de chaque puce.



**Figure 22 :** Logigramme de l'étude à la recherche de gènes candidats soumis à l'empreinte parentale.

Les données brutes d'intensités de fluorescence ont été analysées selon l'algorithme du modèle dynamique (DM) fourni par le fabricant (Di el al., 2005). Les résultats de génotypage pour les différents SNP ont été hiérarchiquement filtrés après transfert dans un tableau Excel suivant un algorithme d'analyse « maison » car il n'en existe pas pour ce type d'analyse puisque les puces ne sont normalement pas utilisées pour cette application et qu'aucune étude équivalente n'a été publiée. Le filtrage hiérarchique a été effectué dans l'ordre qui suit :

1. Selon le score du génotypage : seuls les SNP pour lesquels le score était inférieur à 0,05 à la fois dans l'ADNg et l'ADNc ont été retenus (~150000 SNP).
2. Selon l'informativité : seuls les SNP pour lesquels le génotype de l'ADNg était AB et celui de l'ADNc AA ou BB ont été retenus (15125 SNP).
3. Selon la redondance du résultat entre les échantillons : seuls les SNP montrant un résultat concordant entre au moins 2 placentas ont été retenus (1088 SNP).



4. Selon la position du SNP : seuls les SNP intragéniques ont été retenus (1082 SNP).
5. Selon la redondance du résultat dans le gène : seuls les gènes avec une expression monoallélique répondant aux critères ci-dessus pour au moins deux SNP dans un même gène ont été retenus (151 gènes candidats).
6. Une liste des 18 meilleurs gènes candidats a été sélectionnée en se basant sur le nombre de SNP montrant une expression monoallélique et leurs positions dans le gène. Les gènes avec le plus grand nombre de SNP de préférence situés dans les régions exoniques ont été privilégiés pour cette validation initiale.

Le contrôle positif de l'algorithme établi était la détection de GSE connus tels que IGF2 et l'expression monoallélique des gènes du chromosome X hors les régions pseudoautosomales (PAR) qui sont communes aux chromosomes X et Y et qui échappent à l'inactivation fonctionnelle de l'un des deux X chez les femelles.

#### *La validation des gènes candidats*

Pour chaque gène candidat, un séquençage Sanger a été réalisé à partir de plus de 30 échantillons d'ADNg issus de placentas humains afin d'identifier des individus hétérozygotes pour lesquels un séquençage Sanger de l'ADNc est ensuite réalisé pour vérifier l'expression monoallélique. Pour cette étape, nous avons sélectionné des SNP exoniques et fréquents dans la population caucasienne dans les banques de données publiques. Par la suite, les échantillons d'ADNg issus de la mère ont également été séquencés dans les mêmes conditions pour étudier l'origine parentale de l'allèle exprimé. Ces analyses ont permis de mettre en évidence 7 nouveaux gènes soumis à l'empreinte parentale : *ZFAT*, *ZFAT-AS1*, *GLIS3*, *NTM*, *MAGI2*, *ZC3H12C* et *LIN28B* (Barboux et al., 2012).

## **II. Objectif et organisation du projet**

### **1. Objectif du projet**

Le nombre de GSE identifiés à ce jour représente moins de 1% de la totalité des gènes du génome. Il a été montré que malgré les différents processus biologiques dans lesquels ces gènes sont impliqués, ces derniers interagissent dans le cadre d'un réseau (Varrault et al., 2006). Par ailleurs, il est évident aujourd'hui que ces gènes jouent un rôle majeur dans le développement placentaire et donc foetal en plus d'avoir un rôle direct sur le

développement fœtal. Leur apparition concomitante avec celle du placenta du point de vue évolutif reflète le lien entre le placenta et l’empreinte.

L’identification de nouveaux GSE a un double intérêt :

- (1) du point de vue fondamental, pour mieux comprendre leur régulation épigénétique complexe et les interactions qu’ils ont avec d’autres GSE et gènes non soumis à l’empreinte dans le cadre du réseau de GSE
- (2) d’autre part, une fois identifiés il est possible d’étudier leur rôle potentiel dans certaines pathologies humaines en lien notamment avec la croissance et le comportement.

Nous avons vu dans le chapitre « Méthodes d’identification de nouveaux GSE » les limites de chacune des méthodes utilisées. Le travail préliminaire décrit ci-avant dans le chapitre « Contexte du Projet » montre l’efficacité de l’approche développée par l’équipe pour identifier des GSE puisque 7 gènes sur 18 testés ont été confirmés comme étant des GSE. Le tissu étudié est le placenta humain car c’est celui dans lequel l’empreinte est la plus présente. Ce travail de thèse avait pour but de :

- (1) Poursuivre la validation des GSE candidats non sélectionnés dans la liste initiale de 18 gènes,
- (2) Caractériser la fonction placentaire de ces nouveaux GSE et leur rôle éventuel dans des pathologies de la grossesse liées à l’empreinte.

## **2. Organisation du projet**

Le travail de thèse s’est articulé autour de trois éléments principaux :

- (1) La validation des GSE candidats
- (2) La caractérisation, chez l’humain, d’un des GSE validés
- (3) La caractérisation de ce même GSE dans des modèles animaux et cellulaires

Les résultats des deux premières parties sont détaillés dans l’article prêt à soumettre ci-après. La troisième partie sera exposée dans la section « résultats préliminaires » ci-après.

### III. Résultats

#### 1. Article

L'article présente les résultats de la validation des GSE candidats d'une liste complémentaire (Tableau 6). En effet, en plus des 18 GSE de la liste principale validés dans l'article de Barbaux et al. 2012, nous avons sélectionné 11 autres gènes candidats ayant moins de SNP à expression monoallélique par rapport à la liste principale mais en conservant comme critères : au moins 1 SNP exonique ou 2 SNP hors exons, et une expression documentée dans le placenta.

Sur ces 11 gènes, nous avons validé l'empreinte pour deux gènes, *AIM1* et *DSCAM*. Pour *AIM1*, nous avons des résultats discordants avec des individus montrant une expression monoallélique et d'autres montrant une expression biallélique. L'analyse postérieure des résultats a révélé que l'empreinte est isoforme spécifique pour ce gène. En effet, nous avons choisi 2 SNP pour valider l'empreinte, rs11152999 T>C et rs2297970 G>A. Le premier est spécifique d'un transcrit long et le deuxième est commun aux deux transcrits, long et court . Avec le premier SNP, nous avons seulement deux individus à expression monoallélique mais aucun à expression biallélique. Nous avons donc choisi le 2<sup>ème</sup> SNP pour poursuivre la validation. Deux individus avaient une expression monoallélique et six une expression biallélique. Pendant cette validation, un article a été publié confirmant l'empreinte isoforme-spécifique pour *AIM1* (Das et al., 2013). Nous avons donc jugé préférable d'arrêter les investigations pour ce gène.

Le deuxième GSE identifié est *DSCAM* pour *Down Syndrome Cell Adhesion Molecule*. Nous avons montré que ce gène a une expression paternelle chez tous les individus testés. C'est la première fois que ce gène est identifié comme GSE. Nous avons donc voulu le caractériser d'un point de vue génomique, à savoir identifier la présence d'un élément régulateur de l'empreinte et d'autres GSE potentiels environnants, et d'un point de vue fonctionnel dans le placenta humain.

- (1) Pour répondre à la question sur la régulation épigénétique de l'empreinte, nous avons recherché l'existence d'une DMR intragénique y compris la région promotrice qui est un élément constant caractérisant les GSE (cf Section « Caractéristiques des GSE », p. 46) et qui gouverne l'expression monoallélique. Nous avons ainsi mis en

évidence l'existence d'une DMR dans l'intron 1 et avons montré que cette DMR est méthylée dans les spermatozoïdes humains ce qui est compatible avec l'expression paternelle du gène.

- (2) En outre, nous avons voulu savoir si *DSCAM* fait partie d'un cluster de GSE puisque la majorité des GSE font partie d'un cluster, même si des GSE isolés existent (cf Section « Caractéristiques des GSE », p. 46), et qu'aucun GSE n'a été décrit pour l'instant sur le chromosome 21. Nous avons donc réanalysé les résultats des puces SNP en ciblant l'analyse sur les gènes du chromosome 21 et en baissant le seuil du score à 0,10 au lieu de 0,05 (cf Section « contexte du projet – sélection des gènes candidats », p. 76). Nous avons donc identifié 18 gènes potentiellement soumis à l'empreinte, y compris *DSCAM*, que nous avons testé suivant le protocole de validation des GSE (Tableau 7). La validation de cette série de gènes ciblés sur le chromosome 21 a abouti à l'identification de deux gènes pour lesquels l'expression monoallélique est individu-spécifique (*SAMSN1* et *KRTAP27-1*) et qui méritent d'être confirmée dans d'autres séries pour vérifier s'il s'agit bien d'une empreinte avec une variabilité inter-individuelle ou si ces deux gènes ne sont pas soumis à l'empreinte parentale. Un troisième gène est fortement suspecté d'être soumis à l'empreinte. Il s'agit de *NCAM2* qui appartient à la même famille de protéines que *DSCAM* et pour lequel nous avons observé une expression monoallélique chez 15 individus pour 2 SNP différents (rs232518 C>T et rs2017705 C>T), un seul échantillon de placenta a montré une expression biallélique. Chez les 15 individus à expression monoallélique, nous avons pu étudier la transmission parentale pour 5 d'entre eux. Quatre montrent une expression de l'allèle maternel mais pour un individu l'allèle exprimé n'est pas l'allèle maternel. Nous avons effectué un génotypage par microsatellites des ADNg et ARN placentaires ainsi que de l'ADN maternel pour cet individu afin de vérifier la concordance entre les trois échantillons d'une part, et l'absence de contamination. Les résultats de cette étude n'ont pas montré de discordance inter-échantillons ce qui soutient plutôt une expression monoallélique maternelle avec une variabilité de l'origine parentale de l'allèle exprimé comme cela a déjà été décrit pour un autre GSE (Metsalu et al., 2014).

**Tableau 6** : Liste des GSE candidats identifiés suite à l'analyse initiale des résultats du criblage pangénomique avec les puces Affymetrix 250K Nspl.

GSE candidats	Localisation chromosomique	Nombre SNP à expression monoallélique				Résultat
		Exonique	Intronique	En amont	En aval	
<b>1ère série de validation (Barbaux et al. 2012)</b>						
<i>A2BP1</i>	16		4			Biallélique
<i>ACCN1</i>	17		2			Biallélique
<i>FRMD3</i>	9		3			Biallélique
<i>GLIS3</i>	9		3			GSE
<i>KCNAB1</i>	3		2			Biallélique
<i>LIN28B</i>	6	1	1			GSE
<i>MAGI2</i>	7		3			GSE
<i>NOB1</i>	4	1		1		Biallélique
<i>PLXDC2</i>	10		4			Biallélique
<i>PTPRD</i>	9		2			Biallélique
<i>PTPRG</i>	3			5		Biallélique
<i>ROR1</i>	1		2			Biallélique
<i>STXBP6</i>	14		3			Biallélique
<i>VRK2</i>	2			2		Biallélique
<i>ZC3H12C</i>	11		2	1		GSE
<i>ZFAT</i>	8		6		1	GSE
<i>ZFAT-AS</i>	8					GSE
<i>ZNF597</i>	16	1				GSE
<b>2ème série de validation (travail de la présente thèse)</b>						
<i>AIM1</i>	6		1	1		GSE
<i>CDH11</i>	16		1		3	Biallélique
<i>CSMD1</i>	8		3			Biallélique
<i>CTNND2</i>	5		2			Absence d'amplification de l'ADNc
<i>DSCAM</i>	21		2			GSE
<i>DSCAM-AS</i>	21					Absence d'amplification de l'ADNc
<i>DSCAM-IT</i>	21					Absence d'amplification de l'ADNc
<i>FAM5C</i>	1				6	Biallélique
<i>LRRC46</i>	17	1				Biallélique
<i>NPAS3</i>	14		2			Biallélique
<i>PCDH7</i>	4			2		Biallélique
<i>ROBO2</i>	3		2			Biallélique
<i>KAZN</i>	1		1	1		Pas d'individus hétérozygotes
<b>Gènes non validés</b>						
<i>ADAMTS20</i>	12				2	
<i>ARL6IP1</i>	11			2	1	
<i>ARMC1</i>	4		1		1	
<i>BIN1</i>	2		1		1	
<i>C16orf82</i>	16			2		
<i>C22orf34</i>	22				3	
<i>C2orf34</i>	2		2			
<i>C2orf50</i>	2			2		
<i>C5orf15</i>	5				2	
<i>C6orf118</i>	6				3	
<i>C6orf170</i>	6				2	
<i>C8orf71</i>	8			2		
<i>CADM2</i>	3		1	1		
<i>CCDC149</i>	4		2			
<i>CCDC34</i>	11				2	

**Tableau 6 (suite) :** Liste des GSE candidats identifiés suite à l'analyse initiale des résultats du criblage pangénomique avec les puces Affymetrix 250K Nspl.

GSE candidats	Localisation chromosomique	Nombre SNP à expression monoallélique				Résultat
		Exonique	Intronique	En amont	En aval	
<b>Gènes non validés</b>						
CCDC59	12				3	
CD83	6			2		
CDH6	5			2		
CDH7	18			2		
CNTN5	11		1	1		
CNTNAP2	7		2			
CNTNAP5	2		1	1		
CPEB2	4			2		
CTSC	11				2	
DCLK1	13		3			
DEC1	9		1	1		
DNAH8	6		2			
DOCK4	7		1		1	
DSC3	18			3		
EFNA5	5		1		1	
EN1	2				2	
EPHA4	2		1		1	
EPHA7	6				2	
FAM169B	15				2	
FAM19A5	22			3		
FAM46A	6				2	
FMN2	1		1	1		
G2E3	14			2		
GALNTL6	4			2		
GLS	2			2		
GRHL3	1			2		
GRIN2A	16		2			
GRM3	7			2		
GRM7	3			3		
H2AFY2	10		1	1		
hCG_1745121	7		1		2	
HNRNPK	3			2		
IL20RA	6		1		1	
IPO8	12				3	
IRX5	16			2		
KC6	18				2	
KCNA4	11				2	
KCNMA1	10		3			
KCNMB2	3		1	1		
KIAA0774	13		2			
KLF12	13		1		1	
KRT3	12	1				
LMCD1	3		1	1		
LOC100129055	10	1				
LOC283050	10				2	
LOC388965/CTNNA2	2			2		
LOC642597	18		1		1	
LOC647979	20	1				

**Tableau 6 (suite) :** Liste des GSE candidats identifiés suite à l'analyse initiale des résultats du criblage pangénomique avec les puces Affymetrix 250K Nspl.

GSE candidats	Localisation chromosomique	Nombre SNP à expression monoallélique				Résultat
		Exonique	Intronique	En amont	En aval	
<b>Gènes non validés</b>						
<i>LPHN3</i>	4			2		
<i>LRRIQ3</i>	1				3	
<i>LRRN1</i>	3			2		
<i>MAP2</i>	2			2		
<i>MCPH1</i>	8			3		
<i>MGC45800</i>	4			5		
<i>MGMT</i>	10		1	2		
<i>MON1B</i>	16			2		
<i>MS4A6E</i>	11	1		1		
<i>MUC20</i>	3	1				
<i>MUC4</i>	3		1		1	
<i>NALCN</i>	13		1		1	
<i>NCRNA00052</i>	15			2		
<i>NDUFA10</i>	2				2	
<i>NEBL</i>	10				2	
<i>NKAIN2</i>	6		2			
<i>NKAIN3</i>	8		1		2	
<i>NRG1</i>	8		1	1		
<i>NUDT9P1</i>	10				2	
<i>OCA2</i>	15	1				
<i>OR2A5</i>	7	1				
<i>PCDH15</i>	10				2	
<i>PCSK2</i>	20			2		
<i>PFKP</i>	10			2		
<i>PKHD1</i>	6		2			
<i>PLCB1</i>	20		3			
<i>PLSCR2</i>	3				2	
<i>PLXNA4</i>	7				2	
<i>PPAT</i>	3	1				
<i>PPIAL4G</i>	11			3		
<i>PRELID2</i>	5				2	
<i>PRMT6</i>	1			2		
<i>RAB38</i>	11		1		1	
<i>RORB</i>	9			2		
<i>RP11-35N6.1</i>	9		4			
<i>C3orf58</i>	3				3	
<i>LCORL</i>	4			2		
<i>RPL36</i>	9			2		
<i>RPL7</i>	5				2	
<i>RPS2</i>	4				3	
<i>RPS7</i>	17	1				
<i>RRAGC</i>	1				2	
<i>SASH1</i>	6			2		
<i>SFMBT2</i>	10				2	
<i>SGCD</i>	5		1	1		
<i>SGCZ</i>	8		2		1	
<i>SH2D1B</i>	1				2	
<i>SH3GL2</i>	9		2			

**Tableau 6 (suite) :** Liste des GSE candidats identifiés suite à l'analyse initiale des résultats du criblage pangénomique avec les puces Affymetrix 250K Nspl.

GSE candidats	Localisation chromosomique	Nombre SNP à expression monoallélique			Résultat
		Exonique	Intronique	En amont En aval	
<i>SLC35F1</i>	6		3		
<i>SLC9A7</i>	12				2
<i>SMOC2</i>	6		1	1	
<i>SNTG1</i>	8		2		
<i>SOX11</i>	2			2	
<i>SOX9</i>	17			3	
<i>SP3</i>	2				2
<i>SPON1</i>	11		1	1	
<i>SRBD1</i>	2				2
<i>SYNPR</i>	3		2		
<i>TAC1</i>	7			2	
<i>HNRNPK</i>	2	1			
<i>THSD7B</i>	2			2	
<i>TMEM106B</i>	7			2	
<i>TMEM132C</i>	12			3	
<i>TMSL3</i>	2	1			
<i>TRIB2</i>	2			2	
<i>TRPS1</i>	8				2
<i>TSPAN11</i>	12			2	
<i>TYRP1</i>	9			3	
<i>UNC13C</i>	15		1	1	
<i>UNC5CL</i>	6				2
<i>VPS13C</i>	15		1		1
<i>WBSCR28</i>	7	1			
<i>XKR4</i>	8		2		
<i>XPNPEP1</i>	10				3
<i>ZCCHC10</i>	5				5
<i>ZFPM2</i>	8			2	
<i>ZIC4</i>	3				2
<i>ZNF717</i>	3	1			
<i>ZNF804A</i>	2		1	1	
<i>ZWINT</i>	10				2
<i>ZYG11B</i>	1		1	1	



**Tableau 7 :** Liste des GSE candidats identifiés suite à la ré-analyse ciblée sur le chromosome 21 des résultats du criblage pangénomique avec les puces Affymetrix 250K Nspl.

GSE candidats	Coordonnées génomiques (hg38)	Résultat
<i>LIP1</i>	chr21:14,108,893-14,210,891	Absence d'amplification de l'ADNc
<i>SAMSN1</i>	chr21:14,485,228-14,546,343	50 % Bialléliques, 50 % Monoalléliques chez 4 individus
<i>NCAM2</i>	chr21:20,998,315-21,543,329	GSE ?
<i>ADAMTS5</i>	chr21:26,917,912-26,966,513	Biallélique
<i>GRIK1</i>	chr21:29,536,933-29,939,625	Absence d'amplification de l'ADNc
<i>KRTAP27-1</i>	chr21:30,337,013-30,337,694	50 % Bialléliques, 50 % Monoalléliques chez 6 individus
<i>KRTAP13-4</i>	chr21:30,430,230-30,431,026	Problèmes de séquençage de la région
<i>TIAM1</i>	chr21:31,118,416-31,558,977	Biallélique
<i>ERG</i>	chr21:38,380,027-38,661,694	Absence d'amplification de l'ADNc
<i>PSMG1</i>	chr21:39,174,769-39,183,851	Biallélique
<i>BRWD1</i>	chr21:39,184,176-39,313,819	Biallélique
<i>WRB</i>	chr21:39,380,244-39,397,889	Biallélique
<i>B3GALT5</i>	chr21:39,556,442-39,673,137	Problèmes de séquençage de la région
<b><i>DSCAM</i></b>	<b>chr21:40,010,999-40,847,139</b>	<b>GSE</b>
<i>PLAC4</i>	chr21:41,175,231-41,185,239	Biallélique
<i>ZBTB21</i>	chr21:41,986,831-42,010,357	Biallélique
<i>SLC37A1</i>	chr21:42,499,632-42,581,440	Biallélique
<i>PCBP3</i>	chr21:45,643,694-45,942,454	Biallélique

(3) Nous nous sommes également intéressés au profil d'expression de *DSCAM* dans les placentas normaux et pathologiques. L'immunohistochimie sur des coupes de placentas humains a montré une expression localisée au niveau du syncytiotrophoblaste et des cellules endothéliales. L'analyse de l'expression de *DSCAM* dans les placentas pathologiques (issus de grossesses avec RCIU et/ou Pré-éclampsie) n'a pas révélé de différence significative des niveaux d'expression par rapport aux placentas normaux. Ce qui rend peu probable l'implication de *DSCAM* dans ces pathologies mais n'exclut pas un effet sur le placenta en cas de défaut de l'empreinte. L'expression de cette molécule dans le syncytiotrophoblaste nous a mené à nous interroger sur un éventuel rôle dans la syncytialisation qui s'exprimerait par une différence d'expression de *DSCAM* entre le stade cytotrophoblaste et syncytiotrophoblaste. Nous avons donc étudié des cellules trophoblastiques primaires qui ont la capacité de se syncytialiser spontanément après quelques jours de culture. La comparaison des niveaux d'expression de *DSCAM* après 24, 48 et 72 h de culture n'a montré aucune différence avec néanmoins une limite qui est que le niveau d'expression basal est déjà très faible.

# ARTICLE

**Title:** A genome-wide search for new imprinted genes in the human placenta identifies the first imprinted gene on chromosome 21

**Running title:** *DSCAM* on chromosome 21 is imprinted in the human placenta

**Authors:**

Laïla El Khattabi (1, 2, 3, 4), Stéphanie Backer (1, 2, 3), Amélie Pinard (1, 2, 3, \*), Marie-Noëlle Dieudonné (5), Vassilis Tsatsaris (6), Daniel Vaiman (1, 2, 3), Luisa Dandolo (1, 2, 3), Evelyne Bloch-Gallego (1, 2, 3), Hélène Jammes (7), Sandrine Barboux (1, 2, 3)

**Affiliations :**

1. INSERM, U1016, Institut Cochin, Paris, France
2. CNRS UMR8104, Paris, France
3. Université Paris Descartes, Sorbonne Paris Cité, Paris, France
4. Laboratoire de Cytogénétique, Hôpital Cochin, AP-HP, Paris, France
5. Université Versailles-Saint Quentin en Yvelines, France
6. Maternité Cochin-Port Royal, AP-HP, Paris, France
7. INRA, UMR1198 Biologie du Développement et de la Reproduction, Jouy en Josas, France

\* Present address : Aix Marseille Université, Inserm, GMGF UMR\_S 910, 13385, Marseille, France

Correspondence to : Dr Sandrine Barbaux, INSERM U 1016, 24 rue du Faubourg Saint  
Jacques, 75014 Paris, France. E-mail : sandrine.barbaux@inserm.fr

## **Abstract**

We identified, through a genome-wide search for new imprinted genes in the human placenta, *DSCAM* (Down Syndrome Cellular Adhesion Molecule) as a paternally imprinted gene. We identified a Differentially Methylated Region, located within intron 2 and methylated on the maternal allele, compatible with its paternal expression. This gene encodes a cellular adhesion molecule mainly known for its role in neurone development. We found that *DSCAM* is present in endothelial cells and the syncytiotrophoblast layer of the human placenta where its functions remain unclear. This is the first imprinted gene located on human chromosome 21.

## Introduction

Imprinting is a complex heritable epigenetic regulation of gene expression resulting in a monoallelic expression of a set of genes according to the parent-of-origin of the allele. These genes belong to various families including growth factors, transcription factors, enzymes, receptors, and are involved in placental and foetal development. This is supported by the fact that i) nuclear transfer experiments leading to gynogenote or androgenote embryos result in lethal abnormal fetal or placental development [1, 2], ii) abnormal imprinting of some imprinted genes is associated with placental and foetal growth pathologies such as preeclampsia and intra-uterine growth restriction [3], iii) several imprinted genes are imprinted only in placental tissue [4], and iv) placenta and imprinting appearance are concomitant in terms of evolution [5].

About 150 imprinted genes (IGs) were identified in mice and close to 100 in humans ([www.mousebook.org](http://www.mousebook.org), [www.geneimprint.com](http://www.geneimprint.com) and [igc.otago.ac.nz/home.html](http://igc.otago.ac.nz/home.html)). Some of them have been identified following the molecular characterisation of chromosomal rearrangements or uniparental disomies causing clinical syndromes (Prader-Willi syndrome and Beckwith-Wiedeman syndrome, for instance). IGs are interesting for their particular epigenetic regulation and their role in foeto-maternal nutrients exchanges and postnatal offspring and maternal behaviour. Attempts to identify new IGs in humans and mice by computational means [6, 7], global transcriptome or epigenome analysis [8, 9] or using parthenogenetic tissues [3, 4] have been reported, sometimes using putative genomic specificities such as gene organization or presence of repeated elements, to provide lists of candidate genes. However, validation approaches to confirm the imprinted status of these

computed genes have globally failed to identify numerous new IGs [5]. In order to identify new genes that could be imprinted or at least monoallelic in the human placenta, we designed a strategy using high throughput genotyping arrays. A first validation step allowed the characterisation of 7 candidates as new IGs in the human placenta, one of them being shown as also imprinted in mice [10]. Here, we report the results of an extended analysis of our previous data, exploring 30 candidate genes and leading to the discovery of the first imprinted gene located on human chromosome 21, Down Syndrome Cellular Adhesion Molecule or *DSCAM*.

## **Materials and methods**

### Biological samples

Human placentas from pathologic (IUGR, preeclampsia) or normal pregnancies as well as maternal blood were collected at the Port Royal maternity (Paris, France). Placental samples were treated within 30 min after delivery by caesarean section, dissected to remove the maternal side and snap frozen before further extraction. Maternal peripheral white blood cells were isolated and frozen for DNA extraction. Clinical details have already been described elsewhere [11].

Primary cytotrophoblasts were isolated from first trimester legal abortions. Cells were treated and cultured for up to 3 days as described elsewhere [12].

Sperm samples were obtained from normozoospermic men attending the Cochin infertility clinic. After a density gradient, pellets of sperm were performed to extract genomic DNA.

For all samples, patients had given their informed consent before use of their tissues and all protocols had been approved by the local Ethics Committees.

### Nucleic Acids Extraction and Reverse Transcription

Genomic DNA was extracted from cells or tissues according to a previously described protocol [13]. RNA was extracted from human placental tissues or from animal samples using Trizol (Life Technologies) followed by chloroform extraction and precipitation. Concentration and purity were tested on a Nanodrop spectrophotometer (Nanodrop, Thermo Scientific). 2 to 5 µg of RNA were treated with DNase in order to eliminate potential DNA residues before Reverse Transcription. RNA was then exposed to MMLV Reverse Transcriptase in the presence of hexamer random primers.

### PCR and real-time PCR

Endpoint PCRs were performed using the GoTaq Flexi or HotStart GoTaq Flexi (Promega) following the manufacturer's conditions. Real-time RT-PCR was performed on a LightCycler480 apparatus (Roche) using the LightCycler480 SYBr Green I Master Mix (Roche). Primers, applications, and particular conditions are described in Supplemental Table



3. Results were analysed following the  $\Delta\Delta C_t$  method. The statistical analysis was performed using a two-tailed t-Test and considered significant with a p value  $<0.05$ .

#### STR analysis

We performed a DNA genotyping using the PowerPlex® 16 HS multiplex STR System. The analysis was performed on 2 ng of placental genomic DNA, 2 ng of maternal gDNA, and 750 ng of placental RNA. After loci co-amplification, microsatellites repeats alleles were analysed by standard semiautomatic methods on an ABI PRISM 377 automate (Applied Biosystems). Genotyping data were analyzed using the Gene Mapper version 3.5 software package (Applied Biosystems).

#### Methylation level analysis

Genomic DNA was treated with sodium bisulfite in order to modify unmethylated cytosines using the EZ DNA Methylation-Gold Kit (Zymo Research). 80 ng of bisulfite converted DNA were used as templates for PCR amplification in classical conditions. PCR products were either directly sequenced at the Eurofins Cochin platform after purification using the Macherey-Nagel kit, or used as templates for pyrosequencing reactions, or cloned using the TOPO TA cloning kit (Life Technologies) and sequenced. Pyrosequencing was performed using the PyroMark 24 Gold Kit on a PyroMark Q24 System (Qiagen) according to the manufacturer's recommendations. Duplicate measures were obtained on different PCR

products amplified on bisulfite-treated DNA samples performed on 2 occasions. A difference of more than 5% between duplicates was considered uncertain and repeated. Genomic DNAs from 25 placental samples and 6 sperm samples were analysed.

Primers sequences and applications are listed in Supplemental Table 3.

## Cell culture

The Human Umbilical Vein Endothelial Cell (HUVEC) cell line was grown in DMEM medium and the human choriocarcinoma cell line BeWo was cultured in F12K medium, both supplemented with penicillin/streptomycin (plus 0.1% normocin for HUVECs) and 10% Fetal Calf Serum, at 37°C and 5% CO<sub>2</sub>. BeWo cells were treated with 12µM forskolin for 48 h to induce syncytialisation .

## Immunohistochemistry and staining, Western blot

Paraffin-embedded placental sections from term normal pregnancies or from earlier terminated pregnancies were rehydrated and boiled in citrate buffer. They were treated to block endogenous peroxidases with hydrogen peroxide and unspecific binding sites with BSA. The primary antibody NBP1-59208 against DSCAM (Novus Biotechnie) was incubated in PBS-0.1% Tween-1% BSA overnight at 4°C. After washing, revelation was obtained using the Novolink kit (Leica).

Ten to 40  $\mu\text{g}$  of protein lysate from BeWo or HUVEC cells were run on a 6 % polyacrylamide gel and transferred to a PVDF membrane. Revelation was performed with the same antibody than above, an anti-rabbit secondary antibody and the ECL kit (Millipore).

## **Results**

### *Identification of candidate imprinted genes*

We used data from our previously described work based on a whole genome screening strategy using genotyping microarrays [10]. Briefly, this consisted in hybridizing in parallel pairs of 250K genotyping Affymetrix arrays with either genomic DNA or cDNA extracted from five human placentas. Genomic DNA genotyping allows the sorting of informative SNPs (heterozygous) for which we followed the allelic expression based on the cDNA genotyping data (summarized in Supplemental Figure 1). Hence, we were able to identify genes with a putative monoallelic expression. We established a pipeline to select a list of candidate monoallelic/imprinted genes that we previously described. Here, we extended our analysis to an additional list of candidate genes suspected to be monoallelically expressed. After a positive hit on chromosome 21, we specifically reanalysed data for this chromosome using a less stringent filter on the quality score of genotyping results (0.1 instead of 0.05) to extract more candidates.

### *Validation of candidate imprinted genes*

We explored the imprinted status of 30 genes listed in Supplemental Table 1. The validation was carried out by Sanger sequencing of at least one SNP selected to be i) exonic and ii) with a high heterozygotes frequency, in a collection of placental samples.

We analysed the *AIM1* (Absent in Melanoma 1) gene on chromosome 6 using the rs11152999 SNP in exon 1 and could identify 2 heterozygous placentas exhibiting a monoallelic expression. Given the weak frequency of this polymorphism in Caucasian populations, we analysed another SNP, namely rs2297970, in exon 12. Conflicting results included 2 cases of placental monoallelic expression but also 6 cases of biallelic expression. We observed that the first SNP is present only on the long isoform of the gene while the second is located in an exon common to both the long and short isoforms, the 2 isoforms being the result of alternative promoters usage. This result has also been described by [14] and confirms that the imprinting of *AIM1* is isoform-specific and limited to the longest transcript that uses a different upstream promoter.

We analysed the *NCAM2* (Neural Cellular Adhesion Molecule 2) gene on chromosome 21. We used 2 SNPs rs232518 in exon 9 and rs2017705 in exon 16 (Supplemental Table 2). Most of the heterozygous placentas showed a monoallelic expression (n=15) and in 4 informative cases, the expressed allele was the one inherited from the mother. However, one case showed a biallelic expression, on repeated experiments and though the positions of cDNA specific primers could not allow the amplification of potentially contaminating gDNA. Such polymorphic imprinting has already been described for other imprinted genes [15]. Another conflicting case is one showing a monoallelic expression but with a paternal origin of the expressed allele. To exclude a problem regarding the amplified samples, we used an analysis of Short Tandem Repeats to confirm that the RNA sample and the genomic DNA extracted

were indeed from the same individual and that the maternal DNA was compatible with the child DNA. All 3 samples were correct (Supplemental Figure 2). The exact status of the *NCAM2* gene is therefore still under question.

For a few candidates, we could not confirm the monoallelic expression or could not conclude because of lack of informative SNP, insufficient informative subjects or insufficient expression level (Supplemental Table 1).

We concentrated on a promising candidate, named *DSCAM* for Down Syndrome Cellular Adhesion Molecule, located on chromosome 21, at the 21q22 band. We explored the imprinted status of this gene by amplifying around 2 SNPs in exon 5, rs 2297270 and rs 41395652. We observed 10 cases of heterozygous carriers of these polymorphisms that showed a monoallelic expression of the *DSCAM* gene in placental cDNA. In the analysis a third SNP, rs 34336407 in exon 9 of *DSCAM*, 4 additional heterozygotes also exhibited a monoallelic expression in placental cDNA. No case of biallelic expression has been observed among all samples tested (Supplemental Table 2). In 7 cases, the genotyping of maternal genomic DNA was informative to follow the parental transmission of the alleles. In all 7 cases (5 cases for the SNPs in exon 5 and 3 cases for the SNP in exon 9), the expressed allele was the one transmitted by the father (Figure 1). We therefore concluded that the *DSCAM* gene was maternally imprinted in the human placenta.

No expression could be obtained in lymphocytes samples to extend the imprinting analysis.

At the *DSCAM* locus, an antisense gene (*DSCAM-ASI*) composed of 2 or 3 exons is observed as well as a 4 exon intronic transcript (*DSCAM-ITI*) (Figure 2, A). Both overlap the very large intron 3 of *DSCAM* on opposite strands and are non-coding transcripts. Though we tried to determine the imprinted status of *DSCAM-ASI* and *DSCAM-ITI* using rs 2837597 and rs

2837753 as markers respectively, the expression of these genes was too weak in placental tissue and did not allow their amplification and further investigations.

### *Epigenetic regulation of the imprinting of DSCAM*

We studied the methylation of CpG islands within and around *DSCAM* in order to identify a Differentially Methylated Region (DMR) typical of an imprinted gene. An *in silico* analysis of the genic sequence using the CpGplot software predicted the presence of 23 CpG islands (Fig 2, A). We limited our subsequent investigations to the most relevant sites. Hence, we first selected three predicted CpG islands for which we studied the methylation status, one in the promoter region of the long *DSCAM* isoform (named CpG23) and two (named CpG9 and CpG10) that co-localize with both the promoter region of the antisense gene *DSCAM-AS1* situated within the same locus and the promoter region of the shorter *DSCAM* isoform. We first performed Sanger sequencing after bisulfite treatment of these regions. We could not find any differential methylation of the CpG23, CpG9 and CpG10, these being mainly and globally methylated. A genome-wide search for DMR using methylation arrays had noted the presence of such a 500 bp region about 20 kb from CpG island 22, located within the very large intron 1 of the *DSCAM* gene and thereafter called #22bis [16]. Direct sequencing after bisulfite treatment of placental DNA provided a double peak sequence typical of a DMR at this location (Fig 2, B). We performed a Pyrosequencing approach to precisely evaluate the methylation levels. In 25 placental genomic samples, the average methylation level ranges between 30 and 50 % as often observed with imprinted genes, although variations are observed at different individual CpGs (Fig 2, C). We then cloned amplicons from this CpG #22bis region and sequenced 10-20 clones in some individuals. This confirmed that unmethylated CpG dinucleotides tend to concentrate on one allele while methylated ones are

also in phase on the other allele, suggesting a close to 50/50 allelic repartition (Fig 2, D). However, because of the presence of only 3 rare SNPs in the Caucasian population within the region(rs 795664, rs732279 and rs 114498), we were not able to identify heterozygous carriers and to assess the parental origin of the hypomethylated and hypermethylated alleles. We therefore attempted to analyse the methylation level of this region in DNAs extracted from 6 sperm samples, representing only the paternal allele. In these cases, the methylation level is much lower, at least 30% below that of placental DNA (Fig 2, C). This supports the idea that the methylated allele is the maternal one, consistent with a paternal expression of *DSCAM*. We therefore concluded that a DMR is present in intron 1 of the *DSCAM* gene, with a methylation imprint on the maternal allele.

Another maternally methylated DMR has been described [17] about 600 kb from the *DSCAM* gene, near the *WRB* gene (Tryptophan Rich Basic protein). We checked the imprinted status of this gene and of *PLAC4* (Placenta specific protein 4), located 300 kb from *DSCAM* and highly expressed in the placenta using rs 35946782 and rs 9305729 respectively. But both appeared to have a biallelic expression in the placenta, as confirmed by others for *WRB* [17].

#### *DSCAM cellular expression*

In the literature, *DSCAM* expression is mainly described in neurons. However, the placenta is a transitory organ devoid of neuronal cells. We therefore analysed *DSCAM* expression in placental tissue by immunohistochemistry to identify which cell types expressed *DSCAM*. A positive and strong labelling was observed in endothelial cells of placental capillaries and in syncytiotrophoblasts. The same result was observed whatever the term from early samples (13 and 19 weeks of amenorrhea) to late placentas (30-40 weeks) (Figure 3). *DSCAM* immunolabelling appears mainly in the cytoplasmic compartment of positive cells. A diffuse labelling is seldom detected in mesenchymal cells. Positive signals disappear in the presence

of a blocking DSCAM peptide but not in the presence of an irrelevant peptide. Similar observations were obtained on sections from IUGR affected pregnancies.

DSCAM expression was also validated by Western blot, using samples from cell lines models: extracts from both the HUVEC endothelial cell line and the choriocarcinoma trophoblastic cell line BeWo, revealed the expression of a 200 kDa signal that confirms the specific DSCAM expression and the specificity of the antibody (Figure 3, E).

#### *DSCAM expression in placental tissues*

As an imprinted gene expressed in the placenta, *DSCAM* becomes a candidate for a role in pregnancy-related pathologies. We therefore evaluated the expression level of *DSCAM* using real time RT-PCR in comparison to 3 reference genes recommended for the analysis of placental tissue (*SDHA*, *TBP* and *YWHAZ*) [18]. We did not find any significant difference for *DSCAM* expression between placentas from normal and pathological pregnancies (idiopathic growth restriction, growth restriction of vascular origin or pre-eclampsia) despite an important inter-individual variability (Fig 4, A).

In order to decipher a possible role of *DSCAM* in trophoblast syncytialisation, primary human trophoblastic cells from first trimester conceptuses were isolated and grown in a short-term culture (72h) to spontaneously differentiate into syncytiotrophoblast. By real time RT PCR, we evaluated the level of *DSCAM* expression in these samples, during differentiation. The expression of *DSCAM* was overall low and no difference could be detected during the process of syncytialisation (Fig 4, B). In a test of *in vitro* syncytialisation using BeWo cells treated with forskolin, *DSCAM* levels were too low to show any effect.

## **Discussion**



Imprinted genes are a subset of genes presenting a very peculiar mode of expression that is interesting both from the fundamental point of view and in relation to human pathologies. Since the identification of the first imprinted genes in the 80's, the list has reached about 150 in mice and 100 in human, without certitude about the real definitive number. Our goal was to identify new imprinted genes in the human placenta, particularly in relation to placental pathologies. We developed a strategy that already revealed the existence of 7 novel imprinted genes [10]. We pursued here this work that now leads to the identification of two novel imprinted genes, *AIM1* and *DSCAM*, and one monoallelically expressed gene, *NCAM2*, with a variable parent-of-origin expressed allele.

The first one, *AIM1*, has recently been described as imprinted for the first time simultaneously to our study [14]. We confirm the isoform-specific imprinting of this gene, as the long protein-coding transcript expressed from an upstream promoter is imprinted whereas the short isoform has a biallelic expression. This imprinting model has already been described for the *MEST* gene in human and pig [19, 20] and *GNAS* for example [21]. Recent data suggest that this phenomenon is rather common to many imprinted genes [22]. The *NCAM2* gene also shows a monoallelic maternal expression in most studied cases but some exceptions question its exact status. Such an example of a gene with a random parental origin of the expressed allele has also been reported for the *ZFAT* gene that we found as imprinted in our previously published analyses [23]. In such cases, extended cohorts should be analysed to determine the exact pattern of expression of these genes.

We also identified and validated the *DSCAM* gene as imprinted in the human placenta and showing an exclusive expression of the paternal allele. We could not determine if *DSCAM-AS1* and *DSCAM-IT1*, two non-coding RNAs present at the *DSCAM* locus were also

imprinted, because of a too weak expression level in placental samples. Antisense RNAs to imprinted genes are often submitted to the same regulation, as for *ZFAT-ASI/ZFAT*, *IGF2AS/IGF2* [24], and contribute to the complex regulation of these loci.

We explored the *DSCAM* locus to identify a DMR that could be responsible for its monoallelic expression. The high methylation levels that we observed at selected CpG islands located at the promoters of *DSCAM* and *DSCAM-ASI* were not compatible with a DMR. A genome-wide methylation study by Fang et al orientated us toward a region located in intron 1 [16]. By different methods, we observed that this region presents the characteristics of an imprinted DMR, with an allele globally methylated and the other one unmethylated. Sperm DNA analyses showed a very weak methylation level, compatible with that observed by Fang et al [16]. We concluded that this intron 1 fragment is a maternally imprinted DMR, compatible with a paternal expression of the *DSCAM* gene. We could identify in this region ([www.genomatix.de](http://www.genomatix.de)) 2 putative binding sites for the well-known insulator CTCF as well as for ZAC1/PLAGL1, the coordinator of the mouse imprinted gene network that could relate *DSCAM* to other imprinted genes.

The *DSCAM* gene encodes a cellular adhesion molecule from the Immunoglobulin family known to play roles in neural development and function in mammals and insects. In *Drosophila*, this gene exhibits a particularly rich alternative splicing producing tens of thousands of isoforms [25] while in Arthropods, it is known for its immunologic function [26]. *DSCAM* has been shown to be involved in axon guidance, dendrite arborisation involving self-avoidance and tiling [27]. Placentas are devoid of neurons and our investigations concluded that *DSCAM* was present in endothelial cells and

syncytiotrophoblasts in the human placenta. Therefore a new function for DSCAM has to be found in placental tissue.

DSCAM is a receptor for the secreted guidance factor Netrin 1 [28], which also binds to other receptors such as Neogenin, Deleted in Colorectal Cancer (DCC) and Uncoordinated-5 homolog (UNC5B). These genes have been extensively studied in the central nervous system but they are however also expressed in the placenta [29, 30]. DCC, similarly to DSCAM, is expressed by syncytiotrophoblasts, while Netrin 1 is detected within early cytotrophoblasts and later outside syncytiotrophoblasts. In a preliminary approach using isolated human cytotrophoblasts, we could not see any difference of expression of *DSCAM* between cyto- and syncytiotrophoblasts *in vitro*. To strictly exclude a role of DSCAM in syncytialisation, studies should be reproduced on larger numbers. Interestingly, placentas of Down syndrome patients present an abnormal development of the syncytiotrophoblast layer [31]. In our collection of human samples, no difference of *DSCAM* expression level could be identified between placentas from normal pregnancies and from pregnancies complicated by IUGR or preeclampsia. However, Netrin 1 expression level has been shown to be reduced in placentas from fetal growth restricted babies [32].

Netrins are reported to have broader roles in many tissues, modulating cell-cell and cell-matrix interactions [33]. They are involved in other branching phenomena in the lung and in the vasculature. Indeed, neurogenesis and vasculogenesis share common pathways and actors [34]. Rare publications investigated the role of Netrin 1 in a placental and vascular context. Xie et al showed that Netrin 1 has a pro-angiogenic effect in rat placenta as well as in cultured endothelial cells [35].

Mice carrying a targeted deletion or a mutation of the *Dscam* gene generally pursue their development until birth, with variations in the phenotype and in its severity according to

backgrounds and responsible mutations. Some die shortly after birth from respiratory distress due to an abnormal set up of neurons controlling the diaphragm [36]. In some cases, part of the homozygous null progeny seems to be missing at birth [37]. This shows that *Dscam* is not strictly compulsory to intrauterine life, its absence is compatible with a placental and foetal development, though it could be subnormal. Many other imprinted genes knock-out mice produce viable progeny as well, such as *Igf2*, *H19*, etc [38, 39]. *Dscam* null mice phenotype was mainly studied in a neuro-developmental context, including abnormal retinal development, abnormal amacrine cell morphology, etc [36, 37, 40]. Effects on reproduction were not detailed though a delay of gestation was observed [36] as well as an abnormal maternal behaviour [37]. The imprinted status of the *Dscam* gene should be explored in both the placental compartment and brain structures in mice. In particular, it should be studied in relation to other imprinted genes belonging to the imprinted gene network (IGN) [41].

The localisation of this gene on chromosome 21 is a novelty as no imprinted gene had been assigned to this smallest human autosome so far. Though we specifically targeted our previous results on this chromosome after the positive result of *DSCAM* and analysed some additional candidates, we have not been able to identify any other reliable imprinted gene on chromosome 21 since the imprinted status of *NCAM2* has to be more deeply characterised. Stoll et al had suggested that this chromosome was devoid of imprinted genes as they couldn't find any evidence of phenotypic difference between Down syndrome patients having an extra chromosome 21 from either paternal or maternal origin [42]. However the phenotypic heterogeneity of the disease makes correlations difficult [43]. The non-disjunction error responsible for the disease is very frequently of maternal origin (90%) [44]. In these cases, the level of a paternally expressed gene such as *DSCAM* wouldn't be affected by the presence of the supernumerary chromosome. *DSCAM* has been proposed as a candidate gene responsible

for intellectual disability [45], and for cardiac and visceral malformations [46, 47] in Down Syndrome.

In sum, we identified *DSCAM* as a new imprinted gene in the human placenta. Its role and function in this tissue is still unclear. The exploration of its imprinted status as well as its function in other animal models would be very useful to determine which role it plays in placental tissue and how it could interact with other imprinted genes within the IG Network. Given the limited data available, we hypothesize that *DSCAM* could promote placental growth. Indeed, according to the theory of parental conflict, imprinted genes expressed from the paternal copy, as *DSCAM*, tend to promote growth of the foeto-placental unit. Furthermore, its ligand, Netrin-1, has a pro-angiogenic effect in the placenta, which may promote placental vascularisation and hence placental development and growth.

### **Titles and legends to figures:**

Figure 1: Imprinting analysis of the *DSCAM* gene.

Sequences around polymorphisms rs 2297270 (black arrows) and rs 41395652 (grey arrows) in exon 5, from a representative individual's genomic DNA (A), placental cDNA (B) and maternal genomic DNA (C).

Figure 2: Epigenetic analysis of the *DSCAM* gene.

- A) Map of the *DSCAM* locus on chromosome 21. Upper bands represent exons of the *DSCAM* (black), *DSCAM-ASI* and *DSCAM-ITI* (grey) genes. CpG islands located within and around the gene are shown as stars. CpG island #22bis (in grey) is detailed with each band representing a CpG dinucleotide (position 40841468-40842223 on chromosome 21 (Genome version hg38)).
- B) Sequencing of the CpG #22bis island region from placental genomic DNA treated by bisulfite.
- C) DNA methylation level of CpG island #22bis by pyrosequencing. The grey squares represent the average of 25 placental samples, +/- standard deviation, while black triangles were data obtained from 6 sperm samples, +/- standard deviation.
- D) DNA methylation level of CpG island #22bis region by cloning-sequencing from a representative placental sample. White dots represent unmethylated cytosines and black dots, methylated cytosines.

Figure 3: Expression of *DSCAM* in placental tissues.

A to E) Immunohistochemistry on placental slides with an anti-*DSCAM* antibody (60x magnification). The gestational age of placentas is notified in the lower right corner in weeks of amenorrhea. Scale bar 10  $\mu$ m. ST: syncytiotrophoblasts; EC: endothelial cells

A and C) Staining from a representative sample; B) negative control, in the absence of the primary antibody; D) negative control in the presence of the blocking *DSCAM* peptide; E) control in the presence of a blocking peptide irrelevant to *DSCAM* (Sly peptide). F) Western blot using an anti-*DSCAM* antibody on cellular extracts from the HUVEC and BeWo cell lines (MW, molecular weight marker).

Figure 4: *DSCAM* expression levels in placental samples.

- A) Real time RT-PCR in pathological and control placental samples. (IUGR: intrauterine growth restriction (n=13), vIUGR: intrauterine growth restriction of vascular origin (n=15); PE: preeclampsia (n=12); PE+IUGR: combination of the maternal and foetal pathologies (n=5); controls (n=20)). The expression of *DSCAM* was normalized using 3 housekeeping genes.
- B) Real time RT-PCR in human isolated trophoblasts during spontaneous syncytialisation *in vitro*. (24h, n=7; 48h, n=6; 72h, n=4)

Supplemental Table 1: Candidate genes explored regarding their imprinted status.

Supplemental Table 2: Genes presenting a validated monoallelic expression, SNPs analysed, patients studied and parental origin of the alleles.

Supplemental Table 3: Primers and PCR conditions used.

Supplemental Figure 1: Flow chart of the strategy to identify genes expressed monoallelically.

Supplemental Figure 2: STR based genotyping results of the sample with discordant parental origin of the expressed *NCAM2* allele.

- I. Allelic repeats for D18S51, D21S11, TH01, D3S1358, Penta E for the maternal gDNA (a), the placental gDNA (b) and the placental RNA (c).
- II. Allelic repeats for FGA, TPOX, D8S1179, vWA and Amelogenin for the maternal gDNA (a), the placental gDNA (b) and the placental RNA (c).
- III. Allelic repeats for CSF1PO, D16S539, D7S820, D13S317, D5S818 and Penta D for the maternal gDNA (a), the placental gDNA (b) and the placental RNA (c).

### **Acknowledgments:**

We are indebted to Dr Ahmed Ziyat and Brigitte Lefèvre for their help. We are thankful to all patients and medical staff participating to the Biogestelle collection at the Cochin Hospital. We are grateful to Dr Deborah Bourc'his, Juliane Glaser, Raquel Perez Palacios, Guillaume Pidoux, Elma El Khouri, Charlotte Moretti, Béatrice Parfait and Paul Monnier for their help. The platforms of the Cochin Institute, Imagerie Cellulaire and GenomIC are thanked for their help and support. This work was supported by the ANR TWOTHYME.

### **References:**

1. McGrath, J. and D. Solter, *Completion of mouse embryogenesis requires both the maternal and paternal genomes*. Cell, 1984. **37**(1): p. 179-83.
2. Barton, S.C., M.A. Surani, and M.L. Norris, *Role of paternal and maternal genomes in mouse development*. Nature, 1984. **311**(5984): p. 374-6.
3. Tycko, B., *Imprinted genes in placental growth and obstetric disorders*. Cytogenet Genome Res, 2006. **113**(1-4): p. 271-8.
4. Prickett, A.R. and R.J. Oakey, *A survey of tissue-specific genomic imprinting in mammals*. Mol Genet Genomics, 2012. **287**(8): p. 621-30.
5. Keverne, B., *Monoallelic gene expression and mammalian evolution*. Bioessays, 2009. **31**(12): p. 1318-26.
6. Luedi, P.P., et al., *Computational and experimental identification of novel human imprinted genes*. Genome Res, 2007. **17**(12): p. 1723-30.



7. Luedi, P.P., A.J. Hartemink, and R.L. Jirtle, *Genome-wide prediction of imprinted murine genes*. *Genome Res*, 2005. **15**(6): p. 875-84.
8. Pollard, K.S., et al., *A genome-wide approach to identifying novel-imprinted genes*. *Hum Genet*, 2008. **122**(6): p. 625-34.
9. Wang, X., P.D. Soloway, and A.G. Clark, *A survey for novel imprinted genes in the mouse placenta by mRNA-seq*. *Genetics*, 2011. **189**(1): p. 109-22.
10. Barbaux, S., et al., *A genome-wide approach reveals novel imprinted genes expressed in the human placenta*. *Epigenetics*, 2012. **7**(9): p. 1079-90.
11. Gascoïn-Lachambre, G., et al., *Cullins in human intra-uterine growth restriction: expressional and epigenetic alterations*. *Placenta*, 2010. **31**(2): p. 151-7.
12. Benaitreau, D., et al., *Adiponectin promotes syncytialisation of BeWo cell line and primary trophoblast cells*. *Reprod Biol Endocrinol*, 2010. **8**: p. 128.
13. Miller, S.A., D.D. Dykes, and H.F. Polesky, *A simple salting out procedure for extracting DNA from human nucleated cells*. *Nucleic Acids Res*, 1988. **16**(3): p. 1215.
14. Das, R., et al., *DNMT1 and AIM1 Imprinting in human placenta revealed through a genome-wide screen for allele-specific DNA methylation*. *BMC Genomics*, 2013. **14**: p. 685.
15. Lambertini, L., et al., *A sensitive functional assay reveals frequent loss of genomic imprinting in human placenta*. *Epigenetics*, 2008. **3**(5): p. 261-9.
16. Fang, F., et al., *Genomic landscape of human allele-specific DNA methylation*. *Proc Natl Acad Sci U S A*, 2012. **109**(19): p. 7332-7.
17. Docherty, L.E., et al., *Genome-wide DNA methylation analysis of patients with imprinting disorders identifies differentially methylated regions associated with novel candidate imprinted genes*. *J Med Genet*, 2014. **51**(4): p. 229-38.
18. Meller, M., et al., *Evaluation of housekeeping genes in placental comparative expression studies*. *Placenta*, 2005. **26**(8-9): p. 601-7.
19. Li, X., et al., *Isoform-specific imprinting of the MEST gene in porcine parthenogenetic fetuses*. *Gene*, 2015. **558**(2): p. 287-90.
20. Pedersen, I.S., et al., *Promoter switch: a novel mechanism causing biallelic PEG1/MEST expression in invasive breast cancer*. *Hum Mol Genet*, 2002. **11**(12): p. 1449-53.
21. Peters, J., et al., *Imprinting control within the compact Gnas locus*. *Cytogenet Genome Res*, 2006. **113**(1-4): p. 194-201.
22. Stelzer, Y., et al., *Differentiation of human parthenogenetic pluripotent stem cells reveals multiple tissue- and isoform-specific imprinted transcripts*. *Cell Rep*, 2015. **11**(2): p. 308-20.
23. Metsalu, T., et al., *Using RNA sequencing for identifying gene imprinting and random monoallelic expression in human placenta*. *Epigenetics*, 2014. **9**(10): p. 1397-409.
24. Kanduri, C., *Long noncoding RNAs: Lessons from genomic imprinting*. *Biochim Biophys Acta*, 2015.
25. Zipursky, S.L., W.M. Wojtowicz, and D. Hattori, *Got diversity? Wiring the fly brain with Dscam*. *Trends Biochem Sci*, 2006. **31**(10): p. 581-8.
26. Schmucker, D. and B. Chen, *Dscam and DSCAM: complex genes in simple animals, complex animals yet simple genes*. *Genes Dev*, 2009. **23**(2): p. 147-56.
27. Millard, S.S. and S.L. Zipursky, *Dscam-mediated repulsion controls tiling and self-avoidance*. *Curr Opin Neurobiol*, 2008. **18**(1): p. 84-9.
28. Liu, G., et al., *DSCAM functions as a netrin receptor in commissural axon pathfinding*. *Proc Natl Acad Sci U S A*, 2009. **106**(8): p. 2951-6.
29. Dakouane-Giudicelli, M., et al., *Characterization and expression of netrin-1 and its receptors UNC5B and DCC in human placenta*. *J Histochem Cytochem*, 2010. **58**(1): p. 73-82.
30. Dakouane-Giudicelli, M., et al., *Identification and localization of netrin-4 and neogenin in human first trimester and term placenta*. *Placenta*, 2012. **33**(9): p. 677-81.
31. Malassine, A., et al., *Human trophoblast in trisomy 21: a model for cell-cell fusion dynamic investigation*. *Adv Exp Med Biol*, 2011. **714**: p. 103-12.

32. Qian-hua, W., et al., *Reduced expression of netrin-1 is associated with fetal growth restriction*. Mol Cell Biochem, 2011. **350**(1-2): p. 81-7.
33. Lai Wing Sun, K., J.P. Correia, and T.E. Kennedy, *Netrins: versatile extracellular cues with diverse functions*. Development, 2011. **138**(11): p. 2153-69.
34. Park, K.W., et al., *The axonal attractant Netrin-1 is an angiogenic factor*. Proc Natl Acad Sci U S A, 2004. **101**(46): p. 16210-5.
35. Xie, H., et al., *Effects of netrin-1 and netrin-1 knockdown on human umbilical vein endothelial cells and angiogenesis of rat placenta*. Placenta, 2011. **32**(8): p. 546-53.
36. Amano, K., et al., *DSCAM deficiency causes loss of pre-inspiratory neuron synchronicity and perinatal death*. J Neurosci, 2009. **29**(9): p. 2984-96.
37. Fuerst, P.G., et al., *Neurite arborization and mosaic spacing in the mouse retina require DSCAM*. Nature, 2008. **451**(7177): p. 470-4.
38. Ripoché, M.A., et al., *Deletion of the H19 transcription unit reveals the existence of a putative imprinting control element*. Genes Dev, 1997. **11**(12): p. 1596-604.
39. Varmuza, S. and K. Miri, *What does genetics tell us about imprinting and the placenta connection?* Cell Mol Life Sci, 2015. **72**(1): p. 51-72.
40. Schramm, R.D., et al., *A novel mouse Dscam mutation inhibits localization and shedding of DSCAM*. PLoS One, 2012. **7**(12): p. e52652.
41. Varrault, A., et al., *Zac1 regulates an imprinted gene network critically involved in the control of embryonic growth*. Dev Cell, 2006. **11**(5): p. 711-22.
42. Stoll, C., et al., *No evidence for genomic imprinting in liver-born Down syndrome patients*. Acta Genet Med Gemellol (Roma), 1996. **45**(1-2): p. 265-71.
43. Muranjan, M., T. Chaudhari, and B.R. Vundinti, *Phenotypic heterogeneity and parental origin of extra chromosome 21 in Down syndrome*. Indian Pediatr, 2010. **47**(5): p. 429-32.
44. Ballesta, F., et al., *Parental origin and meiotic stage of non-disjunction in 139 cases of trisomy 21*. Ann Genet, 1999. **42**(1): p. 11-5.
45. Yamakawa, K., et al., *DSCAM: a novel member of the immunoglobulin superfamily maps in a Down syndrome region and is involved in the development of the nervous system*. Hum Mol Genet, 1998. **7**(2): p. 227-37.
46. Jannot, A.S., et al., *Chromosome 21 scan in Down syndrome reveals DSCAM as a predisposing locus in Hirschsprung disease*. PLoS One, 2013. **8**(5): p. e62519.
47. Barlow, G.M., et al., *Down syndrome congenital heart disease: a narrowed region and a candidate gene*. Genet Med, 2001. **3**(2): p. 91-101.

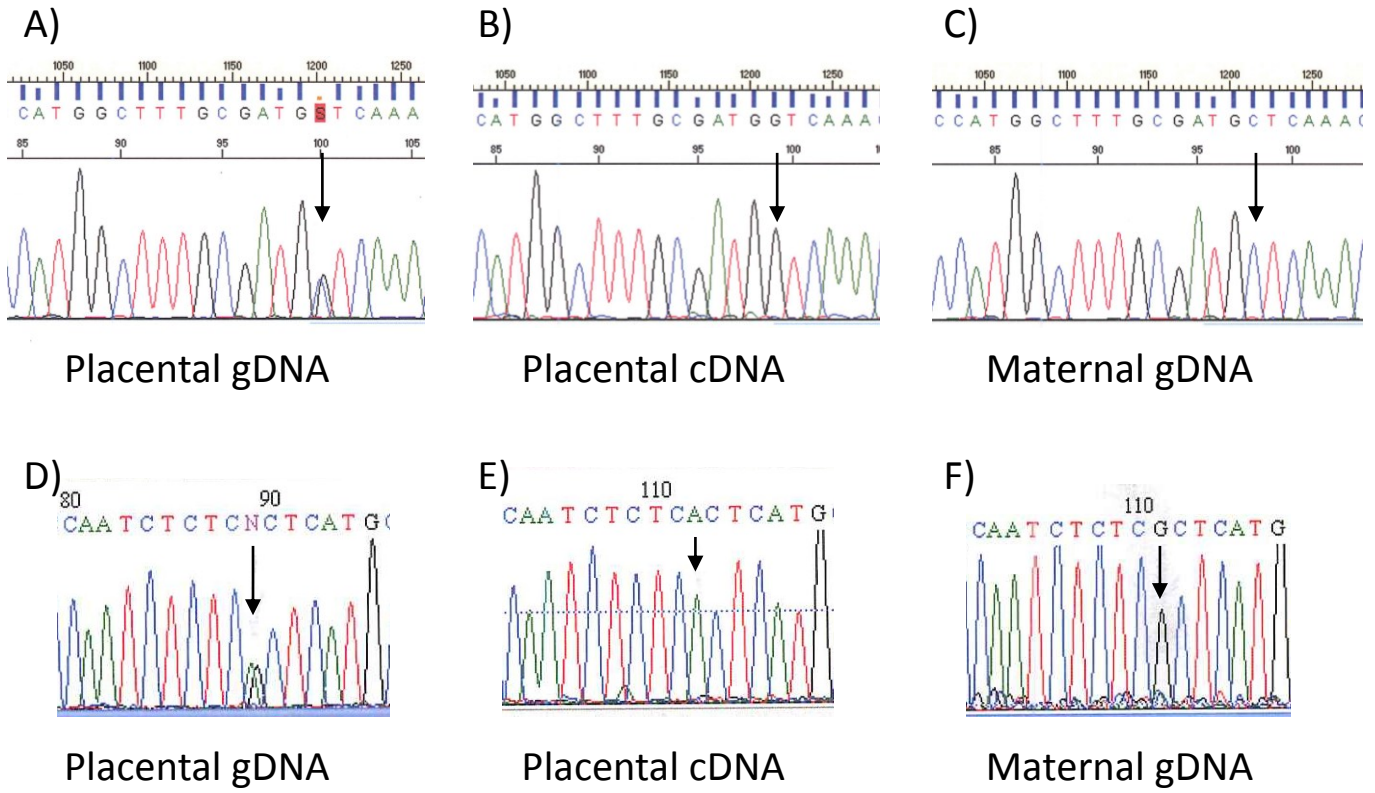


Figure 1

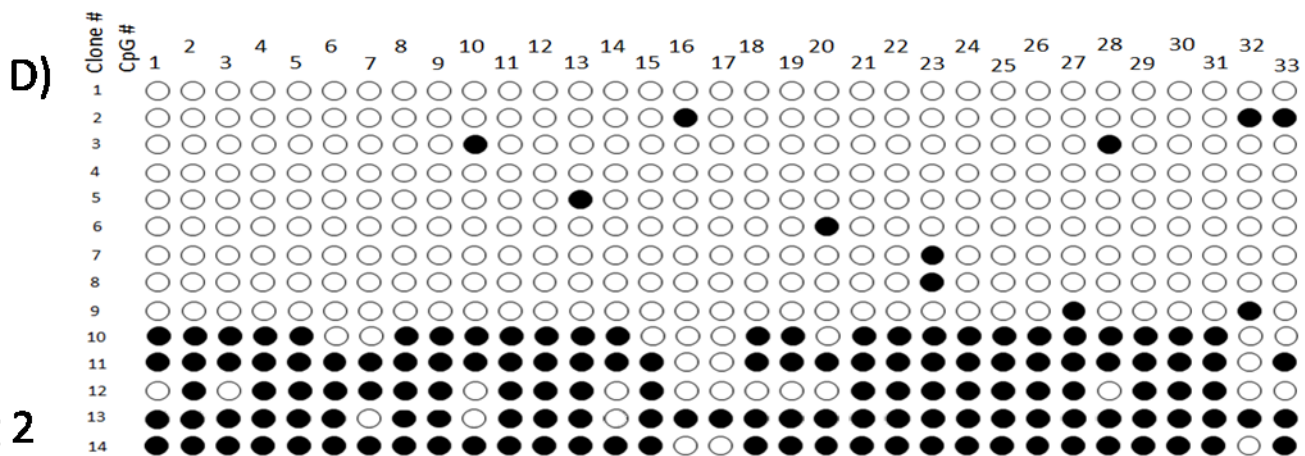
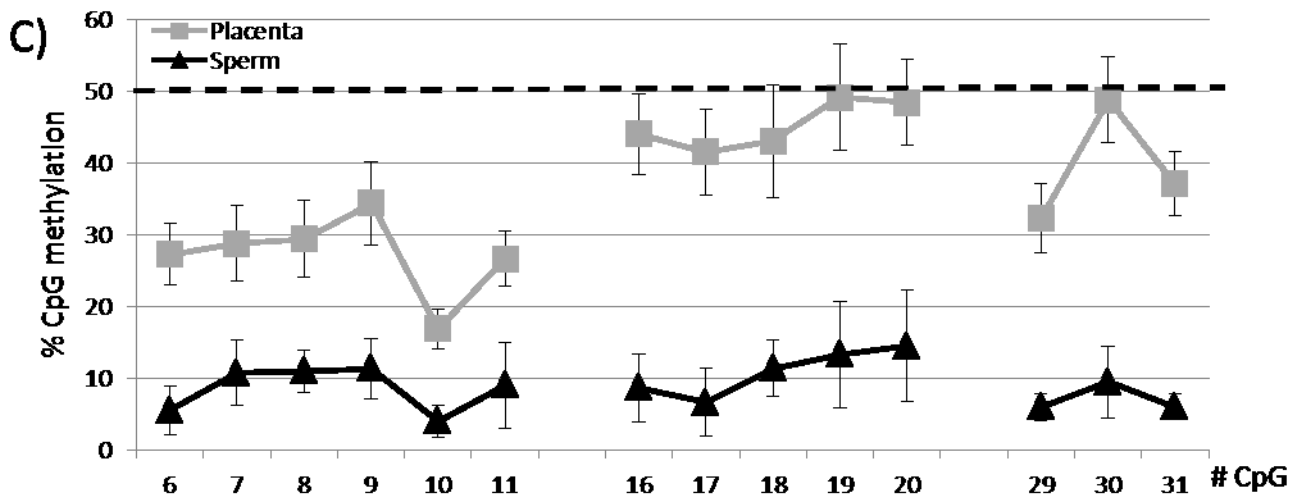
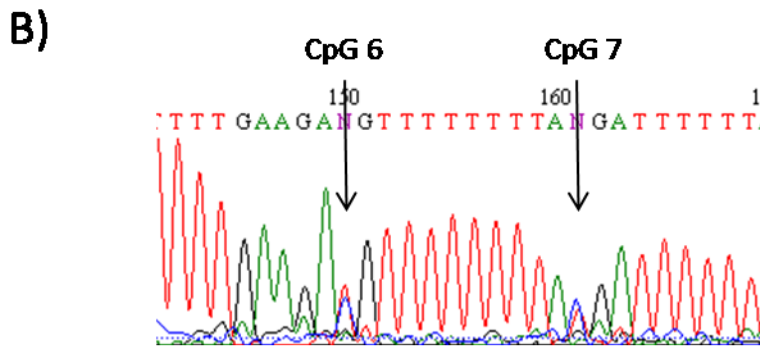
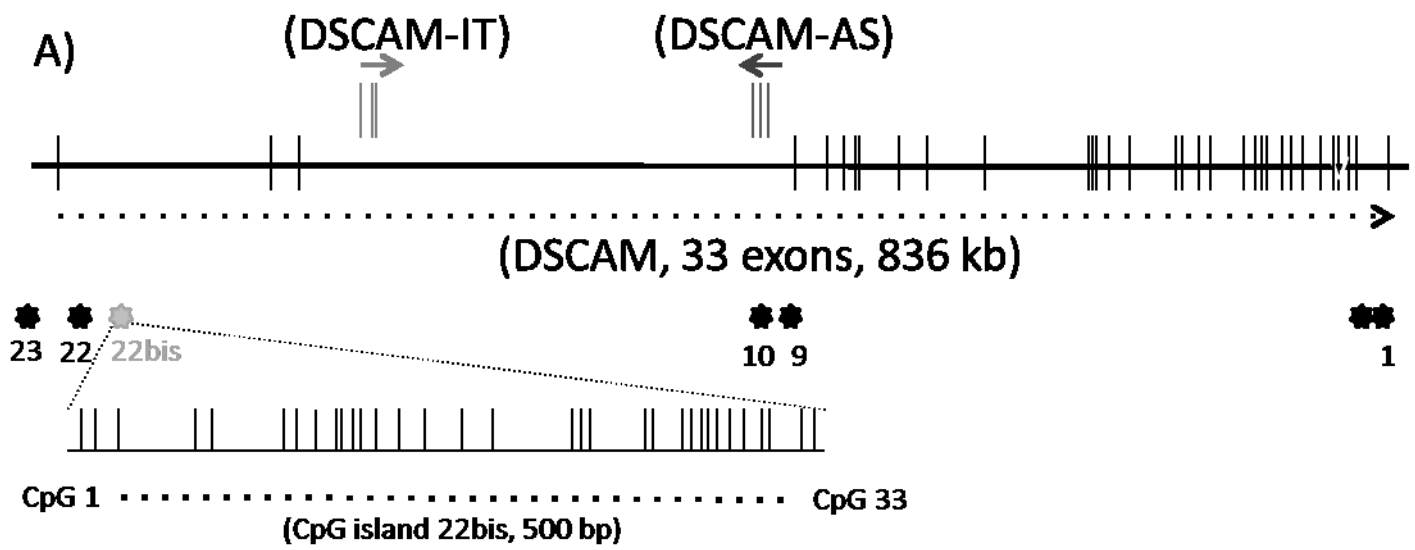


Fig 2

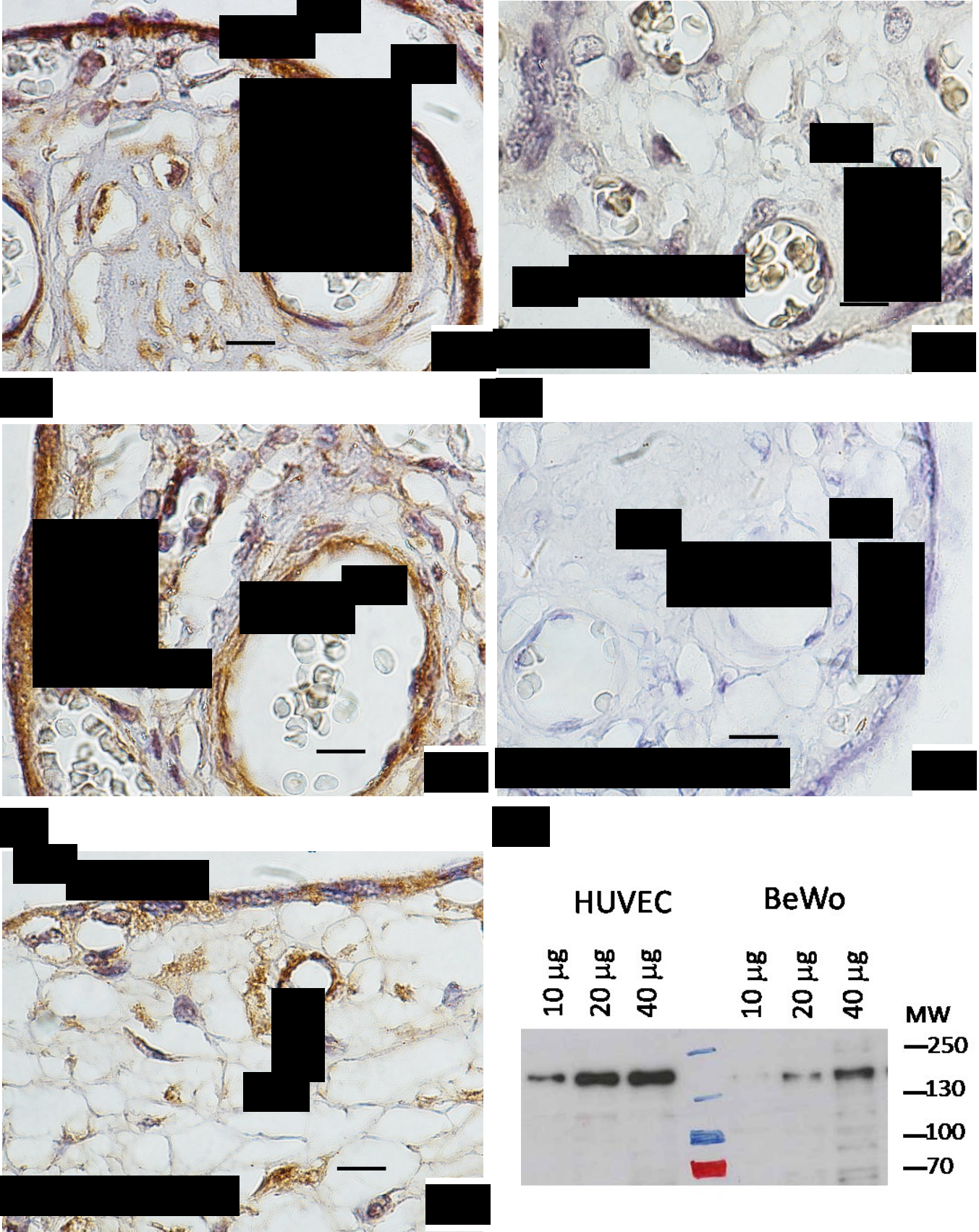
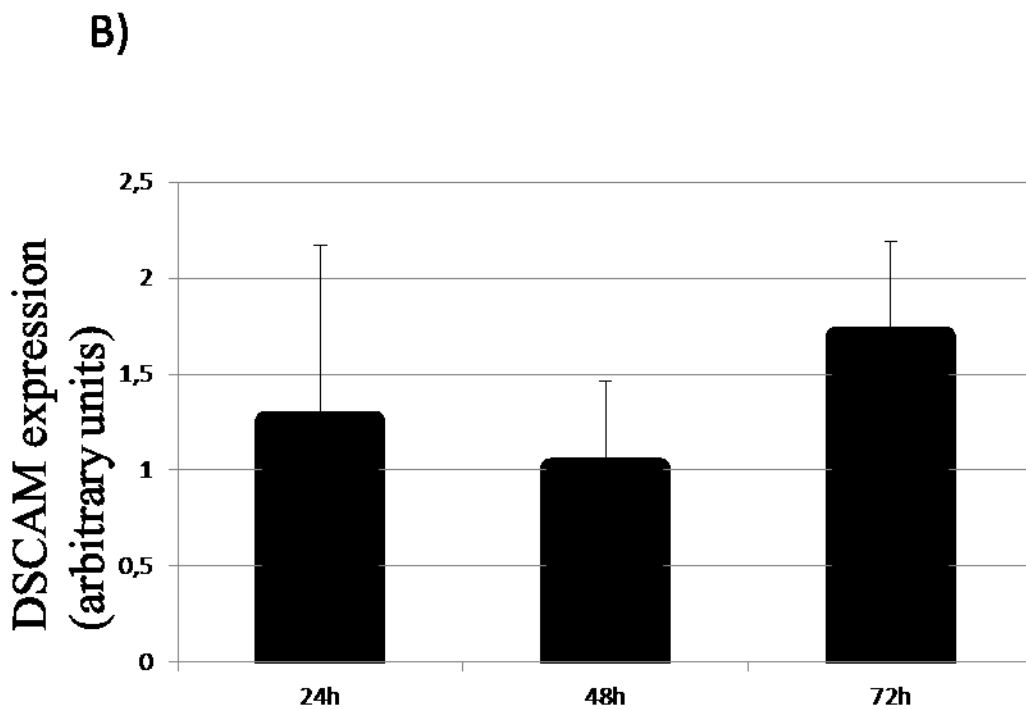
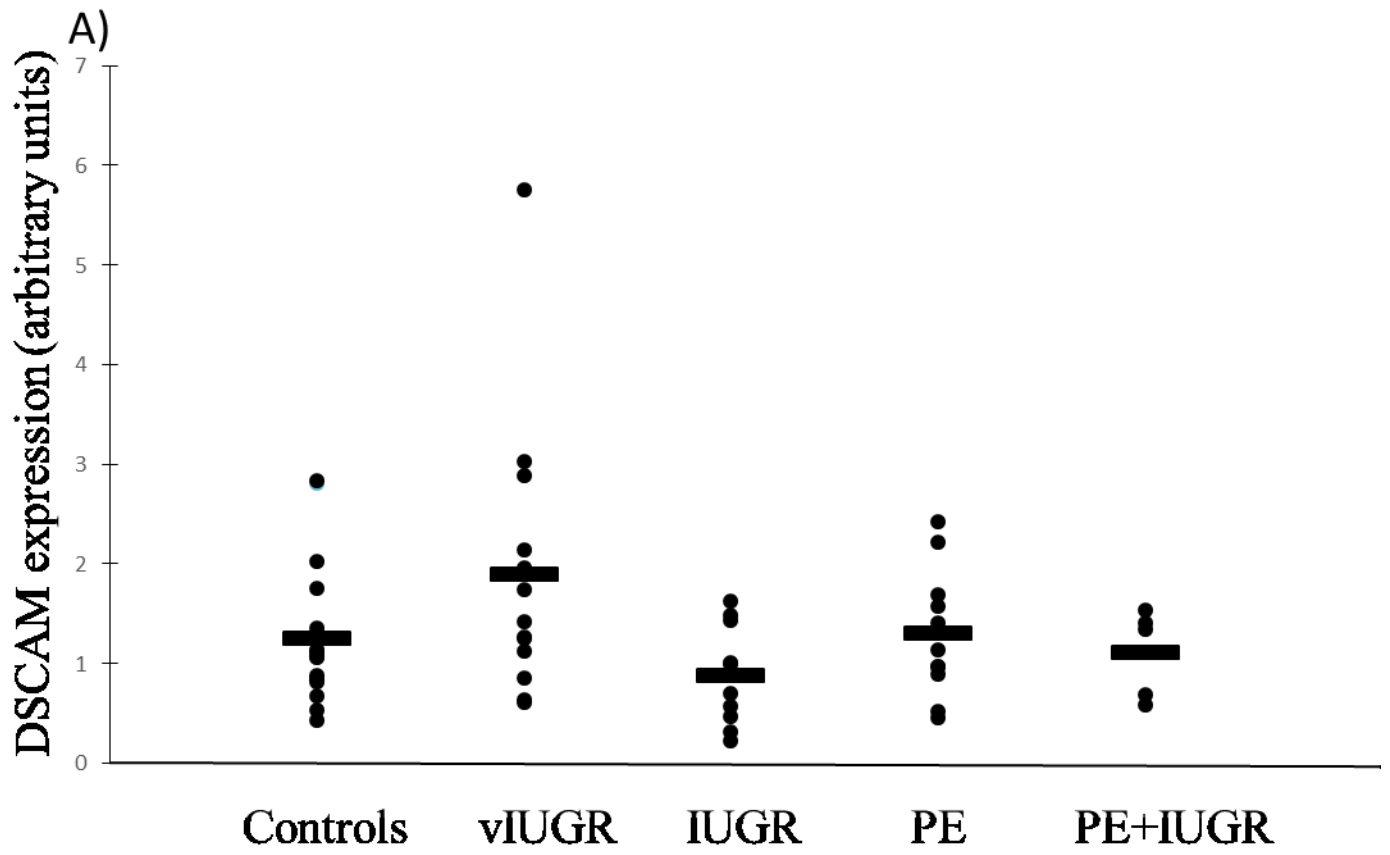
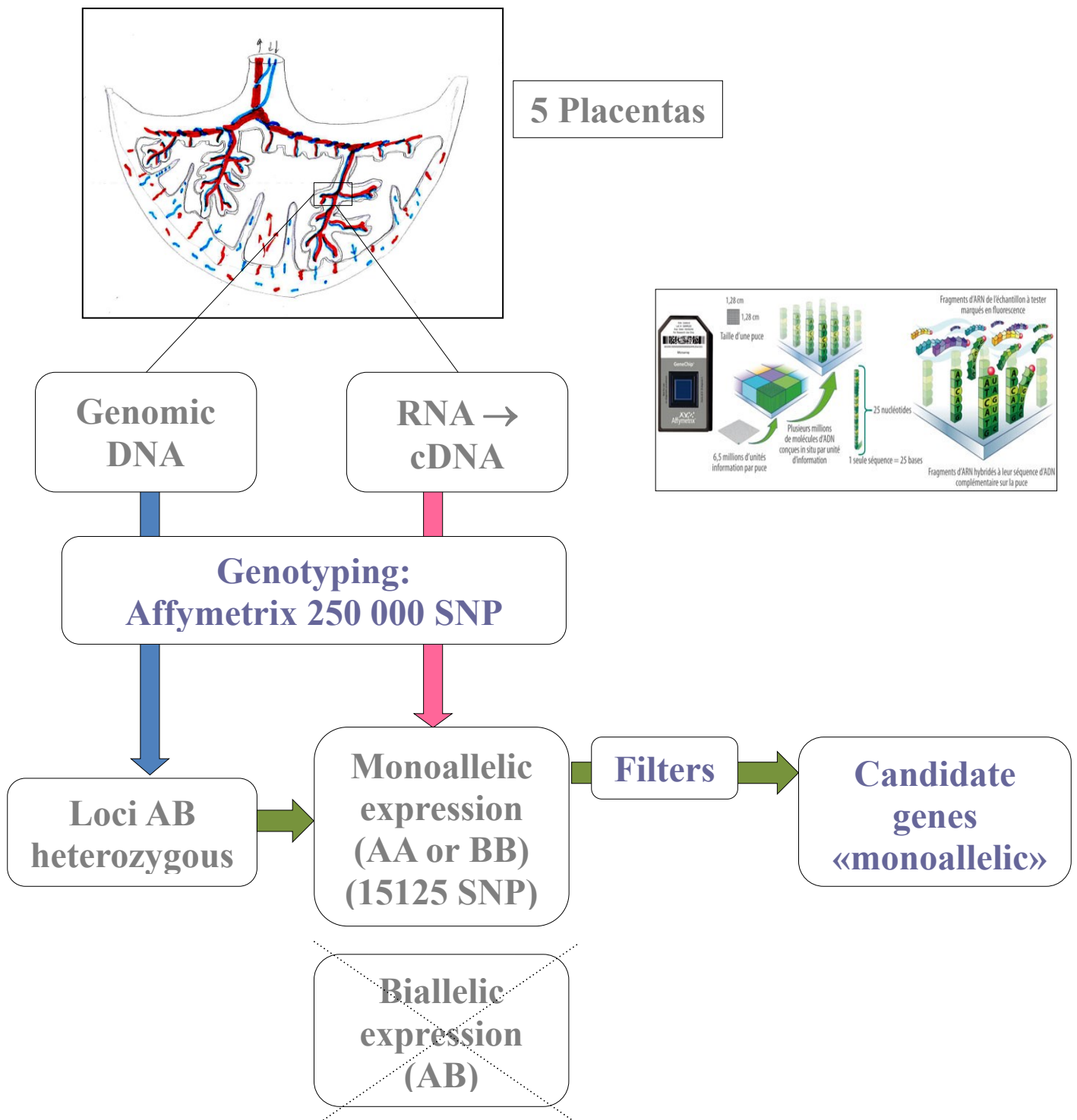


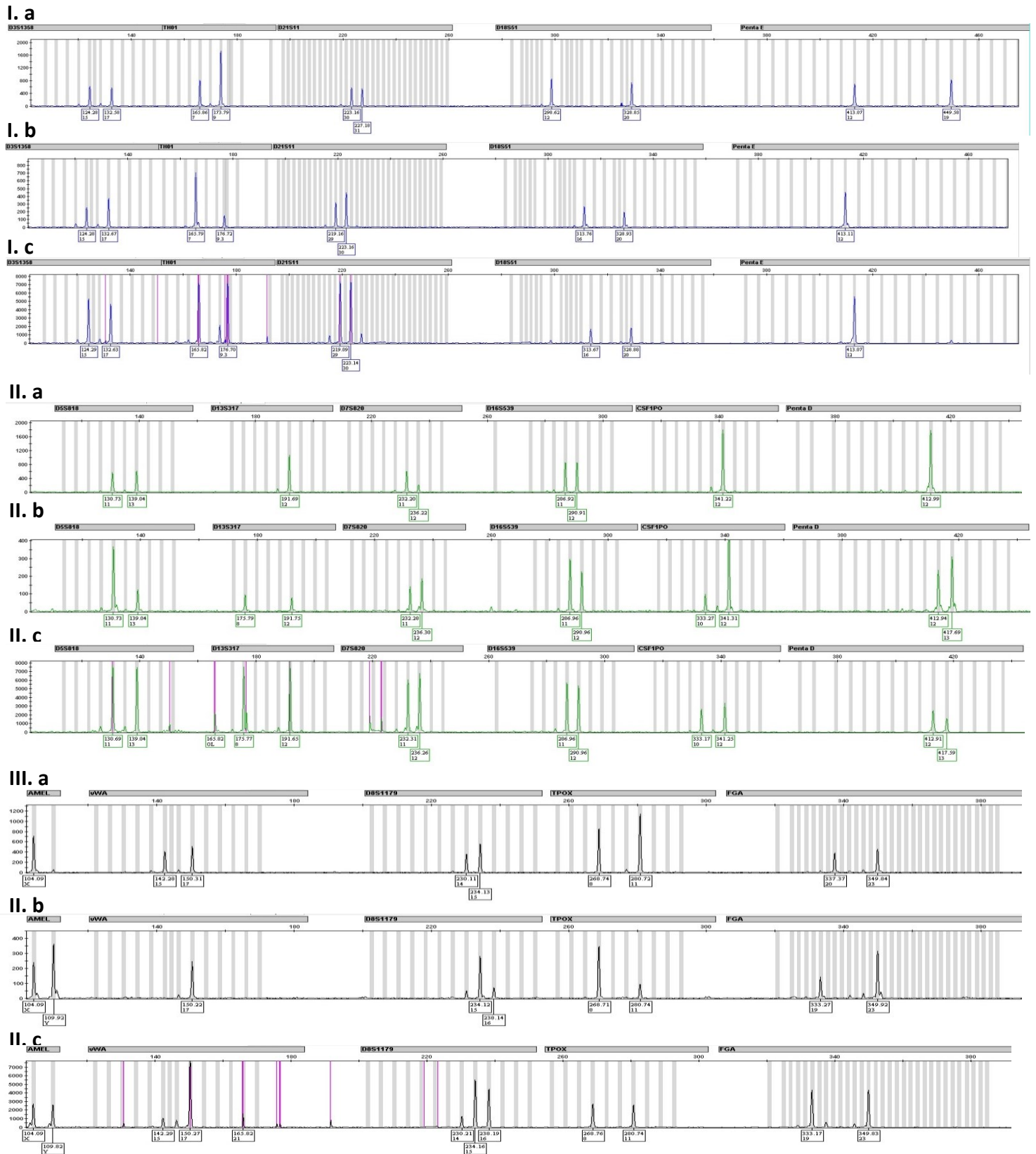
Figure 3



**Figure 4**



Supplemental Figure 1



Supplemental Figure 2



Supplemental Table 1

Gene	Chromosome	Argument for analysis	Status	Problem
<i>CSMD1</i>	8	Initial arrays	Biallelic	
<i>CTNND2</i>	5	Initial arrays	?	Lack of cDNA amplification
<i>KAZN</i>	1	Initial arrays	?	Absence of heterozytes
<i>PARK2</i>	6	Initial arrays	?	Lack of cDNA amplification
<i>PACRG</i>	6	Initial arrays	?	Lack of cDNA amplification
<i>CDH11</i>	16	Initial arrays	Biallelic	
<i>FAM5C</i>	1	Initial arrays	Biallelic	
<i>LRRC46</i>	17	Initial arrays	Biallelic	
<i>NPAS3</i>	14	Initial arrays	Biallelic	
<i>PCDH7</i>	4	Initial arrays	Biallelic	
<i>ROBO2</i>	3	Initial arrays	Biallelic	
<b><i>AIM1</i></b>	6	Initial arrays, proximity to <i>LIN28B</i> (1Mb)	Imprinted	
<b><i>DSCAM</i></b>	21	Initial arrays	Imprinted	
<b><i>NCAM2</i></b>	21	Re-analysis of chr 21	Conflicting	Imprinted or rando monoallelic?
<i>BRWD1</i>	21	Re-analysis of chr 21	Biallelic	
<i>TIAM1</i>	21	Re-analysis of chr 21	Biallelic	
<i>PSMG1</i>	21	Re-analysis of chr 21	Biallelic	
<i>KRTAP13-4</i>	21	Re-analysis of chr 21	?	Sequencing problem
<i>KRTAP27-1</i>	21	Re-analysis of chr 21	Conflicting	Insufficient cases
<i>B3GALT5</i>	21	Re-analysis of chr 21	?	Sequencing problem
<i>ERG</i>	21	Re-analysis of chr 21	?	Lack of cDNA amplification
<i>GRIK1</i>	21	Re-analysis of chr 21	?	Lack of cDNA amplification
<i>SLC37A1</i>	21	Re-analysis of chr 21	Biallelic	
<i>PCBP3</i>	21	Re-analysis of chr 21	Biallelic	
<i>LIP1</i>	21	Re-analysis of chr 21	?	Lack of cDNA amplification
<i>SAMSN1</i>	21	Re-analysis of chr 21	Conflicting	Insufficient cases
<i>ADAMTS5</i>	21	Re-analysis of chr 21	Biallelic	
<i>ZBTB21</i>	21	Literature (imprinted in horse)	Biallelic	
<i>WRB</i>	21	proximity to <i>DSCAM</i> , DMR	Biallelic	
<i>PLAC4</i>	21	proximity to <i>DSCAM</i>	Biallelic	

Supplemental Table 2

Gene	SNP	Placentas tested	Heterozygous placentas	Expression status	Parental origin
<i>AIM1</i>	rs11152999, T>C	30	2	monoallelic	not tested
	rs2297970, G>A	23	2	monoallelic	not tested
			6	biallelic	not relevant
<i>DSCAM</i>	rs41395652, C>T	40	4	monoallelic	paternal
			4	monoallelic	not informative
	rs2297270, G>C	40	1	monoallelic	paternal
			1	monoallelic	not informative
rs34336407, G>A	42	3	monoallelic	paternal	
		1	monoallelic	not informative	
<i>NCAM2</i>	rs232518, C>T	41	3	monoallelic	maternal
			1	monoallelic	paternal
			8	monoallelic	not informative
			1	biallelic	not relevant
	rs2017705, C>T	39	1	monoallelic	maternal
			3	monoallelic	not informative

Supplemental Table 3

Gene	Species	Primer F	Sequence	Primer R	Sequence	Use
<i>DSCAM</i>	Human	gDNAintron4 left	AACTCCTCATTCTGCCTCTAA	exon5 right	TCACGGTCTTCTGGAACCTC	gDNA Genotyping rs2297270/41395652
<i>DSCAM</i>	Human	exon9 left	CTCAGGGGACTTACCCATCA	gDNAintron9 right	TTAGGGGAATTCATGGTTGG	gDNA Genotyping rs34336407
<i>DSCAM</i>	Human	cDNAexon5 left	AACAGCGCCAGACTTTTTGT	exon5 right	TCACGGTCTTCTGGAACCTC	cDNA Genotyping rs2297270/41395652
<i>DSCAM</i>	Human	exon9 left	CTCAGGGGACTTACCCATCA	cDNAexon10 right	GGGTAACCCTCAGCAGAACA	cDNA Genotyping rs34336407
<i>DSCAM</i>	Human	exon 31 F	GCGTCAGCCCCTCGCAAGAC	exon 32 R	CCTGCCGTCAACTGCTCCGA	qPCR
<i>SDHA</i>	Human	F	TGG GAA CAA GAG GGC ATC TG	R	CCA CCA CTG CAT CAA ATT CAT G	qPCR
<i>TBP</i>	Human	F	TGC ACA GGA GCC AAG AGT GAA	R	CAC ATC ACA GCT CCC CAC CA	qPCR
<i>YWHAZ</i>	Human	F	ACT TTT GGT ACA TTG TGG CTT CAA	R	CCG CCA GGA CAA ACC AGT AT	qPCR
<i>DSCAM</i>	Human	CpG9 F2	ttTTGATTGGtATTtAGGTTTATGT	CpG9 R2	TTaCATACTTTTCAATCCCTCTTaC	DMR search, Bisulfite DNA
<i>DSCAM</i>	Human	CpG10 F	AGTGGGGTGTGTATGTAAGGATTT	CpG10 R	CCACCTTAATCCTCATCACCTTTA C	DMR search, Bisulfite DNA
<i>DSCAM</i>	Human	CpG23 F3	tAAGGATTTGGAAGGAGtATTTt	CpG23 R3	CTCAaTCACATaaATCCCTCTaC	DMR search, Bisulfite DNA
<i>DSCAM</i>	Human	CpGFang F4	gaggaTtgTt TTatgTagat gg	CpGFang R4 (+/- biotin)	ccttctaaacaccttaaatc	DMR search, Bisulfite DNA, Pyroseq
<i>DSCAM</i>	Human	CpGFang F5	gaaTTtaga aggtgTTtag aaagg	CpGFang R1 (+/- biotin)	tcaccctAtcaccatcaAca	DMR search, Bisulfite DNA, Pyroseq
<i>DSCAM</i>	Human	Pyro1	ggtTaggggTttTtTtg			Pyrosequencing primer
<i>DSCAM</i>	Human	Pyro2	ggtTTagTagTtttg			Pyrosequencing primer
<i>DSCAM</i>	Human	Pyro3	gaaaggaaaTaTaggtgtg			Pyrosequencing primer
<i>DSCAM</i>	Human	Pyro5	ggaggtggTagTtT			Pyrosequencing primer
<i>DSCAM</i>	Human	SNP2-S (biotin)	GCTCCTTGAGGATTTCCAATCTCT	SNP2-AS	CGGGCTATGCAGGTGTAATTCC	Pyrosequencing rs34336407
<i>DSCAM</i>	Human	SNP2-R	AATCCCATTGTGCAT			Pyrosequencing rs34336407
<i>AIM1</i>	Human	gDNAexon12 F	CAAACCCATCTGCTTCTAAAGG	gDNAexon12 R	CATGAACCAGGAACACACATT	gDNA genotyping rs2297970
<i>AIM1</i>	Human	cDNAexon12 F	TGGATCCAAAGGTTC	cDNAexon12 R	CCTCCTGAAACTCTGCAAGAC	cDNA genotyping rs2297970
<i>AIM1</i>	Human	Exon2 F	TGGTCCTCAGCCTAACCAAG	Exon2 R	CTTCCCACAGCTTCTGGACT	gDNA/CDNA genotyping rs11152999
<i>PLAC4</i>	Human	Exon1 F	TGGAGTGAAGGGGGTTAAA	Exon1 R	TGCAGTGGCACAATCGTAAT	gDNA/CDNA genotyping rs9305729

Supplemental Table 3 (suite)

Gene	Species	Primer F	Sequence	Primer R	Sequence	Use
<i>WBR</i>	Human	Exon1 F	TCTGCCAAAAAGTCCCAAAC	Exon1 R	AGTGAAAGCTCGGACAGCTC	gDNA/CDNA genotyping rs35946782
<i>NCAM2</i>	Human	Intron8 F	ggtcaccatttgcacacac	Exon9 R	TTTGTATGCCCTCCAATTCT	gDNA Genotyping rs232518
<i>NCAM2</i>	Human	Exon7 F	ATGTCTGCAGGGCCACAAAT	Exon10 R	CCAGTGAATTGATGCTGGTGG	cDNA Genotyping rs232518
<i>NCAM2</i>	Human	Exon16 F	TAATGGTCTTGGGCTTGGAG	Intron16 R	aggtcagagagcaggcagtg	gDNA Genotyping rs2017705
<i>NCAM2</i>	Human	Exon16 F	TAATGGTCTTGGGCTTGGAG	Exon17 R	GGCTCCCATCTTCGTGATTA	cDNA Genotyping rs2017705
<i>Dscam</i>	Mouse	Exon17 F	GATCATTCGCGGCTACCAA	Exon18 R	TAGATGACCCTGAACCCCTG	gDNA genotyping rs46901054, rs46747365, rs46432878
<i>Dscam</i>	Mouse	Exon17int F	TAGGCTACCGGGAGTACAGC	Exon18int R	GGTTGCTATGGCTTGGACAT	cDNA genotyping rs46901054, rs46747365, rs46432878
<i>Igf2</i>	Mouse	Exon3 F	GGCCCTCCTGGAGACATACTG	Intron3 R	gagctccatctaggtgc	gDNA genotyping rs248081537
<i>Igf2</i>	Mouse	Exon3 F	GGCCCTCCTGGAGACATACTG	Exon4 R	CTTTGAGCTCTTTGGCAAGGAT	cDNA genotyping rs248081537

Nous venons de décrire le travail qui a mené à l'identification de trois nouveaux GSE, *AIM1*, *DSCAM* et *NCAM2*, sur 28 gènes candidats analysés (article ci-avant). Nous nous sommes intéressés plus particulièrement à l'un d'entre eux, *DSCAM*. Ce gène est localisé en 21q22.2 dans la région critique contenant les gènes candidats pour la déficience intellectuelle et les anomalies viscérales (atrésie intestinale et maladie de Hirschsprung) associées à la trisomie 21. Il code pour une protéine de la superfamille des Immunoglobulines de la classe des molécules d'adhésion cellulaire des neurones. Cette protéine est largement exprimée dans le système nerveux central. Elle est conservée au cours de l'évolution et a été surtout étudiée chez la Drosophile qui en possède plus de 38000 isoformes, issues d'épissage alternatif, participant à la formation des axones pendant l'embryogenèse et à l'établissement des connexions neuronales par la reconnaissance interneuronale (Hattori et al., 2007; Schmucker et al., 2000). Des travaux chez la souris montrent également le rôle de *DSCAM* dans les fonctions neuronales et la formation de la rétine chez les mammifères (Fuerst et al., 2008; Huang et al., 2015; Li et al., 2015; Maynard and Stein, 2012; Zhang et al., 2015). Le rôle que joue *DSCAM* dans le placenta reste inconnu. Nous avons donc voulu l'étudier, guidés par le profil d'expression observé par l'immunohistochimie des coupes placentaires. L'expression de *DSCAM* dans les cellules endothéliales nous a fait suspecter une fonction dans l'angiogenèse placentaire, d'autant plus que le ligand de *DSCAM* qui est la Netrin-1 (*NTN1*) possède des fonctions connues dans ce processus (Chen et al., 2014; Ding et al., 2014). Nous avons donc mené des études fonctionnelles sur un modèle de cellules endothéliales humaines immortalisées (*Human Umbilical Vein Endothelial Cells*, HUVEC) dont les résultats sont exposés dans la partie « Résultats préliminaires ». Par ailleurs, nous avons profité de la disponibilité, au sein de l'Institut, de modèles murins déficients en *Ntn1* et *Dcc*, respectivement le ligand de *DSCAM* et un autre récepteur de *NTN1* avec lequel *DSCAM* peut collaborer (Huang et al., 2015), pour étudier le phénotype de croissance fœtale et placentaire de ces modèles.

## **2. Résultats préliminaires**

Dans cette partie, nous présentons des résultats que nous sommes en train de confirmer.

### **2.1. Extension de l'étude du statut d'empreinte à d'autres mammifères**

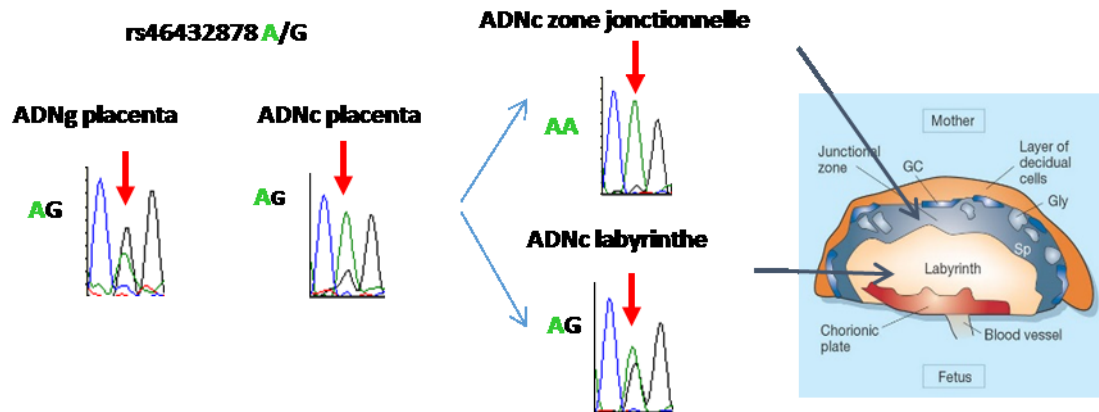
Pour un GSE donné, l'empreinte peut varier d'une espèce à l'autre (cf Section « Dynamique de

l’empreinte parentale », p. 49). L’étude de l’empreinte dans d’autres espèces a un intérêt fondamental pour suivre sa conservation au cours de l’évolution et un intérêt expérimental pour envisager des modèles animaux pour les études fonctionnelles. Nous avons donc investigué le statut d’empreinte de *DSCAM* chez la souris et le bovin.

#### *Chez la souris*

Notre méthode d’identification des GSE est basée sur la recherche d’une expression monoallélique en suivant l’expression des allèles au niveau de SNP. Ceci implique de disposer d’individus hétérozygotes. Or, les souris disponibles dans les laboratoires ont un fond génétique homozygote. Nous avons généré des souris polymorphes en croisant deux souches, *Mus musculus domesticus* et *Mus musculus molossinus*, dont les génotypes homozygotes sont divergents. Nous avons obtenu des générations F1 hétérozygotes chez lesquelles nous avons prélevé aux stades E13,5 ou E16,5 différents tissus : des régions du système nerveux central (Télocéphale, Mésencéphale, Cervelet, Bulbe rachidien, Moëlle épinière), du poumon et du placenta entier ou séparé en zone jonctionnelle et labyrinthe.

Le génotypage de l’ADNc du placenta entier a montré une expression biallélique mais avec un biais important qui se reflète par la différence de hauteur de pic entre les deux allèles (Figure 23). Cette observation nous a conduit à suspecter une hétérogénéité de l’empreinte entre les différents compartiments du placenta murin à savoir le labyrinthe et la zone jonctionnelle. L’étude de l’ADNc de ces deux régions a démontré l’existence d’une expression monoallélique limitée à la zone jonctionnelle et une absence d’empreinte dans le labyrinthe. Ces résultats sont en cours de confirmation sur un nombre plus important de placentas, obtenus dans les deux sens de croisement. L’étude de l’empreinte dans d’autres tissus est également en cours.



**Figure 23 :** Résultats du séquençage de *Dscam* dans le placenta murin. L'ADNc du placenta total montre une expression bi-allélique avec un allèle majoritaire. L'analyse des deux compartiments placentaires (la zone jonctionnelle et le labyrinthe) met en évidence une expression mono-allélique dans la zone jonctionnelle et bi-allélique dans le labyrinthe.

#### Chez le bovin

L'empreinte chez cette espèce a un intérêt économique puisque les techniques d'Assistance Médicale à la Procréation appliquées aux bovins, pour sélectionner les races les plus recherchées (qualité de la viande, qualité de production laitière...), entraînent un taux élevé de pathologie de l'empreinte dite syndrome de surcroissance ou *Large Offspring Syndrome*.

Après analyse d'une première série d'échantillons de placentas bovins, nous n'avons pas identifié d'individu hétérozygote pour pouvoir étudier l'expression allélique. Nous avons donc récupéré des échantillons de placentas issus de croisement inter-espèces entre *Bos taurus* et *Bos indicus* afin d'avoir un maximum de polymorphismes dans la descendance comme pour les souris. Ces échantillons sont également en cours d'analyse.

### 2.2. Analyse du profil d'expression de *DSCAM* dans l'embryon murin

Nous avons réalisé une hybridation *in situ* sur coupes d'embryons murins sauvages afin de caractériser le profil d'expression de *Dscam*. Des embryons de souris Swiss ont été prélevés à E13,5 avec leur placenta, fixés dans du Paraformaldéhyde à 4 % puis enrobés dans du sucre pour congélation. Des coupes sagittales et coronales ont été réalisées au Cryostat. L'hybridation *in situ* a été effectuée, avec une sonde anti-ARNm *Dscam* que nous avons synthétisée, selon une procédure précédemment publiée (Causeret et al., 2004).

Nous avons observé une expression restreinte au système nerveux central, au poumon (au niveau de la lame basale des bronches/bronchioles), au cœur et au placenta.

### **2.3. Etudes fonctionnelles à partir des cellules HUVEC**

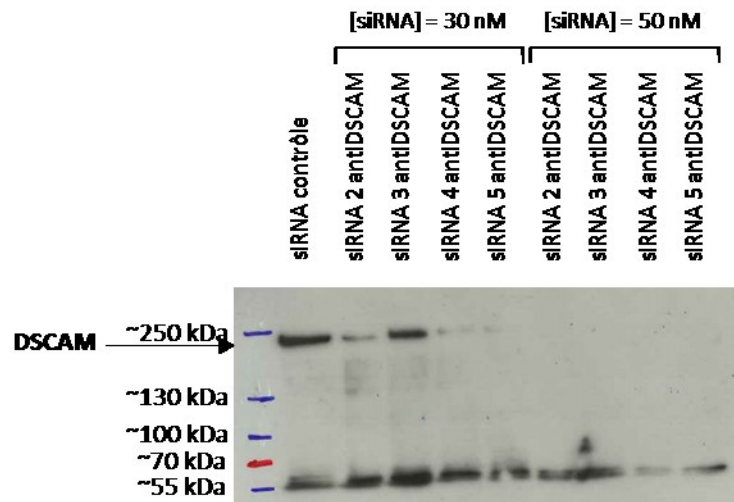
L'immunohistochimie sur coupes de placentas humains a montré une expression de DSCAM dans le syncytiotrophoblaste et les cellules endothéliales. Nous avons choisi le modèle de la lignée des cellules endothéliales de la veine ombilicale humaine (*Human Umbilical Vein Endothelial Cell*, HUVEC) pour effectuer les études fonctionnelles car, d'une part, ces cellules proviennent de la veine ombilicale et sont donc assez proches des capillaires villosités fœtaux du placenta, et d'autre part elles sont largement caractérisées dans la littérature ce qui facilite l'accès aux informations, réactifs et protocoles les concernant.

Afin d'explorer le rôle de DSCAM dans les cellules endothéliales, nous avons étudié les conséquences de son inhibition, par le système d'ARN interférence, sur les fonctions des HUVEC.

#### *Mise en silence de DSCAM par ARN interférence*

Nous avons mis en culture des HUVEC dans le milieu DMEM+ (Lifetech) complété avec 10 % de Sérum de veau fœtal (Lifetech) et 1 % d'antibiotiques (Pénicilline-Streptomycine, Lifetech). Les cellules ont ensuite été transfectées par le siRNA-4 anti-DSCAM qui induisait une disparition complète de la protéine en Western Blot (Qiagen) en utilisant la lipofectamine (RNAimax, Invitrogen). Après 72 h, les cellules sont utilisées pour différents tests : (1) l'analyse protéique par Western Blot afin de vérifier l'efficacité de l'inhibition (Figure 24), (2) l'étude morphologique, (3) le test d'angiogenèse dans le Matrigel, et (4) le test de la blessure. Les mêmes expériences ont été réalisées en utilisant un siRNA contrôle (Santa Cruz Biotechnology) et en absence de transfection.

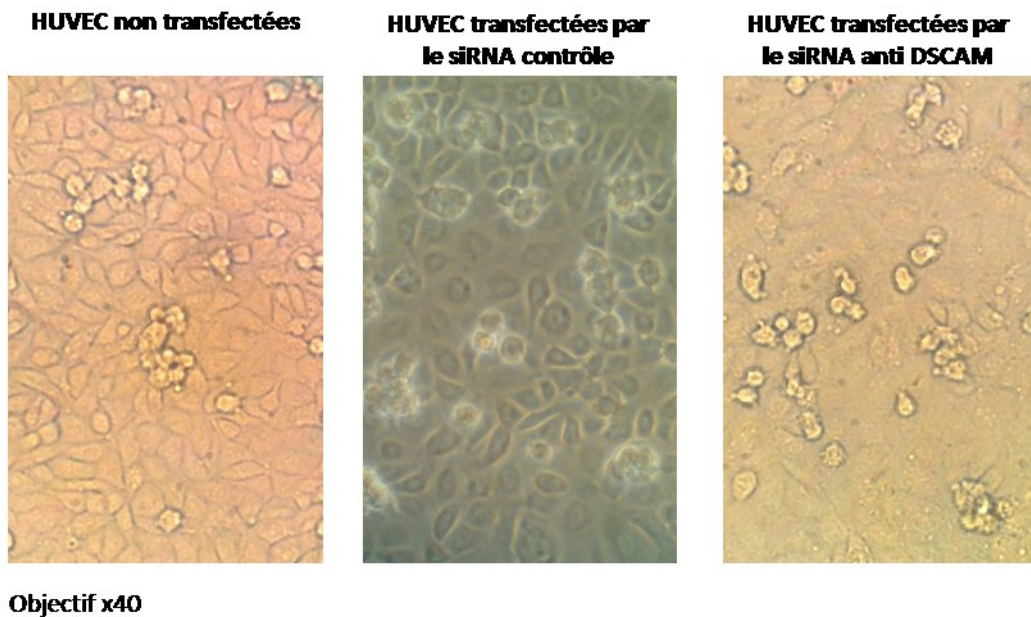




**Figure 24 :** Western Blot avec protéines extraites de cultures de cellules HUVEC transfectées avec différents siRNA. Conditions expérimentales : Gel à 6 % de polyacrylamide, dépôt de 30 ug de protéines, membrane de transfert : PVDF, marquage avec un anticorps anti-DSCAM dilué au 1/1000<sup>e</sup> (NBP1-59208, Novus Biologicals).

#### *Les paramètres cellulaires standards*

Les cellules transfectées avec le siRNA-4 anti-DSCAM n'ont montré aucune différence en termes de prolifération, mortalité ou morphotype, par rapport aux cellules transfectées avec le siRNA contrôle ou les cellules non transfectées. Néanmoins, nous avons observé un changement d'aspect au niveau des jonctions intercellulaires avec une impression d'effacement de ces jonctions dans les cellules transfectées avec le siRNA-4 anti-DSCAM (Figure 25). Cette observation doit être vérifiée avec un co-immunomarquage d'une protéine spécifique de la jonction (VE-Cadhérine) et de DSCAM.

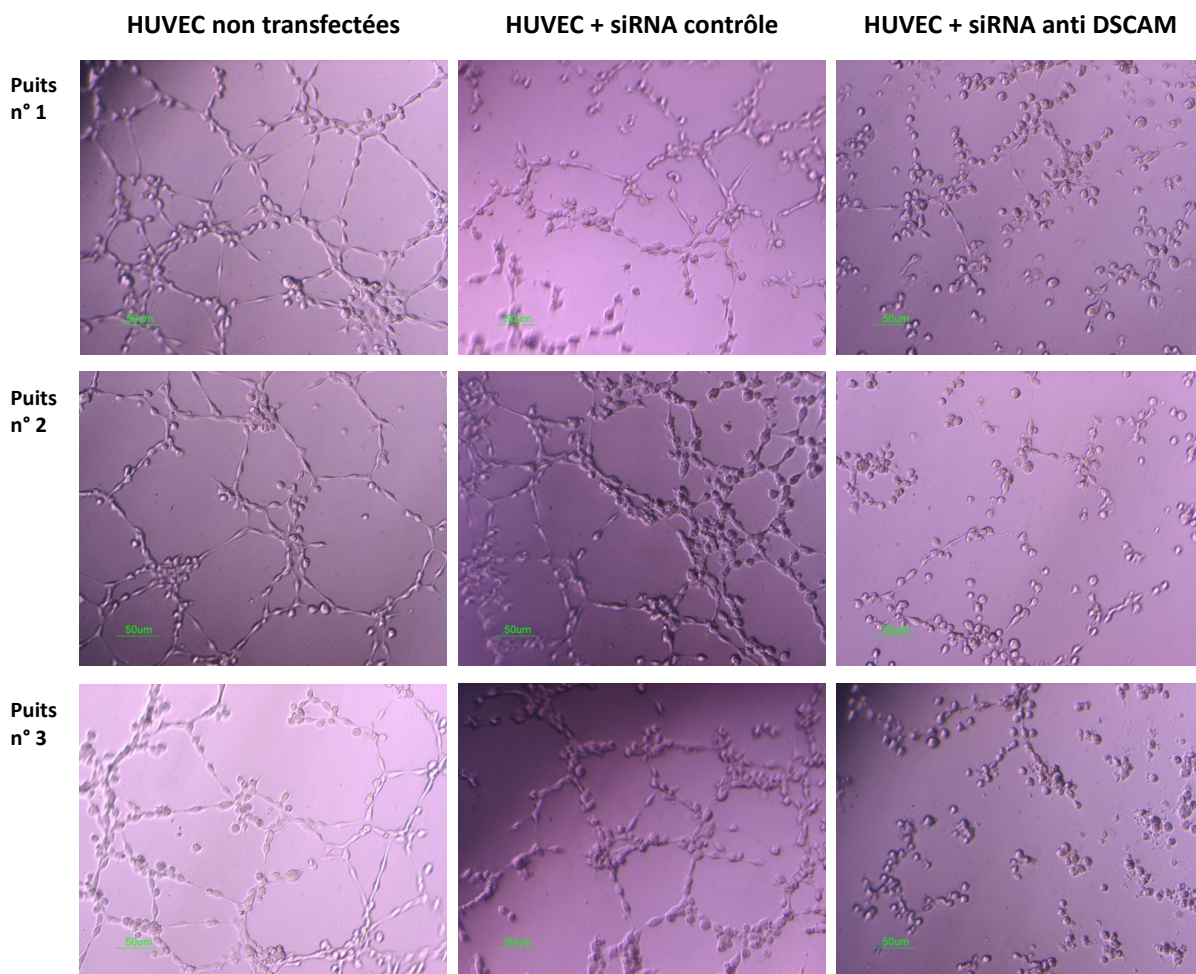


**Figure 25** : Observation microscopique des cultures confluentes des cellules HUVEC sous différentes conditions : absence de transfection, transfection avec un siRNA contrôle et transfection avec le siRNA 4 anti-DSCAM.

#### *Le test d'angiogenèse dans le Matrigel*

Le Matrigel est un milieu qui reproduit la matrice extracellulaire. Les cellules endothéliales et *a fortiori* les HUVEC sont capables de former une structure tridimensionnelle bien définie correspondant à des tubes vasculaires après quelques heures de culture dans ce milieu. Ce processus d'angiogenèse met en œuvre différentes fonctions cellulaires à savoir la migration et la traversée de la matrice extracellulaire, la prolifération, la reconnaissance intercellulaire et la capacité à former des vaisseaux/tubes.

Contrairement aux cellules non transfectées ou à celles transfectées avec un siRNA contrôle, les cellules transfectées avec le siRNA-4 anti-DSCAM ne forment pas de réseau tubulaire. Elles ont perdu leur propriété angiogénique. Seules de rares formations de tubes sont observées pouvant provenir d'une sous-population d'HUVEC non transfectées (Figure 26).

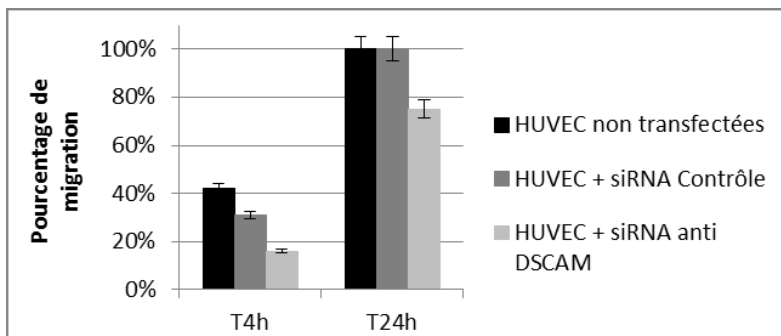
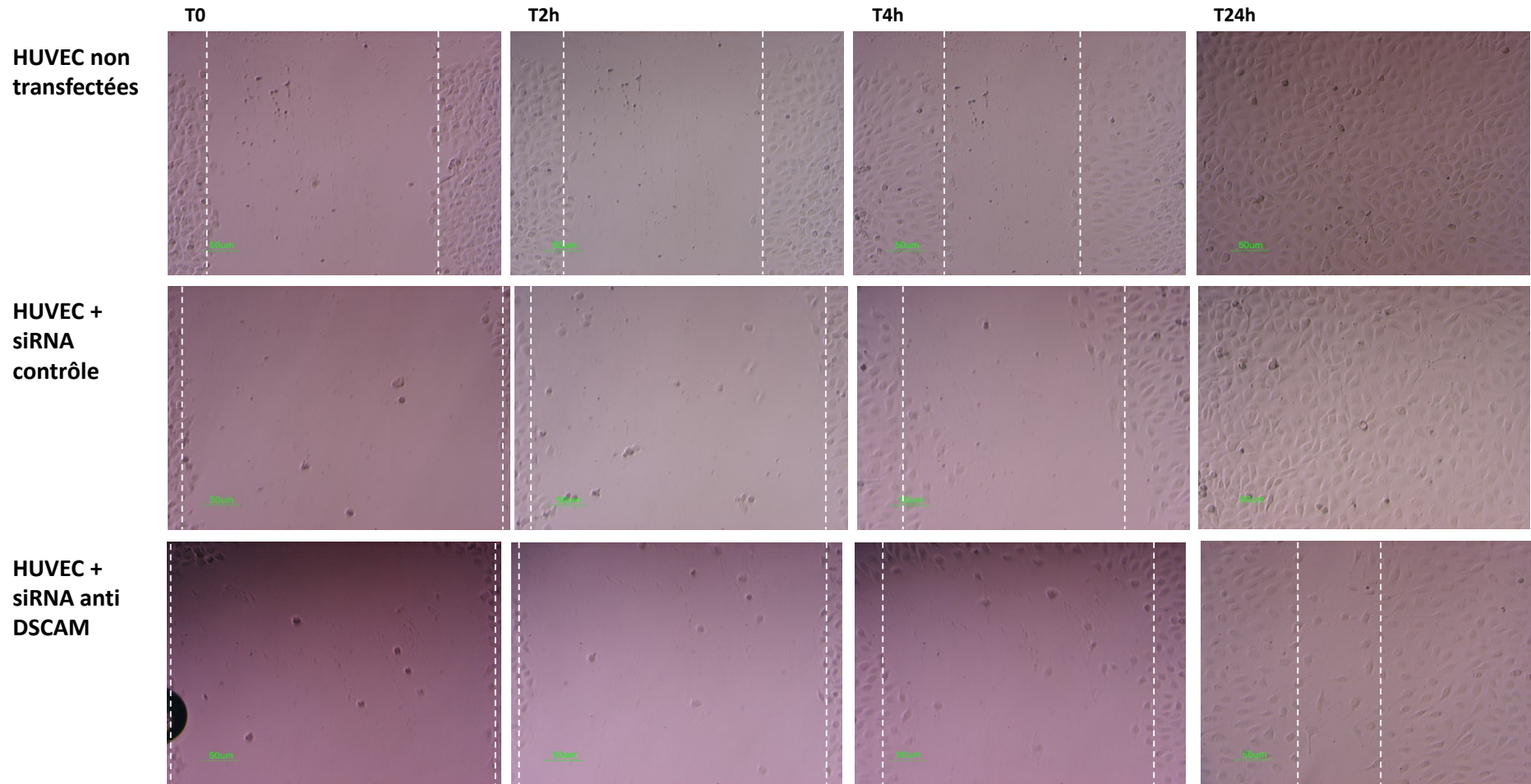


**Figure 26 :** Test d'angiogenèse/tubulogenèse dans le Matrigel. Test réalisé en triplicat pour chacune des conditions (HUVEC non transfectées, HUVEC transfectées avec un siRNA contrôle, HUVEC transfectées avec le siRNA 4 anti DSCAM).

### Test de la blessure

Ce test mesure la capacité de migration des cellules. Il consiste à créer une brèche dans le tapis cellulaire, puis évaluer sa fermeture dans le temps en présence d'un agent antiméiotique pour ne juger que la fermeture liée à la migration des cellules de part et d'autre de la brèche les unes vers les autres et non celle liée à leur prolifération.

Ce premier résultat montre une tendance à la diminution des capacités migratoires des cellules HUVEC transfectées avec le siRNA-4 anti-DSCAM (Figure 27).



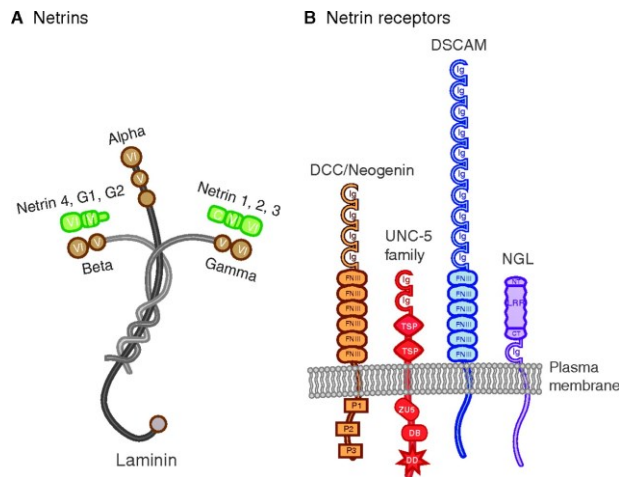
**Figure 27 :** Test de la blessure sur des cultures de cellules HUVEC non transfectées ou après transfection par un siRNA contrôle ou le siRNA anti DSCAM.

(A) Observations microscopiques de la blessure à différents temps. Objectif x20

(B) Comparaison quantitative des capacités de migration des cellules. Le pourcentage de migration est le ratio entre la différence entre les distances séparant les deux bords de la blessure, au temps T et au temps T0 et la distance à T0.

## 2.4. Etude des modèles murins KO *Dcc* et KO *Ntn1*

DSCAM est un récepteur trans-membranaire de la NTN1 qui se lie à d'autres récepteurs trans-membranaires de la même famille des immunoglobulines que DSCAM, notamment DCC (pour *Deleted in Colorectal Cancer*) (Figure 28). Il a été montré que DCC et DSCAM interagissent pour médier l'action de la NTN1 dans les cellules neuronales (Huang et al., 2015).



**Figure 28** : Structures de la Netrin et de ses récepteurs (Lai Wing Sun et al., 2011) .

Les souris knockout *Dscam* meurent 24h après la naissance d'un défaut respiratoire. Dans certains fonds génétiques, une mortalité prénatale est observée ou parfois un allongement de la gestation. La morphologie du placenta n'a sûrement pas été analysée. Ce modèle n'est pas disponible pour plus d'étude dans le contexte reproductif.

Nous avons profité de la disponibilité de modèles murins déficients en *Dcc* ou *Ntn1*, dans l'équipe d'Evelyne Bloch-Gallego à l'Institut, pour étudier l'éventuel impact d'un défaut de la voie d'action de *Dscam* sur le développement fœtal et placentaire. Nous avons ainsi analysé 19 fœtus et placentas issus de 4 portées de souris *Ntn1*<sup>+/-</sup> x *Ntn1*<sup>+/-</sup>, et 62 fœtus et placentas issus de 10 portées *Dcc*<sup>+/-</sup> x *Dcc*<sup>+/-</sup> (Tableaux 8 et 9).

Au total, le taux de résorption observé était de 21 % dans les portées *Ntn1* et 11 % dans les portées *Dcc*. La répartition des génotypes étaient de 11 % d'homozygotes +/+, 63 % d'hétérozygotes et 26 % d'homozygotes -/- pour les *Ntn1*, et 30,4 %, 48,2 % et 21,4 % respectivement pour les *Dcc*. Sur les 14 portées disséquées, seules les données issues des portées du stade E13,5 ont pu être analysées en raison d'un manque de représentativité des différents génotypes aux autres stades. La comparaison des poids des fœtus et placentas à E13,5 entre les différents génotypes n'a montré aucune

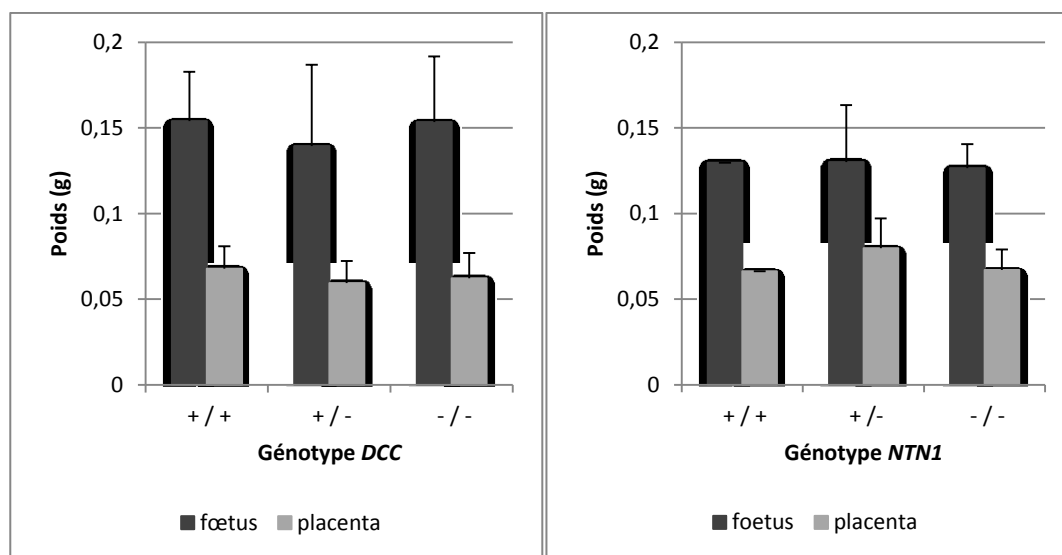
différence significative ni dans les portées *Ntn1* ni dans les portées *Dcc* (Figure 29).

**Tableau 8 : Caractéristiques des portées issues des croisements de souris *Dcc+/-* x *Dcc+/-*.**

Portée	A		B		C		D		G		H		I		J		K		N	
Stade	13,5		16,5		13,5		13,5		13,5		13,5		13,5		18,5		18,5		16,5	
N foetus	11		1		5		5		5		9		2		11		6		7	
N résorption	1		0		1		0		0		2		1		1		2ou3		0	
N +/+	4		Génotypage non réalisé		4		0		Génotypage non réalisé		3		1		3		0		2	
N +/-	5				1		4				2		1		5		4		5	
N -/-	2				0		1				4		0		3		2		0	
Poids (g)	Foetus Placenta		Foetus Placenta		Foetus Placenta		Foetus Placenta		Foetus Placenta		Foetus Placenta		Foetus Placenta		Foetus Placenta		Foetus Placenta		Foetus Placenta	
Génotypes +/+	0,1646 0,0626		0,0658 0,0389						0,2070 0,0670		0,1021 0,0560		1,0974 0,0947				0,4930 0,1313			
	0,1481 0,0760		0,0660 ?						0,1600 0,0676				1,0508 0,0881				0,3769 0,1025			
	0,1468 0,0816		0,0643 0,0253						0,1459 0,0857				1,2057 0,0881							
	0,1565 0,0462		0,0495 0,0260																	
Génotypes +/-	0,1670 0,0687		0,0544 0,0311		0,0969 0,0436				0,1837 0,0594		0,0922 0,0525		1,0273 0,0891		1,0992 0,0960		0,4405 0,1443			
	0,1774 0,0612				0,0718 0,0373				0,1873 0,0830				0,9330 0,0873		1,1274 0,0942		0,3495 0,1378			
	0,1764 0,0688				0,0962 0,0478								1,0214 0,0868		1,0268 0,0875		0,4094 0,1002			
	0,1755 0,0676				0,0763 0,0542								0,9128 0,0851		1,1224 0,0977		0,4287 0,1050			
	0,1728 0,0693												1,0117 0,0980				0,2285 0,0981			
Génotypes -/-	0,1822 0,0739				0,0886 0,0361				0,1485 0,0520				1,0304 0,0795		0,9557 0,0903					
	0,1452 0,0744								0,1940 0,0770				1,0195 0,0907		1,0550 0,0832					
									0,1269 0,0593				1,0209 0,0902							
									0,1890 0,0631											

**Tableau 9 :** Caractéristiques des portées issues des croisements de souris *Ntn1*<sup>+/-</sup> x *Ntn1*<sup>+/-</sup>.

Portée	L		M		O		P	
Stade	13,5		13,5		13,5		16,5	
N foetus	8		6		3		2	
N résorption	1		1		3		0	
N +/+	1		0		0		1	
N +/-	4		4		3		1	
N -/-	3		2		0		0	
Poids (g)	Foetus	Placenta	Foetus	Placenta	Foetus	Placenta	Foetus	Placenta
Génotypes +/+	0,1296	0,0662					0,6443	0,1727
	0,1453	0,0882	0,1487	0,0746	0,0960	0,0840	0,6128	0,0998
Génotypes +/-	0,1344	0,0859	0,1572	0,0868	0,1260	0,0982		
	0,0498	0,0334	0,1608	0,0946	0,1124	0,0827		
	0,1425	0,0800	0,1574	0,0695				
	0,1108	0,0596						
Génotypes -/-	0,1304	0,0605						
	0,1382	0,0809						



**Figure 29 :** Comparaison des poids des foetus et placentas des portées *Dcc* et *Ntn1* au stade E13,5.

En revanche, l'étude histologique des coupes de placentas (colorations HE, PAS et Trichrome de Masson) a révélé une désorganisation architecturale, une anomalie de maturation des villosités



avec une syncytialisation incomplète et un retard de différenciation trophoblastique. En revanche, aucune anomalie n'a été observée dans le spongiotrophoblaste et pour les cellules géantes.

## **DISCUSSION ET PERSPECTIVES**

## DISCUSSION

### I. L'approche utilisée pour l'identification de nouveaux GSE

Dans ce travail, nous avons cherché à identifier de nouveaux GSE dans le placenta humain en utilisant une stratégie précédemment développée dans l'équipe et qui a fait ses preuves (Barbaux et al. 2012) puisqu'elle avait permis la validation de 7 nouveaux GSE parmi 18 gènes candidats identifiés (cf Tableau 6). Nous avons poursuivi la validation de gènes candidats préalablement identifiés par cette méthode et sommes parvenus à valider l'empreinte pour deux autres gènes, *AIM1* et *DSCAM*, sur une liste de 11 gènes candidats (cf Tableau 6). La proportion de GSE validés dans la présente série est plus faible que dans la précédente (18 % versus 39 %). Ceci peut s'expliquer par le fait que la première série comprenait les « meilleurs » gènes candidats à savoir ceux pour lesquels il y avait le plus grand nombre de SNP à expression monoallélique et le plus de SNP exoniques. Néanmoins, ce pourcentage reste supérieur à celui obtenu avec d'autres méthodes d'identification de nouveaux GSE. Par ailleurs, il faut y ajouter l'identification de *NCAM2* à partir de la liste ciblée sur le chromosome 21 qui contenait 17 autres GSE candidats. Ce faible rendement de cette deuxième recherche s'explique par l'utilisation de scores moins stringents dans l'algorithme d'analyse.

Les puces de génotypage utilisées pour cette étude possèdent malheureusement un faible taux de SNP exoniques, la majorité des résultats ayant été obtenus grâce à des SNP en position intronique révélés par des messagers probablement non matures. L'évolution de la technologie permet désormais de disposer de puces assurant une bien meilleure couverture du génome (jusqu'à 6 millions de SNP), ciblant plus fréquemment les exons. Un des avantages de cette approche réside dans l'hybridation parallèle des ADNc et ADNg, foetal comme maternel, sur le même matériel. Le séquençage massif de l'ARN, autre approche de pointe envisagée pour découvrir de nouveaux GSE, doit être complété par un génotypage par une autre méthode. Dans les 2 approches, un point clé est l'estimation des taux relatifs de chaque allèle et la détermination d'un seuil signant une expression monoallélique ou au moins fortement déséquilibrée. Des améliorations informatiques sont probablement nécessaires pour mieux analyser ces données. Un écueil possible de la recherche de GSE concerne le tissu utilisé. Certains GSE sont très spécifiques d'un tissu et d'un temps donné (sac vitellin par exemple) et une étude exhaustive de divers tissus serait nécessaire pour découvrir tous les GSE, ce qui est impossible dans l'espèce humaine. La découverte de cas

d’empreinte circonscrits à des régions très réduites (trophoblaste villosus versus extra-villosus ou labyrinthe versus zone jonctionnelle) laisse aussi supposer que nos approches sur un tissu forcément complexe, aux types cellulaires nombreux comme le placenta, sont incapables de détecter ces phénomènes pointus. La question du nombre exact de gènes soumis à empreinte, de leur nature et spécificité, chez l’homme comme chez la souris, reste donc irrésolue, laissant encore de nombreuses inconnues sur le fonctionnement de ces gènes particuliers.

*AIM1* montre une empreinte isoforme-spécifique limitée à la forme la plus longue, comme cela a déjà été observé pour d’autres GSE tels que *Mest* (Blagitko et al., 2000; Kamei et al., 2007). Ce gène ayant été identifié comme GSE par une autre équipe au moment où nous menions cette étude (Das et al., 2013), nous avons décidé de privilégier l’étude de *DSCAM*.

## II. *DSCAM*, nouveau GSE

Nous avons montré que ce gène était un GSE à expression paternelle dans le placenta humain. Nous avons également mis en évidence l’existence d’une DMR dans l’intron 1. Le niveau de méthylation n’est pas exactement à 50 % comme le voudrait la théorie mais la variation du taux de méthylation expérimental par rapport au taux théorique est connue et peut être soit d’origine technique soit d’origine biologique (Jones and Liang, 2009; Laird et al., 2004). Les DMR sont une caractéristique constante des GSE car la méthylation différentielle est à la base de l’expression monoallélique que ce soit dans un cluster de GSE ou dans un GSE isolé (cf Section « Caractéristiques des GSE », p. 46). Dans les GSE, les DMR sont en général méthylées sur l’allèle réprimé. Nous n’avons pas pu déterminer l’origine parentale de l’allèle méthylé en raison de l’absence de SNP avec une fréquence d’hétérozygotie suffisante dans la région. En revanche, nous avons pu étudier l’état de méthylation de cette région dans les spermatozoïdes humains. Ce qui a montré une hypométhylation, déjà observée dans l’étude de Fang et al. (Fang et al., 2012). Par déduction, on peut raisonnablement penser que l’allèle méthylé dans le placenta provient de la mère ce qui est compatible avec l’origine paternelle de l’allèle exprimé observée dans nos résultats. L’analyse de la séquence de la DMR associée à *DSCAM* met en évidence la présence de sites de fixation pour CTCF et pour un autre GSE, *PLAGL1*. CTCF est une protéine pléiotropique dont l’une des fonctions est de participer à la régulation de l’empreinte en exerçant un effet insulateur et en organisant les boucles chromatiniques (cf Section « Régulation de l’empreinte – Deux grands modèles de régulation de l’empreinte », p. 46). *PLAGL1* code pour un facteur de transcription qui interagit avec de nombreux

GSE et gènes non soumis à l’empreinte dans le cadre du réseau de GSE dont les fonctions diverses concourent à contrôler la croissance embryonnaire (Varrault et al., 2006). Par ailleurs, le locus de *DSCAM* héberge deux ARN non codants, *DSCAM-AS1* et *DSCAM-IT1*, ce qui concorde également avec une autre caractéristique des GSE (cf Section « Caractéristiques des GSE », p. 46). Ces ARN non codants peuvent jouer un rôle dans la régulation de l’empreinte locale comme c’est le cas pour *H19* par exemple. Dans ce cas, ils sont également soumis à l’empreinte et exprimés à partir de l’allèle opposé à celui à partir duquel s’exprime le GSE sens. Nous avons tenté d’investiguer le statut d’empreinte de ces deux ARNnc mais n’avons pas réussi à amplifier l’ADNc pour le séquencer. Ceci est probablement dû à un trop faible niveau d’expression mais un faible niveau d’expression n’est pas un signe d’absence d’action.

### III. Le chromosome 21 et l’empreinte

Même si des GSE isolés existent, la plupart d’entre eux sont organisés en clusters de GSE distants de quelques mégabases. *DSCAM* est localisé sur le chromosome 21 en 21q22.2 et aucun GSE n’a été identifié à ce jour sur ce chromosome. Nous avons donc étendu la recherche d’autres GSE à tout le chromosome 21. Ce qui a abouti à l’identification d’un autre GSE très probable qui est *NCAM2*, *Neural Cell Adhesion Molecule*, et qui aurait une expression maternelle. Ce gène est situé à environ 19 Mb de *DSCAM*, ce qui n’est pas en faveur d’une régulation commune de l’empreinte dans le cadre d’un cluster mais plutôt de GSE isolés.

Pour *NCAM2*, nous avons observé deux situations de discordance : une expression biallélique pour un cas contre 19 cas monoalléliques et un cas d’expression monoallélique mais a priori d’origine paternelle et non maternelle comme les trois autres cas informatifs. Concernant le cas d’expression biallélique, il pourrait s’expliquer soit par une empreinte à variabilité inter-individuelle comme cela a déjà été rapporté pour d’autres GSE tels que *PEG3* ou *MEST*-isoforme 2 (Lambertini et al., 2008; Tycko, 2006), soit par une situation pathologique de perte de l’empreinte qui peut être induite par de nombreuses causes endogènes ou exogènes récapitulées dans la Section « Pathologies humaines liées à l’empreinte parentale » (p. 69), ou alors par une contamination, lors du prélèvement du placenta, par la décidue maternelle où le gène pourrait s’exprimer de façon biallélique. Deux causes ont d’ores et déjà pu être éliminées par la revue des données cliniques, à savoir la PE et le RCIU. La contamination par la décidue maternelle reste très peu probable en raison des précautions prises au moment du prélèvement en prélevant en surface de la face fœtale. Concernant la discordance

sur l'origine parentale de l'allèle exprimé, il faut rappeler que dans notre méthode de détermination de l'origine parentale de l'allèle exprimé nous procédons par déduction car seul le matériel génétique de la mère est disponible. Par conséquent, seules les situations où la mère est homozygote pour le SNP étudié nous permettent de déterminer l'origine de l'allèle exprimé dans le placenta. L'expression *a priori* paternelle de *NCAM2* est en réalité un cas où l'allèle exprimé n'est pas celui de la mère. Plusieurs hypothèses ont été envisagées : une erreur de prélèvement, une grossesse issue d'un don d'ovocyte non déclaré par la mère ou non renseigné dans le formulaire de l'étude, ou encore une altération de l'empreinte. Il se pourrait également que l'expression monoallélique de ce GSE se fasse aléatoirement à partir de l'allèle maternel ou paternel même dans un même tissu. Nous avons pu réaliser un génotypage des échantillons d'ADNg placentaire, d'ADNg maternel et d'ARN placentaire (à la recherche de traces minimales d'ADN placentaire), par des marqueurs microsatellites. Nous avons confirmé la concordance entre ces trois échantillons ce qui élimine toute cause d'erreur d'échantillon ou d'identité parentale. Par ailleurs, aucune pathologie de la grossesse n'est connue pour cet individu. Ces données soutiennent donc l'hypothèse d'une variabilité inter-individuelle de l'origine de l'allèle exprimé. Ce phénomène est rare mais a déjà été rapporté pour *ZFAT* par exemple (Metsalu et al., 2014). Même si une anomalie de l'empreinte dans cet échantillon ne peut être exclue.

*NCAM2* appartient à la même famille de protéines que *DSCAM*. Cette famille n'avait pas été représentée dans le groupe des GSE jusqu'ici. Or, un travail récent de Al Adhami et al. (Al Adhami et al., 2015), de la même équipe que Varrault et al. à l'origine du réseau de GSE (Varrault et al., 2006), rapporte une sur-représentation inattendue dans le réseau de GSE de gènes non soumis à l'empreinte liés à la matrice extra-cellulaire, l'adhésion cellulaire et la communication inter-cellulaire en faveur d'une nouvelle voie d'action des GSE. *DSCAM* et *NCAM2* pourraient donc jouer un rôle dans le réseau via cette voie là.

Pour deux autres gènes candidats sur le chromosome 21, *SAMSN1* et *KRTAP27-1*, l'expression monoallélique est retrouvée chez 50 % des individus hétérozygotes ce qui pourrait signer une empreinte à variabilité inter-individuelle mais le nombre d'individus hétérozygotes est trop faible pour conclure. Une étude sur une population plus large aiderait à répondre à la question.

Ces résultats sont inédits puisqu'il n'y avait pas de locus connu comme étant soumis à l'empreinte sur le chromosome 21. Par ailleurs, il n'existe pas d'observations cliniques dans la littérature en

faveur de pathologies cliniques associées à des anomalies chromosomiques sur le 21 avec un effet parental laissant suspecter une éventuelle empreinte. Aucun cas pathologique associé à une disomie uniparentale (DUP) du chromosome 21 n'a été décrit. Néanmoins, Handerson et al. s'étaient posés la question de l'existence de GSE sur le chromosome 21 (Henderson et al., 1994) suite à une étude visant à rechercher de façon systématique des situations de DUP dans des produits de fausses-couches. La DUP du chromosome 21 (DUP21) avait été retrouvée dans deux cas. Les auteurs ont alors posé l'hypothèse de l'existence de GSE importants pour le développement dans le chromosome 21. Cette hypothèse a ensuite été remise en question par des rapports de sujets sains portant une DUP21, maternelle pour un cas (Rogan et al., 1999) et paternelle pour un autre (Blouin et al., 1993). Dans le cas de la DUP21 paternelle (Blouin et al., 1993), les auteurs ont décrit un phénotype normal avec des données à la naissance, y compris le poids, qui étaient normales. En revanche, pour le sujet portant la DUP21 maternelle (Rogan et al., 1999), les auteurs ont rapporté un phénotype normal mais aucune donnée notamment sur le poids n'a été communiquée. Dans les deux cas, aucun examen du placenta n'avait été réalisé. Les délétions/duplications de la région sont quant à elles très rares et dans ces rares cas l'origine parentale de l'anomalie n'a pas été recherchée ce qui rend impossible la recherche d'un éventuel effet parental. Enfin, la question se pose également pour les situations de Trisomie 21 où l'on pourrait s'attendre à une différence entre le phénotype associé à une Trisomie d'origine paternelle et celui associé à une Trisomie d'origine maternelle, mais là encore il y a très peu d'études comparatives du phénotype de la Trisomie 21 selon l'origine parentale du chromosome 21 surnuméraire (Muranjan et al., 2010; Stoll et al., 1995). Très peu de cas de Trisomie 21 d'origine paternelle ont pu être analysés en raison de la fréquence largement majoritaire des cas d'origine maternelle (près de 95 %) et en plus ces études ne se sont intéressées qu'au phénotype postnatal. Une équipe a analysé l'origine parentale des Trisomies 21 diagnostiquées en prénatal et a conclu, de façon intéressante, à une fréquence plus élevée des fœtus portant un chromosome 21 surnuméraire d'origine paternelle par rapport au pourcentage en postnatal (Muller et al., 2000). Les pertes fœtales seraient donc plus importantes lorsque la Trisomie 21 est d'origine paternelle. Or, nous avons montré que *DSCAM* est un GSE à expression paternelle. Un lien pourrait donc exister entre la surexpression de *DSCAM* dans les Trisomies 21 d'origine paternelle et la mortalité fœtale plus élevée. Aucune des études ne s'est malheureusement intéressée au placenta.

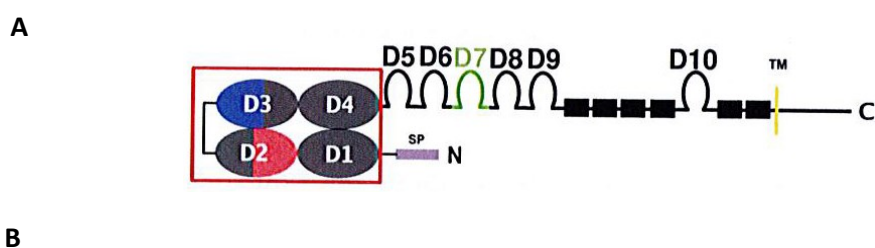
## IV. Rôle de *DSCAM* dans le placenta ?

### 1. Fonctions connues de *DSCAM*

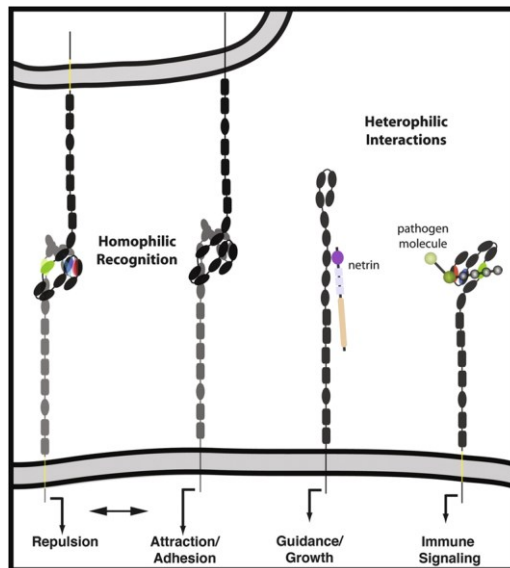
Ce gène (OMIM 602523) code pour une molécule d'adhésion qui appartient à une nouvelle classe de molécules d'adhésion des cellules neurales de la superfamille des Immunoglobulines. La structure de cette protéine a été essentiellement étudiée chez la Drosophile et met en évidence une protéine à plusieurs domaines : un domaine intracytoplasmique dans la partie C-terminale, un domaine transmembranaire, et une partie extracellulaire composée de 10 domaines Immunoglobuline-like, 6 domaines Fibronectine type III et un peptide signal à l'extrémité N-terminale (Figure 30a).

*DSCAM* est largement exprimée dans le système nerveux central au cours du développement (Yamakawa et al., 1998) et serait à l'origine de la déficience intellectuelle, des anomalies viscérales ainsi que des anomalies cardiaques associés à la Trisomie 21. La sévérité de la déficience intellectuelle étant variable et les anomalies viscérales et cardiaques inconstantes, il serait intéressant de comparer les phénotypes en fonction de l'origine parentale du chromosome 21 surnuméraire.

*DSCAM* est connue pour son rôle dans le guidage des neurones lors du développement du système nerveux central. Le modèle le plus étudié est la Drosophile qui possède plusieurs dizaines de milliers d'isoformes de *DSCAM* issues d'un épissage alternatif. Cette diversité importante d'isoformes garantirait la spécificité des connexions neuronales en permettant aux neurones exprimant la même isoforme de se reconnaître (interactions homophiliques) pour participer au même processus (« self-self recognition »), et en empêchant les neurones ayant des isoformes différentes d'interagir les uns avec les autres (« self-self avoidance ») (Hattori et al., 2007). Ces interactions cellulaires participent à la migration des neurones et à la croissance axonale, processus qui sont indispensables à l'arborisation et l'établissement du réseau neuronal mais aussi au développement de la rétine (Fuerst et al., 2008).







**Figure 30** : Représentations schématiques (A) de la structure du récepteur *Dscam* (Meijers et al., 2007) et (B) des interactions homophiles et hétérophiles possibles (Schmucker et Chen, 2009).

Contrairement à la *Drosophile*, seules trois isoformes ont été identifiées chez l'Homme et une seule chez la souris ([www.ensembl.org](http://www.ensembl.org)). Ces quelques isoformes n'assurent donc pas la même spécificité cellulaire que les près de 40000 de la *Drosophile*. Ce qui n'empêche pas DSCAM d'avoir des fonctions équivalentes d'interactions homophiliques participant à la migration et à l'établissement des projections neuronales, mais surtout des interactions hétérophiliques avec son Ligand, la Netrin-1 (NTN1). NTN1 est un chemoattractant qui se lie à plusieurs récepteurs dont DSCAM. Lors du développement du système nerveux central, NTN1 guide les projections axonales (Ly et al., 2008). Elle est également sécrétée au niveau de la plaque neurale lors de la gastrulation pour attirer les axones vers le tube neural. Il a été montré que NTN1 est également impliquée dans l'angiogenèse (Larrivee et al., 2007; Lu et al., 2004; Wilson et al., 2006).

## 2. Hypothèses de fonctions placentaires de *DSCAM*

Le placenta étant dépourvu de cellules nerveuses, l'action de *DSCAM* passerait donc par un autre type cellulaire mettant en œuvre des voies de signalisation communes à celles de l'innervation, *via* des processus d'adhésion, prolifération et mobilité cellulaire. Néanmoins, ce rôle n'a pas encore été documenté dans la littérature.

Les souris *Dscam*<sup>-/-</sup> dans un fond génétique de C57BL/6 décèdent en général dans les 24 h suivant la naissance en raison d'un problème respiratoire lié à un dysfonctionnement des neurones qui contrôlent la contraction du diaphragme (Amano et al., 2009). Dans un autre fond génétique le

knockdown de *Dscam* n'est pas létal mais des anomalies respiratoires du même type persistent. Un rallongement du temps gestationnel a également été rapporté. En revanche, les données sur l'état macroscopique et microscopique du placenta n'ont pas été récoltées. Une anomalie placentaire n'est donc pas exclue. Pour preuve, la placentomégalie chez les souris déficientes en *H19* n'a été observée qu'après ré-examen des souris dans une seconde publication (Keniry et al., 2012).

Dans ce travail, nous avons observé que *DSCAM* s'exprime dans le syncytiotrophoblaste et l'endothélium des vaisseaux de l'axe mésenchymateux des villosités. *DSCAM* pourrait donc exercer son action en participant soit à la mise en place de l'angiogenèse placentaire via son expression dans les cellules endothéliales soit à la formation du placenta même via l'expression dans le trophoblaste, voire même avoir les deux fonctions. Son expression dans le syncytiotrophoblaste et l'endothélium qui sont à l'interface des échanges foeto-maternels est compatible avec une éventuelle fonction dans ces échanges comme c'est le cas de nombreux GSE.

### **2.1. Rôle de *DSCAM* dans le développement de la structure placentaire**

Dans ce travail, l'étude du statut d'empreinte de *Dscam* dans le placenta murin a mis en évidence une empreinte limitée à la zone jonctionnelle. Cette divergence de l'empreinte entre les deux compartiments (labyrinthe et zone jonctionnelle) peut être la conséquence d'une perte de l'empreinte au niveau du labyrinthe dans un intérêt évolutif ou bien une acquisition de l'empreinte nécessaire à la différenciation de cellules de la zone jonctionnelle, d'autant plus que les deux compartiments n'ont pas la même origine embryologique (Maltepe et al., 2010). Le rôle de l'empreinte dans la différenciation cellulaire a déjà été suggéré pour *CTNNA3* qui a une expression monoallélique maternelle dans le cytotrophoblaste villositaire du placenta humain et une expression biallélique dans le trophoblaste extra-villositaire, en faveur de l'implication de l'empreinte dans la différenciation trophoblastique (van Dijk et al., 2004). D'un point de vue fonctionnel, l'absence de l'empreinte dans le labyrinthe qui est le lieu des échanges foeto-maternels et sa présence dans la zone jonctionnelle dont les fonctions restent encore mal connues mais qui participe à la structure placentaire soutient la fait que *Dscam* appartiendrait au groupe de GSE à fonction placentaire impliqués dans le développement de la structure du placenta et non ses fonctions physiologiques, comme c'est le cas d'autres GSE tels que *Ascl2* et *Peg10* (Guillemot et al., 1994; Ono et al., 2006). Ce qui n'empêche pas qu'un défaut d'un GSE de ce groupe puisse entraîner secondairement un défaut du labyrinthe et donc altérer les fonctions physiologiques du placenta. Dans le cas de *Ascl2/Mash2*,

les souris déficientes montrent une altération des deux compartiments mais des expériences sur embryons chimériques avec un mélange de cellules mutantes et de cellules normales sont en faveur d'un rôle de ce GSE uniquement dans la structure de la zone jonctionnelle suggérant que le défaut du labyrinthe est un phénomène secondaire (Tanaka et al., 1997).

Dans le modèle de cellules trophoblastiques primaires humaines en culture, nous n'avons pas observé de différence d'expression de *DSCAM* avant et après syncytialisation. Des essais sur un modèle immortalisé de cellules BeWo, syncytialisant sous l'effet de la Forskolin, n'ont pas été plus concluants. Ce gène n'interviendrait donc pas dans la fusion des cytotrophoblastes en syncytiotrophoblaste mais il pourrait jouer un rôle dans la différenciation ou la migration cellulaire nécessaire à l'invasion trophoblastique de la décidue maternelle. De rares données issues d'études de placentas portant une trisomie 21 montrent des anomalies de la maturation des villosités et de la différenciation trophoblastique (Pidoux et al., 2012). Ces anomalies pourraient être liées à *DSCAM* ou à d'autres protéines exprimées à partir d'autres gènes du chromosome 21 ou même d'autres chromosomes. Des études plus approfondies des cohortes de placentas avec trisomie 21 sont nécessaires.

## **2.2. Rôle de *DSCAM* dans l'angiogenèse placentaire**

Le réseau vasculaire et le réseau nerveux suivent des trajectoires parallèles suggérant qu'ils sont attirés par un même attractant. Park et al. ont suggéré que le chemoattractant commun est NTN1 qui est le ligand de *DSCAM* (Park et al., 2004). NTN1 stimule la vasculogénèse de façon spécifique comparé aux facteurs de croissance type VEGF, en stimulant la migration, la prolifération et l'adhésion des cellules endothéliales et des cellules musculaires lisses. La voie moléculaire par laquelle NTN1 va médier ses fonctions dans l'angiogenèse n'est pas encore totalement élucidée. Des interactions avec son récepteur UNC5B semblent induire une inhibition de l'angiogenèse contrôlant ainsi le guidage du réseau vasculaire (Larrivee et al., 2007; Lu et al., 2004). Récemment, CD146 a été identifié comme un nouveau récepteur de NTN1 qui stimule l'angiogenèse (Tu et al., 2015). Un travail récent de Xie et al. suggère que NTN1 aurait un effet proangiogénique dans le placenta (Xie et al., 2011). Les auteurs ont analysé l'effet d'une surexpression ou une sous-expression de NTN1 sur les fonctions endothéliales *in vitro* et *in vivo*, en transfectant des cellules HUVEC et des placentas de rat pendant la gestation. La sous-expression de NTN1 entraînait une diminution de la prolifération, migration et de la capacité de tubulogénèse des HUVEC, et une

diminution de la densité vasculaire dans les placentas transfectés ainsi que la baisse des poids placentaires et fœtaux. La sur-expression induisait l'effet inverse. Nous avons posé le postulat que DSCAM pourrait avoir une action sur l'angiogenèse dans le placenta par son interaction avec NTN1. Nos résultats préliminaires sur les cellules HUVEC sont en faveur de cette hypothèse puisque l'inhibition de DSCAM diminue les capacités migratoire et de tubulogenèse des cellules endothéliales HUVEC. Et ce même en présence d'un milieu conditionné contenant de la NTN1 qui normalement stimule l'angiogenèse. Ce qui soutient l'hypothèse que DSCAM jouerait un rôle proangiogénique dans la vasculogenèse placentaire *via* NTN1.

Au final, le rôle de *DSCAM* dans le placenta suggéré par nos résultats est compatible avec la théorie du conflit parental puisqu'il s'exprime à partir de l'allèle paternel et qu'il stimulerait l'angiogenèse et participerait à la structure placentaire. Ces deux axes d'action potentiels promeuvent la croissance foeto-placentaire. Or, c'est la fonction des GSE à expression paternelle.

### **2.3. DSCAM et pathologies placentaires**

Dans notre cohorte, nous n'avons pas trouvé de différence d'expression de DSCAM entre les placentas issus de grossesses pathologiques (RCIU et PE) et les placentas issus de grossesses normales. Or, ces pathologies peuvent être associées à des anomalies de l'empreinte. Notre résultat peut s'expliquer par l'hétérogénéité des étiologies qui entraînent ces pathologies. En outre, il faut signaler que vue la taille des groupes étudiés nous n'avons pas été en mesure d'apparier les échantillons en fonction du sexe fœtal, du terme exact de l'accouchement, de l'âge maternel, etc. Or, ces facteurs confondants peuvent impacter les niveaux d'expression voire l'état de méthylation.

## PERSPECTIVES

L'approche utilisée dans ce travail mériterait d'être mise en œuvre sur des données issues de nouvelles puces de génotypage qui sont plus adaptées pour l'analyse de l'ARN, qui contiennent bien plus de SNP (6 millions contre 250000) dont un plus grand nombre localisé dans des régions exoniques par rapport aux puces utilisées dans ce travail, et qui permettent un volume d'échantillons analysés plus important et donc une analyse simultanée des trios (ADNg placentaire, ADNc placentaire et ADNg maternel) de plusieurs individus.

Ce travail a initié la caractérisation d'un nouveau GSE dans le placenta humain, *DSCAM*. Etant donné que la souris KO *Dscam* existe, son étude approfondie en portant un intérêt particulier à caractériser la croissance fœtale et placentaire ainsi que l'histologie placentaire aiderait à mieux comprendre la fonction de ce GSE. La place de *DSCAM* dans le réseau des GSE pourrait également être investiguée en analysant par exemple son expression en cas de déficience d'un GSE tel que *H19*. Un autre aspect à explorer est l'expression de *DSCAM* dans des placentas issus de grossesses obtenus par assistance médicale à la procréation.

Concernant le rôle de *DSCAM* dans le placenta, la question reste à élucider. Nos résultats préliminaires sont en faveur d'un rôle pro-angiogénique mais qui a besoin d'être confirmé par des expériences supplémentaires d'inhibition par siRNA anti-*DSCAM*, éventuellement combinée à des expériences de transfection de placentas murins *in vivo*. L'analyse de l'impact sur l'expression des marqueurs de l'angiogenèse (VEGF, Tie2, vWF) serait également utile. Par ailleurs, l'étude histologique des placentas portant un chromosome 21 surnuméraire dont l'origine parentale est connue ainsi que des placentas avec DUP(21) aurait un grand intérêt même si cela est difficilement réalisable en raison des autorisations nécessaires et des difficultés de recrutement.

Il serait également intéressant d'approfondir l'étude des autres GSE potentiels révélés par cette étude. Pour *SAMSN1* et *KRTAP27-1*, un élargissement du nombre d'individus analysés, pour savoir si ce sont bien des GSE à variabilité inter-individuelle, pourra être envisagé grâce à la poursuite des inclusions de nouveaux prélèvements dans l'étude. Pour *NCAM2* également, un élargissement de la taille de la cohorte étudiée pourra aider à confirmer la variabilité inter-individuelle de l'origine parentale de l'allèle exprimé.

# **BIBLIOGRAPHIE**

## BIBLIOGRAPHIE

- Abramowitz, L.K., and Bartolomei, M.S. (2012). Genomic imprinting: recognition and marking of imprinted loci. *Current opinion in genetics & development* 22, 72-78.
- Adalsteinsson, B.T., and Ferguson-Smith, A.C. (2014). Epigenetic control of the genome-lessons from genomic imprinting. *Genes* 5, 635-655.
- Al Adhami, H., Evano, B., Le Digarcher, A., Gueydan, C., Dubois, E., Parrinello, H., Dantec, C., Bouschet, T., Varrault, A., and Journot, L. (2015). A systems-level approach to parental genomic imprinting: the imprinted gene network includes extracellular matrix genes and regulates cell cycle exit and differentiation. *Genome research* 25, 353-367.
- Amano, K., Fujii, M., Arata, S., Tojima, T., Ogawa, M., Morita, N., Shimohata, A., Furuichi, T., Itohara, S., Kamiguchi, H., *et al.* (2009). DSCAM deficiency causes loss of pre-inspiratory neuron synchronicity and perinatal death. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 29, 2984-2996.
- Angiolini, E., Fowden, A., Coan, P., Sandovici, I., Smith, P., Dean, W., Burton, G., Tycko, B., Reik, W., Sibley, C., *et al.* (2006). Regulation of placental efficiency for nutrient transport by imprinted genes. *Placenta* 27 Suppl A, S98-102.
- Arnaud, P., Monk, D., Hitchins, M., Gordon, E., Dean, W., Beechey, C.V., Peters, J., Craigen, W., Preece, M., Stanier, P., *et al.* (2003). Conserved methylation imprints in the human and mouse GRB10 genes with divergent allelic expression suggests differential reading of the same mark. *Human molecular genetics* 12, 1005-1019.
- Babak, T., DeVeale, B., Tsang, E.K., Zhou, Y., Li, X., Smith, K.S., Kukurba, K.R., Zhang, R., Li, J.B., van der Kooy, D., *et al.* (2015). Genetic conflict reflected in tissue-specific maps of genomic imprinting in human and mouse. *Nature genetics* 47, 544-549.
- Baran, Y., Subramaniam, M., Biton, A., Tukiainen, T., Tsang, E.K., Rivas, M.A., Pirinen, M., Gutierrez-Arcelus, M., Smith, K.S., Kukurba, K.R., *et al.* (2015). The landscape of genomic imprinting across diverse adult human tissues. *Genome research* 25, 927-936.
- Barboux, S., Gascoin-Lachambre, G., Buffat, C., Monnier, P., Mondon, F., Tonanny, M.B., Pinard, A., Auer, J., Bessieres, B., Barlier, A., *et al.* (2012). A genome-wide approach reveals novel imprinted genes expressed in the human placenta. *Epigenetics : official journal of the DNA Methylation Society* 7, 1079-1090.
- Barlow, D.P., and Bartolomei, M.S. (2014). Genomic imprinting in mammals. *Cold Spring Harbor perspectives in biology* 6.
- Bartolomei, M.S. (2009). Genomic imprinting: employing and avoiding epigenetic processes. *Genes & development* 23, 2124-2133.
- Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., 3rd, Gingeras, T.R., *et al.* (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* 120, 169-181.
- Bernstein, B.E., Meissner, A., and Lander, E.S. (2007). The mammalian epigenome. *Cell* 128, 669-681.
- Bestor, T.H. (1992). Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *The EMBO journal* 11, 2611-2617.
- Bestor, T.H. (2000). The DNA methyltransferases of mammals. *Human molecular genetics* 9, 2395-2402.
- Bian, C., and Yu, X. (2014). PGC7 suppresses TET3 for protecting DNA methylation. *Nucleic acids research* 42, 2893-2905.
- Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes & development* 16, 6-21.

- Bjornsson, H.T., Albert, T.J., Ladd-Acosta, C.M., Green, R.D., Rongione, M.A., Middle, C.M., Irizarry, R.A., Broman, K.W., and Feinberg, A.P. (2008). SNP-specific array-based allele-specific expression analysis. *Genome research* 18, 771-779.
- Blagitko, N., Mergenthaler, S., Schulz, U., Wollmann, H.A., Craigen, W., Eggermann, T., Ropers, H.H., and Kalscheuer, V.M. (2000). Human GRB10 is imprinted and expressed from the paternal and maternal allele in a highly tissue- and isoform-specific fashion. *Human molecular genetics* 9, 1587-1595.
- Blouin, J.L., Avramopoulos, D., Pangalos, C., and Antonarakis, S.E. (1993). Normal phenotype with paternal uniparental isodisomy for chromosome 21. *American journal of human genetics* 53, 1074-1078.
- Bostick, M., Kim, J.K., Esteve, P.O., Clark, A., Pradhan, S., and Jacobsen, S.E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* 317, 1760-1764.
- Bourc'his, D., and Bestor, T.H. (2004). Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* 431, 96-99.
- Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B., and Bestor, T.H. (2001). Dnmt3L and the establishment of maternal genomic imprints. *Science* 294, 2536-2539.
- Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., *et al.* (2005). The transcriptional landscape of the mammalian genome. *Science* 309, 1559-1563.
- Cattanach, B.M. (1986). Parental origin effects in mice. *Journal of embryology and experimental morphology* 97 Suppl, 137-150.
- Cattanach, B.M., and Kirk, M. (1985). Differential activity of maternally and paternally derived chromosome regions in mice. *Nature* 315, 496-498.
- Causeret, F., Hidalgo-Sanchez, M., Fort, P., Backer, S., Popoff, M.R., Gauthier-Rouviere, C., and Bloch-Gallego, E. (2004). Distinct roles of Rac1/Cdc42 and Rho/Rock for axon outgrowth and nucleokinesis of precerebellar neurons toward netrin 1. *Development* 131, 2841-2852.
- Cedar, H., and Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. *Nature reviews Genetics* 10, 295-304.
- Charalambous, M., da Rocha, S.T., and Ferguson-Smith, A.C. (2007). Genomic imprinting, growth control and the allocation of nutritional resources: consequences for postnatal life. *Current opinion in endocrinology, diabetes, and obesity* 14, 3-12.
- Chen, T., Chen, D., Li, F., and Tan, Z. (2014). Netrin-1 with stem cells promote angiogenesis in limb ischemic rats. *The Journal of surgical research* 192, 664-669.
- Chen, T., Ueda, Y., Xie, S., and Li, E. (2002). A novel Dnmt3a isoform produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation. *The Journal of biological chemistry* 277, 38746-38754.
- Chotalia, M., Smallwood, S.A., Ruf, N., Dawson, C., Lucifero, D., Frontera, M., James, K., Dean, W., and Kelsey, G. (2009). Transcription is required for establishment of germline methylation marks at imprinted genes. *Genes & development* 23, 105-117.
- Choux, C., Carmignac, V., Bruno, C., Sagot, P., Vaiman, D., and Fauque, P. (2015). The placenta: phenotypic and epigenetic modifications induced by Assisted Reproductive Technologies throughout pregnancy. *Clinical epigenetics* 7, 87.
- Chuang, L.S., Ian, H.I., Koh, T.W., Ng, H.H., Xu, G., and Li, B.F. (1997). Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science* 277, 1996-2000.
- Ciccone, D.N., Su, H., Hevi, S., Gay, F., Lei, H., Bajko, J., Xu, G., Li, E., and Chen, T. (2009). KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature* 461, 415-418.



- Constancia, M., Hemberger, M., Hughes, J., Dean, W., Ferguson-Smith, A., Fundele, R., Stewart, F., Kelsey, G., Fowden, A., Sibley, C., *et al.* (2002). Placental-specific IGF-II is a major modulator of placental and fetal growth. *Nature* *417*, 945-948.
- Curley, J.P., Pinnock, S.B., Dickson, S.L., Thresher, R., Miyoshi, N., Surani, M.A., and Keverne, E.B. (2005). Increased body fat in mice with a targeted mutation of the paternally expressed imprinted gene *Peg3*. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* *19*, 1302-1304.
- Dai, L., Peng, C., Montellier, E., Lu, Z., Chen, Y., Ishii, H., Debernardi, A., Buchou, T., Rousseaux, S., Jin, F., *et al.* (2014). Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark. *Nature chemical biology* *10*, 365-370.
- Das, R., Lee, Y.K., Strogantsev, R., Jin, S., Lim, Y.C., Ng, P.Y., Lin, X.M., Chng, K., Yeo, G., Ferguson-Smith, A.C., *et al.* (2013). DNMT1 and AIM1 Imprinting in human placenta revealed through a genome-wide screen for allele-specific DNA methylation. *BMC genomics* *14*, 685.
- De Koning, L., Corpet, A., Haber, J.E., and Almouzni, G. (2007). Histone chaperones: an escort network regulating histone traffic. *Nature structural & molecular biology* *14*, 997-1007.
- DeBaun, M.R., Niemitz, E.L., and Feinberg, A.P. (2003). Association of in vitro fertilization with Beckwith-Wiedemann syndrome and epigenetic alterations of LIT1 and H19. *American journal of human genetics* *72*, 156-160.
- Delaval, K., Govin, J., Cerqueira, F., Rousseaux, S., Khochbin, S., and Feil, R. (2007). Differential histone modifications mark mouse imprinting control regions during spermatogenesis. *The EMBO journal* *26*, 720-729.
- Di, X., Matsuzaki, H., Webster, T.A., Hubbell, E., Liu, G., Dong, S., Bartell, D., Huang, J., Chiles, R., Yang, G., *et al.* (2005). Dynamic model based algorithms for screening and genotyping over 100 K SNPs on oligonucleotide microarrays. *Bioinformatics* *21*, 1958-1963.
- Ding, Q., Liao, S.J., and Yu, J. (2014). Axon guidance factor netrin-1 and its receptors regulate angiogenesis after cerebral ischemia. *Neuroscience bulletin* *30*, 683-691.
- Diplas, A.I., Lambertini, L., Lee, M.J., Sperling, R., Lee, Y.L., Wetmur, J., and Chen, J. (2009). Differential expression of imprinted genes in normal and IUGR human placentas. *Epigenetics : official journal of the DNA Methylation Society* *4*, 235-240.
- Engel, N., Thorvaldsen, J.L., and Bartolomei, M.S. (2006). CTCF binding sites promote transcription initiation and prevent DNA methylation on the maternal allele at the imprinted H19/*Igf2* locus. *Human molecular genetics* *15*, 2945-2954.
- English, C.M., Adkins, M.W., Carson, J.J., Churchill, M.E., and Tyler, J.K. (2006). Structural basis for the histone chaperone activity of *Asf1*. *Cell* *127*, 495-508.
- Fang, F., Hodges, E., Molaro, A., Dean, M., Hannon, G.J., and Smith, A.D. (2012). Genomic landscape of human allele-specific DNA methylation. *Proceedings of the National Academy of Sciences of the United States of America* *109*, 7332-7337.
- Feinberg, A.P., and Tycko, B. (2004). The history of cancer epigenetics. *Nature reviews Cancer* *4*, 143-153.
- Ferguson-Smith, A.C., Moore, T., Detmar, J., Lewis, A., Hemberger, M., Jammes, H., Kelsey, G., Roberts, C.T., Jones, H., and Constancia, M. (2006). Epigenetics and imprinting of the trophoblast -- a workshop report. *Placenta* *27 Suppl A*, S122-126.
- Ferron, S.R., Charalambous, M., Radford, E., McEwen, K., Wildner, H., Hind, E., Morante-Redolat, J.M., Laborda, J., Guillemot, F., Bauer, S.R., *et al.* (2011). Postnatal loss of *Dlk1* imprinting in stem cells and niche astrocytes regulates neurogenesis. *Nature* *475*, 381-385.

- Fitzpatrick, G.V., Soloway, P.D., and Higgins, M.J. (2002). Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. *Nature genetics* 32, 426-431.
- Font de Mora, J., Esteban, L.M., Burks, D.J., Nunez, A., Garces, C., Garcia-Barrado, M.J., Iglesias-Osma, M.C., Moratinos, J., Ward, J.M., and Santos, E. (2003). Ras-GRF1 signaling is required for normal beta-cell development and glucose homeostasis. *The EMBO journal* 22, 3039-3049.
- Fuerst, P.G., Koizumi, A., Masland, R.H., and Burgess, R.W. (2008). Neurite arborization and mosaic spacing in the mouse retina require DSCAM. *Nature* 451, 470-474.
- Gao, W.L., Li, D., Xiao, Z.X., Liao, Q.P., Yang, H.X., Li, Y.X., Ji, L., and Wang, Y.L. (2011). Detection of global DNA methylation and paternally imprinted H19 gene methylation in preeclamptic placentas. *Hypertension research : official journal of the Japanese Society of Hypertension* 34, 655-661.
- Gardiner-Garden, M., and Frommer, M. (1987). CpG islands in vertebrate genomes. *Journal of molecular biology* 196, 261-282.
- Garfield, A.S., Cowley, M., Smith, F.M., Moorwood, K., Stewart-Cox, J.E., Gilroy, K., Baker, S., Xia, J., Dalley, J.W., Hurst, L.D., *et al.* (2011). Distinct physiological and behavioural functions for parental alleles of imprinted Grb10. *Nature* 469, 534-538.
- Gicquel, C., Gaston, V., Mandelbaum, J., Siffroi, J.P., Flahault, A., and Le Bouc, Y. (2003). In vitro fertilization may increase the risk of Beckwith-Wiedemann syndrome related to the abnormal imprinting of the KCN1OT gene. *American journal of human genetics* 72, 1338-1341.
- Gimelbrant, A., Hutchinson, J.N., Thompson, B.R., and Chess, A. (2007). Widespread monoallelic expression on human autosomes. *Science* 318, 1136-1140.
- Graham, C.F. (1974). The production of parthenogenetic mammalian embryos and their use in biological research. *Biological reviews of the Cambridge Philosophical Society* 49, 399-424.
- Grewal, S.I., and Moazed, D. (2003). Heterochromatin and epigenetic control of gene expression. *Science* 301, 798-802.
- Groth, A., Rocha, W., Verreault, A., and Almouzni, G. (2007). Chromatin challenges during DNA replication and repair. *Cell* 128, 721-733.
- Guillemot, F., Nagy, A., Auerbach, A., Rossant, J., and Joyner, A.L. (1994). Essential role of Mash-2 in extraembryonic development. *Nature* 371, 333-336.
- Guo, L., Choufani, S., Ferreira, J., Smith, A., Chitayat, D., Shuman, C., Uxa, R., Keating, S., Kingdom, J., and Weksberg, R. (2008). Altered gene expression and methylation of the human chromosome 11 imprinted region in small for gestational age (SGA) placentae. *Developmental biology* 320, 79-91.
- Hajkova, P. (2010). Epigenetic reprogramming--taking a lesson from the embryo. *Current opinion in cell biology* 22, 342-350.
- Hajkova, P., Ancelin, K., Waldmann, T., Lacoste, N., Lange, U.C., Cesari, F., Lee, C., Almouzni, G., Schneider, R., and Surani, M.A. (2008). Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature* 452, 877-881.
- Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J., and Surani, M.A. (2002). Epigenetic reprogramming in mouse primordial germ cells. *Mechanisms of development* 117, 15-23.
- Hammoud, S.S., Nix, D.A., Zhang, H., Purwar, J., Carrell, D.T., and Cairns, B.R. (2009). Distinctive chromatin in human sperm packages genes for embryo development. *Nature* 460, 473-478.
- Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 405, 486-489.
- Hattori, D., Demir, E., Kim, H.W., Viragh, E., Zipursky, S.L., and Dickson, B.J. (2007). Dscam diversity is essential for neuronal wiring and self-recognition. *Nature* 449, 223-227.

- Haycock, P.C., and Ramsay, M. (2009). Exposure of mouse embryos to ethanol during preimplantation development: effect on DNA methylation in the h19 imprinting control region. *Biology of reproduction* *81*, 618-627.
- Heijmans, B.T., Tobi, E.W., Stein, A.D., Putter, H., Blauw, G.J., Susser, E.S., Slagboom, P.E., and Lumey, L.H. (2008). Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proceedings of the National Academy of Sciences of the United States of America* *105*, 17046-17049.
- Hemberger, M. (2007). Epigenetic landscape required for placental development. *Cellular and molecular life sciences : CMLS* *64*, 2422-2436.
- Henckel, A., and Arnaud, P. (2010). Genome-wide identification of new imprinted genes. *Brief Funct Genomics*.
- Henckel, A., Chebli, K., Kota, S.K., Arnaud, P., and Feil, R. (2012). Transcription and histone methylation changes correlate with imprint acquisition in male germ cells. *The EMBO journal* *31*, 606-615.
- Henckel, A., Nakabayashi, K., Sanz, L.A., Feil, R., Hata, K., and Arnaud, P. (2009). Histone methylation is mechanistically linked to DNA methylation at imprinting control regions in mammals. *Human molecular genetics* *18*, 3375-3383.
- Henderson, D.J., Sherman, L.S., Loughna, S.C., Bennett, P.R., and Moore, G.E. (1994). Early embryonic failure associated with uniparental disomy for human chromosome 21. *Human molecular genetics* *3*, 1373-1376.
- Hirasawa, R., Chiba, H., Kaneda, M., Tajima, S., Li, E., Jaenisch, R., and Sasaki, H. (2008). Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes & development* *22*, 1607-1616.
- Hiura, H., Sugawara, A., Ogawa, H., John, R.M., Miyauchi, N., Miyanari, Y., Horiike, T., Li, Y., Yaegashi, N., Sasaki, H., *et al.* (2010). A tripartite paternally methylated region within the Gpr1-Zdbf2 imprinted domain on mouse chromosome 1 identified by meDIP-on-chip. *Nucleic acids research* *38*, 4929-4945.
- Holliday, R., and Pugh, J.E. (1975). DNA modification mechanisms and gene activity during development. *Science* *187*, 226-232.
- Hu, Y.G., Hirasawa, R., Hu, J.L., Hata, K., Li, C.L., Jin, Y., Chen, T., Li, E., Rigolet, M., Viegas-Pequignot, E., *et al.* (2008). Regulation of DNA methylation activity through Dnmt3L promoter methylation by Dnmt3 enzymes in embryonic development. *Human molecular genetics* *17*, 2654-2664.
- Huang, H., Shao, Q., Qu, C., Yang, T., Dwyer, T., and Liu, G. (2015). Coordinated interaction of Down syndrome cell adhesion molecule and deleted in colorectal cancer with dynamic TUBB3 mediates Netrin-1-induced axon branching. *Neuroscience* *293*, 109-122.
- Iglesias-Platas, I., Martin-Trujillo, A., Petazzi, P., Guillaumet-Adkins, A., Esteller, M., and Monk, D. (2014). Altered expression of the imprinted transcription factor PLAGL1 deregulates a network of genes in the human IUGR placenta. *Human molecular genetics* *23*, 6275-6285.
- Illingworth, R.S., and Bird, A.P. (2009). CpG islands--'a rough guide'. *FEBS letters* *583*, 1713-1720.
- Ishida, M., and Moore, G.E. (2013). The role of imprinted genes in humans. *Molecular aspects of medicine* *34*, 826-840.
- Itier, J.M., Tremp, G.L., Leonard, J.F., Multon, M.C., Ret, G., Schweighoffer, F., Tocque, B., Bluet-Pajot, M.T., Cormier, V., and Dautry, F. (1998). Imprinted gene in postnatal growth role. *Nature* *393*, 125-126.
- Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. *Science* *293*, 1074-1080.
- Jia, D., Jurkowska, R.Z., Zhang, X., Jeltsch, A., and Cheng, X. (2007). Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature* *449*, 248-251.
- Jones, P.A., and Liang, G. (2009). Rethinking how DNA methylation patterns are maintained. *Nature reviews Genetics* *10*, 805-811.

- Joubert, B.R., Haberg, S.E., Nilsen, R.M., Wang, X., Vollset, S.E., Murphy, S.K., Huang, Z., Hoyo, C., Middtun, O., Cupul-Uicab, L.A., *et al.* (2012). 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. *Environmental health perspectives* *120*, 1425-1431.
- Kagami, M., Nagai, T., Fukami, M., Yamazawa, K., and Ogata, T. (2007). Silver-Russell syndrome in a girl born after in vitro fertilization: partial hypermethylation at the differentially methylated region of PEG1/MEST. *Journal of assisted reproduction and genetics* *24*, 131-136.
- Kajii, T., and Ohama, K. (1977). Androgenetic origin of hydatidiform mole. *Nature* *268*, 633-634.
- Kamei, Y., Suganami, T., Kohda, T., Ishino, F., Yasuda, K., Miura, S., Ezaki, O., and Ogawa, Y. (2007). Peg1/Mest in obese adipose tissue is expressed from the paternal allele in an isoform-specific manner. *FEBS letters* *581*, 91-96.
- Kanayama, N., Takahashi, K., Matsuura, T., Sugimura, M., Kobayashi, T., Moniwa, N., Tomita, M., and Nakayama, K. (2002). Deficiency in p57Kip2 expression induces preeclampsia-like symptoms in mice. *Molecular human reproduction* *8*, 1129-1135.
- Kanduri, C. (2015). Long noncoding RNAs: Lessons from genomic imprinting. *Biochimica et biophysica acta*.
- Kanduri, C., Pant, V., Loukinov, D., Pugacheva, E., Qi, C.F., Wolffe, A., Ohlsson, R., and Lobanenkov, V.V. (2000). Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. *Current biology : CB* *10*, 853-856.
- Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* *429*, 900-903.
- Katari, S., Turan, N., Bibikova, M., Erinle, O., Chalian, R., Foster, M., Gaughan, J.P., Coutifaris, C., and Sapienza, C. (2009). DNA methylation and gene expression differences in children conceived in vitro or in vivo. *Human molecular genetics* *18*, 3769-3778.
- Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., *et al.* (2005). Antisense transcription in the mammalian transcriptome. *Science* *309*, 1564-1566.
- Kato, Y., Kaneda, M., Hata, K., Kumaki, K., Hisano, M., Kohara, Y., Okano, M., Li, E., Nozaki, M., and Sasaki, H. (2007). Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. *Human molecular genetics* *16*, 2272-2280.
- Keniry, A., Oxley, D., Monnier, P., Kyba, M., Dandolo, L., Smits, G., and Reik, W. (2012). The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. *Nature cell biology* *14*, 659-665.
- Keverne, E.B. (2015). Genomic imprinting, action, and interaction of maternal and fetal genomes. *Proceedings of the National Academy of Sciences of the United States of America* *112*, 6834-6840.
- Khorkova, O., Hsiao, J., and Wahlestedt, C. (2015). Basic biology and therapeutic implications of lncRNA. *Advanced drug delivery reviews* *87*, 15-24.
- Kim, J.D., Kim, H., Ekram, M.B., Yu, S., Faulk, C., and Kim, J. (2011). Rex1/Zfp42 as an epigenetic regulator for genomic imprinting. *Human molecular genetics* *20*, 1353-1362.
- Koerner, M.V., Pauler, F.M., Huang, R., and Barlow, D.P. (2009). The function of non-coding RNAs in genomic imprinting. *Development* *136*, 1771-1783.
- Korostowski, L., Raval, A., Breuer, G., and Engel, N. (2011). Enhancer-driven chromatin interactions during development promote escape from silencing by a long non-coding RNA. *Epigenetics & chromatin* *4*, 21.
- Koukoura, O., Sifakis, S., Soufla, G., Zaravinos, A., Apostolidou, S., Jones, A., Widschwendter, M., and Spandidos, D.A. (2011). Loss of imprinting and aberrant methylation of IGF2 in placentas from pregnancies complicated with fetal growth restriction. *International journal of molecular medicine* *28*, 481-487.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* *128*, 693-705.

- Kumar, N., Leverence, J., Bick, D., and Sampath, V. (2012). Ontogeny of growth-regulating genes in the placenta. *Placenta* 33, 94-99.
- Lai Wing Sun, K., Correia, J.P., and Kennedy, T.E. (2011). Netrins: versatile extracellular cues with diverse functions. *Development* 138, 2153-2169.
- Laird, C.D., Pleasant, N.D., Clark, A.D., Sneed, J.L., Hassan, K.M., Manley, N.C., Vary, J.C., Jr., Morgan, T., Hansen, R.S., and Stoger, R. (2004). Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. *Proceedings of the National Academy of Sciences of the United States of America* 101, 204-209.
- Lambertini, L., Diplas, A.I., Lee, M.J., Sperling, R., Chen, J., and Wetmur, J. (2008). A sensitive functional assay reveals frequent loss of genomic imprinting in human placenta. *Epigenetics : official journal of the DNA Methylation Society* 3, 261-269.
- Lambertini, L., Marsit, C.J., Sharma, P., Maccani, M., Ma, Y., Hu, J., and Chen, J. (2012). Imprinted gene expression in fetal growth and development. *Placenta* 33, 480-486.
- Larrivee, B., Freitas, C., Trombe, M., Lv, X., Delafarge, B., Yuan, L., Bouvree, K., Breant, C., Del Toro, R., Brechot, N., *et al.* (2007). Activation of the UNC5B receptor by Netrin-1 inhibits sprouting angiogenesis. *Genes & development* 21, 2433-2447.
- Larsen, F., Gundersen, G., Lopez, R., and Prydz, H. (1992). CpG islands as gene markers in the human genome. *Genomics* 13, 1095-1107.
- Latos, P.A., Pauler, F.M., Koerner, M.V., Senergin, H.B., Hudson, Q.J., Stocsits, R.R., Allhoff, W., Stricker, S.H., Klement, R.M., Warczok, K.E., *et al.* (2012). Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. *Science* 338, 1469-1472.
- Lau, M.M., Stewart, C.E., Liu, Z., Bhatt, H., Rotwein, P., and Stewart, C.L. (1994). Loss of the imprinted IGF2/cation-independent mannose 6-phosphate receptor results in fetal overgrowth and perinatal lethality. *Genes & development* 8, 2953-2963.
- Lee, J., Inoue, K., Ono, R., Ogonuki, N., Kohda, T., Kaneko-Ishino, T., Ogura, A., and Ishino, F. (2002). Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells. *Development* 129, 1807-1817.
- Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K., *et al.* (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* 125, 301-313.
- Lei, H., Oh, S.P., Okano, M., Juttermann, R., Goss, K.A., Jaenisch, R., and Li, E. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* 122, 3195-3205.
- Leonhardt, H., Page, A.W., Weier, H.U., and Bestor, T.H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* 71, 865-873.
- Lewis, A., Mitsuya, K., Umlauf, D., Smith, P., Dean, W., Walter, J., Higgins, M., Feil, R., and Reik, W. (2004). Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nature genetics* 36, 1291-1295.
- Li, E., Beard, C., and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. *Nature* 366, 362-365.
- Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69, 915-926.
- Li, S., Sukeena, J.M., Simmons, A.B., Hansen, E.J., Nuhn, R.E., Samuels, I.S., and Fuerst, P.G. (2015). DSCAM promotes refinement in the mouse retina through cell death and restriction of exploring dendrites. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 35, 5640-5654.

- Li, X., Ito, M., Zhou, F., Youngson, N., Zuo, X., Leder, P., and Ferguson-Smith, A.C. (2008). A maternal-zygotic effect gene, *Zfp57*, maintains both maternal and paternal imprints. *Developmental cell* *15*, 547-557.
- Liang, F., Diao, L., Liu, J., Jiang, N., Zhang, J., Wang, H., Zhou, W., Huang, G., and Ma, D. (2014). Paternal ethanol exposure and behavioral abnormalities in offspring: associated alterations in imprinted gene methylation. *Neuropharmacology* *81*, 126-133.
- Lin, S.P., Youngson, N., Takada, S., Seitz, H., Reik, W., Paulsen, M., Cavaille, J., and Ferguson-Smith, A.C. (2003). Asymmetric regulation of imprinting on the maternal and paternal chromosomes at the *Dlk1-Gtl2* imprinted cluster on mouse chromosome 12. *Nature genetics* *35*, 97-102.
- Linder, D., McCaw, B.K., and Hecht, F. (1975). Parthenogenic origin of benign ovarian teratomas. *The New England journal of medicine* *292*, 63-66.
- Liu, Y.J., Nakamura, T., and Nakano, T. (2012). Essential role of *DPPA3* for chromatin condensation in mouse oocytogenesis. *Biology of reproduction* *86*, 40.
- Lu, X., Le Noble, F., Yuan, L., Jiang, Q., De Lafarge, B., Sugiyama, D., Breant, C., Claes, F., De Smet, F., Thomas, J.L., *et al.* (2004). The netrin receptor *UNC5B* mediates guidance events controlling morphogenesis of the vascular system. *Nature* *432*, 179-186.
- Lubinsky, M., Herrmann, J., Kosseff, A.L., and Opitz, J.M. (1974). Letter: Autosomal-dominant sex-dependent transmission of the Wiedemann-Beckwith syndrome. *Lancet* *1*, 932.
- Lucifero, D., Mertineit, C., Clarke, H.J., Bestor, T.H., and Trasler, J.M. (2002). Methylation dynamics of imprinted genes in mouse germ cells. *Genomics* *79*, 530-538.
- Luedi, P.P., Dietrich, F.S., Weidman, J.R., Bosko, J.M., Jirtle, R.L., and Hartemink, A.J. (2007). Computational and experimental identification of novel human imprinted genes. *Genome research* *17*, 1723-1730.
- Luedi, P.P., Hartemink, A.J., and Jirtle, R.L. (2005). Genome-wide prediction of imprinted murine genes. *Genome research* *15*, 875-884.
- Ly, A., Nikolaev, A., Suresh, G., Zheng, Y., Tessier-Lavigne, M., and Stein, E. (2008). *DSCAM* is a netrin receptor that collaborates with *DCC* in mediating turning responses to netrin-1. *Cell* *133*, 1241-1254.
- Ma, D., Shield, J.P., Dean, W., Leclerc, I., Knauf, C., Burcelin, R.R., Rutter, G.A., and Kelsey, G. (2004). Impaired glucose homeostasis in transgenic mice expressing the human transient neonatal diabetes mellitus locus, *TNDM*. *The Journal of clinical investigation* *114*, 339-348.
- Mackay, D.J., Callaway, J.L., Marks, S.M., White, H.E., Acerini, C.L., Boonen, S.E., Dayanikli, P., Firth, H.V., Goodship, J.A., Haemers, A.P., *et al.* (2008). Hypomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in *ZFP57*. *Nature genetics* *40*, 949-951.
- Maeda, N., and Hayashizaki, Y. (2006). Genome-wide survey of imprinted genes. *Cytogenetic and genome research* *113*, 144-152.
- Maltepe, E., Bakardjiev, A.I., and Fisher, S.J. (2010). The placenta: transcriptional, epigenetic, and physiological integration during development. *The Journal of clinical investigation* *120*, 1016-1025.
- Mancini-Dinardo, D., Steele, S.J., Levorse, J.M., Ingram, R.S., and Tilghman, S.M. (2006). Elongation of the *Kcnq1ot1* transcript is required for genomic imprinting of neighboring genes. *Genes & development* *20*, 1268-1282.
- Mao, Z., Xia, W., Chang, H., Huo, W., Li, Y., and Xu, S. (2015). Paternal BPA exposure in early life alters *Igf2* epigenetic status in sperm and induces pancreatic impairment in rat offspring. *Toxicology letters* *238*, 30-38.
- Marcho, C., Bevilacqua, A., Tremblay, K.D., and Mager, J. (2015). Tissue-specific regulation of *Igf2r/Airn* imprinting during gastrulation. *Epigenetics & chromatin* *8*, 10.
- Margueron, R., Trojer, P., and Reinberg, D. (2005). The key to development: interpreting the histone code? *Current opinion in genetics & development* *15*, 163-176.

- Markunas, C.A., Xu, Z., Harlid, S., Wade, P.A., Lie, R.T., Taylor, J.A., and Wilcox, A.J. (2014). Identification of DNA methylation changes in newborns related to maternal smoking during pregnancy. *Environmental health perspectives* 122, 1147-1153.
- Masemola, M.L., van der Merwe, L., Lombard, Z., Viljoen, D., and Ramsay, M. (2015). Reduced DNA methylation at the PEG3 DMR and KvDMR1 loci in children exposed to alcohol in utero: a South African Fetal Alcohol Syndrome cohort study. *Frontiers in genetics* 6, 85.
- Maynard, K.R., and Stein, E. (2012). DSCAM contributes to dendrite arborization and spine formation in the developing cerebral cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 32, 16637-16650.
- McCann, J.A., Zheng, H., Islam, A., Goodyer, C.G., and Polychronakos, C. (2001). Evidence against GRB10 as the gene responsible for Silver-Russell syndrome. *Biochemical and biophysical research communications* 286, 943-948.
- McGrath, J., and Solter, D. (1984). Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* 37, 179-183.
- McMinn, J., Wei, M., Schupf, N., Cusmai, J., Johnson, E.B., Smith, A.C., Weksberg, R., Thaker, H.M., and Tycko, B. (2006). Unbalanced placental expression of imprinted genes in human intrauterine growth restriction. *Placenta* 27, 540-549.
- Meijers, R., Puettmann-Holgado, R., Skiniotis, G., Liu, J.H., Walz, T., Wang, J.H., and Schmucker, D. (2007). Structural basis of Dscam isoform specificity. *Nature* 449, 487-491.
- Messerschmidt, D.M., de Vries, W., Ito, M., Solter, D., Ferguson-Smith, A., and Knowles, B.B. (2012). Trim28 is required for epigenetic stability during mouse oocyte to embryo transition. *Science* 335, 1499-1502.
- Metsalu, T., Viltrop, T., Tiirats, A., Rajashekar, B., Reimann, E., Koks, S., Rull, K., Milani, L., Acharya, G., Basnet, P., *et al.* (2014). Using RNA sequencing for identifying gene imprinting and random monoallelic expression in human placenta. *Epigenetics : official journal of the DNA Methylation Society* 9, 1397-1409.
- Mileva, G., Baker, S.L., Konkle, A.T., and Bielajew, C. (2014). Bisphenol-A: epigenetic reprogramming and effects on reproduction and behavior. *International journal of environmental research and public health* 11, 7537-7561.
- Moore, T., and Haig, D. (1991). Genomic imprinting in mammalian development: a parental tug-of-war. *Trends in genetics : TIG* 7, 45-49.
- Muller, F., Rebiffe, M., Taillandier, A., Oury, J.F., and Mornet, E. (2000). Parental origin of the extra chromosome in prenatally diagnosed fetal trisomy 21. *Human genetics* 106, 340-344.
- Muranjan, M., Chaudhari, T., and Vundinti, B.R. (2010). Phenotypic heterogeneity and parental origin of extra chromosome 21 in Down syndrome. *Indian pediatrics* 47, 429-432.
- Nagano, T., Mitchell, J.A., Sanz, L.A., Pauler, F.M., Ferguson-Smith, A.C., Feil, R., and Fraser, P. (2008). The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* 322, 1717-1720.
- Nakamura, T., Arai, Y., Umehara, H., Masuhara, M., Kimura, T., Taniguchi, H., Sekimoto, T., Ikawa, M., Yoneda, Y., Okabe, M., *et al.* (2007). PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nature cell biology* 9, 64-71.
- Nakamura, T., Liu, Y.J., Nakashima, H., Umehara, H., Inoue, K., Matoba, S., Tachibana, M., Ogura, A., Shinkai, Y., and Nakano, T. (2012). PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* 486, 415-419.
- Nakashima, H., Kimura, T., Kaga, Y., Nakatani, T., Seki, Y., Nakamura, T., and Nakano, T. (2013). Effects of dppa3 on DNA methylation dynamics during primordial germ cell development in mice. *Biology of reproduction* 88, 125.

- Natsume, R., Eitoku, M., Akai, Y., Sano, N., Horikoshi, M., and Senda, T. (2007). Structure and function of the histone chaperone CIA/ASF1 complexed with histones H3 and H4. *Nature* *446*, 338-341.
- Nelissen, E.C., van Montfoort, A.P., Dumoulin, J.C., and Evers, J.L. (2011). Epigenetics and the placenta. *Human reproduction update* *17*, 397-417.
- Oakey, R.J., and Beechey, C.V. (2002). Imprinted genes: identification by chromosome rearrangements and post-genomic strategies. *Trends in genetics : TIG* *18*, 359-366.
- Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* *99*, 247-257.
- Okano, M., Xie, S., and Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature genetics* *19*, 219-220.
- Ono, R., Nakamura, K., Inoue, K., Naruse, M., Usami, T., Wakisaka-Saito, N., Hino, T., Suzuki-Migishima, R., Ogonuki, N., Miki, H., *et al.* (2006). Deletion of Peg10, an imprinted gene acquired from a retrotransposon, causes early embryonic lethality. *Nature genetics* *38*, 101-106.
- Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.P., Allis, C.D., *et al.* (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* *448*, 714-717.
- Pandey, R.R., Mondal, T., Mohammad, F., Enroth, S., Redrup, L., Komorowski, J., Nagano, T., Mancini-Dinardo, D., and Kanduri, C. (2008). Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Molecular cell* *32*, 232-246.
- Park, K.W., Crouse, D., Lee, M., Karnik, S.K., Sorensen, L.K., Murphy, K.J., Kuo, C.J., and Li, D.Y. (2004). The axonal attractant Netrin-1 is an angiogenic factor. *Proceedings of the National Academy of Sciences of the United States of America* *101*, 16210-16215.
- Peters, J., Holmes, R., Monk, D., Beechey, C.V., Moore, G.E., and Williamson, C.M. (2006). Imprinting control within the compact Gnas locus. *Cytogenetic and genome research* *113*, 194-201.
- Peters, J., and Williamson, C.M. (2008). Control of imprinting at the Gnas cluster. *Advances in experimental medicine and biology* *626*, 16-26.
- Phillips, J.E., and Corces, V.G. (2009). CTCF: master weaver of the genome. *Cell* *137*, 1194-1211.
- Pidoux, G., Gerbaud, P., Cocquebert, M., Segond, N., Badet, J., Fournier, T., Guibourdenche, J., and Evain-Brion, D. (2012). Review: Human trophoblast fusion and differentiation: lessons from trisomy 21 placenta. *Placenta* *33* Suppl, S81-86.
- Pires, N.D., and Grossniklaus, U. (2014). Different yet similar: evolution of imprinting in flowering plants and mammals. *F1000prime reports* *6*, 63.
- Plagge, A., Gordon, E., Dean, W., Boiani, R., Cinti, S., Peters, J., and Kelsey, G. (2004). The imprinted signaling protein XL alpha s is required for postnatal adaptation to feeding. *Nature genetics* *36*, 818-826.
- Pollard, K.S., Serre, D., Wang, X., Tao, H., Grundberg, E., Hudson, T.J., Clark, A.G., and Frazer, K. (2008). A genome-wide approach to identifying novel-imprinted genes. *Human genetics* *122*, 625-634.
- Popp, C., Dean, W., Feng, S., Cokus, S.J., Andrews, S., Pellegrini, M., Jacobsen, S.E., and Reik, W. (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* *463*, 1101-1105.
- Pozharny, Y., Lambertini, L., Ma, Y., Ferrara, L., Litton, C.G., Diplas, A., Jacobs, A.R., Chen, J., Stone, J.L., Wetmur, J., *et al.* (2010). Genomic loss of imprinting in first-trimester human placenta. *American journal of obstetrics and gynecology* *202*, 391 e391-398.
- Pray-Grant, M.G., Daniel, J.A., Schieltz, D., Yates, J.R., 3rd, and Grant, P.A. (2005). Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* *433*, 434-438.



- Prickett, A.R., and Oakey, R.J. (2012). A survey of tissue-specific genomic imprinting in mammals. *Molecular genetics and genomics* : MGG 287, 621-630.
- Probst, A.V., Dunleavy, E., and Almouzni, G. (2009). Epigenetic inheritance during the cell cycle. *Nature reviews Molecular cell biology* 10, 192-206.
- Quenneville, S., Verde, G., Corsinotti, A., Kapopoulou, A., Jakobsson, J., Offner, S., Baglivo, I., Pedone, P.V., Grimaldi, G., Riccio, A., *et al.* (2011). In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. *Molecular cell* 44, 361-372.
- Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* 447, 425-432.
- Reik, W., Constancia, M., Fowden, A., Anderson, N., Dean, W., Ferguson-Smith, A., Tycko, B., and Sibley, C. (2003). Regulation of supply and demand for maternal nutrients in mammals by imprinted genes. *The Journal of physiology* 547, 35-44.
- Reik, W., and Walter, J. (2001). Genomic imprinting: parental influence on the genome. *Nature reviews Genetics* 2, 21-32.
- Richards, E.J., and Elgin, S.C. (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell* 108, 489-500.
- Riggs, A.D. (1975). X inactivation, differentiation, and DNA methylation. *Cytogenetics and cell genetics* 14, 9-25.
- Ringrose, L., and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annual review of genetics* 38, 413-443.
- Rogan, P.K., Sabol, D.W., and Punnett, H.H. (1999). Maternal uniparental disomy of chromosome 21 in a normal child. *American journal of medical genetics* 83, 69-71.
- Rousseaux, S., and Khochbin, S. (2015). Histone Acylation beyond Acetylation: Terra Incognita in Chromatin Biology. *Cell journal* 17, 1-6.
- Ruf, N., Bähring, S., Galetzka, D., Pliushch, G., Luft, F.C., Nurnberg, P., Haaf, T., Kelsey, G., and Zechner, U. (2007). Sequence-based bioinformatic prediction and QUASEP identify genomic imprinting of the KCNK9 potassium channel gene in mouse and human. *Human molecular genetics* 16, 2591-2599.
- Salas, M., John, R., Saxena, A., Barton, S., Frank, D., Fitzpatrick, G., Higgins, M.J., and Tycko, B. (2004). Placental growth retardation due to loss of imprinting of Phlda2. *Mechanisms of development* 121, 1199-1210.
- Sanz, L.A., Chamberlain, S., Sabourin, J.C., Henckel, A., Magnuson, T., Hugnot, J.P., Feil, R., and Arnaud, P. (2008). A mono-allelic bivalent chromatin domain controls tissue-specific imprinting at Grb10. *The EMBO journal* 27, 2523-2532.
- Schmucker, D., and Chen, B. (2009). Dscam and DSCAM: complex genes in simple animals, complex animals yet simple genes. *Genes & development* 23, 147-156.
- Schmucker, D., Clemens, J.C., Shu, H., Worby, C.A., Xiao, J., Muda, M., Dixon, J.E., and Zipursky, S.L. (2000). *Drosophila* Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell* 101, 671-684.
- Schoenherr, C.J., Levorse, J.M., and Tilghman, S.M. (2003). CTCF maintains differential methylation at the Igf2/H19 locus. *Nature genetics* 33, 66-69.
- Schubeler, D. (2015). Function and information content of DNA methylation. *Nature* 517, 321-326.
- Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G. (2007). Genome regulation by polycomb and trithorax proteins. *Cell* 128, 735-745.

- Searle, A.G., and Beechey, C.V. (1978). Complementation studies with mouse translocations. *Cytogenetics and cell genetics* *20*, 282-303.
- Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., *et al.* (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* *450*, 908-912.
- Shogren-Knaak, M., and Peterson, C.L. (2006). Switching on chromatin: mechanistic role of histone H4-K16 acetylation. *Cell Cycle* *5*, 1361-1365.
- Sibley, C.P., Coan, P.M., Ferguson-Smith, A.C., Dean, W., Hughes, J., Smith, P., Reik, W., Burton, G.J., Fowden, A.L., and Constancia, M. (2004). Placental-specific insulin-like growth factor 2 (Igf2) regulates the diffusional exchange characteristics of the mouse placenta. *Proceedings of the National Academy of Sciences of the United States of America* *101*, 8204-8208.
- Sims, R.J., 3rd, Chen, C.F., Santos-Rosa, H., Kouzarides, T., Patel, S.S., and Reinberg, D. (2005). Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. *The Journal of biological chemistry* *280*, 41789-41792.
- Singal, R., and Ginder, G.D. (1999). DNA methylation. *Blood* *93*, 4059-4070.
- Sleutels, F., Zwart, R., and Barlow, D.P. (2002). The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* *415*, 810-813.
- Smallwood, S.A., and Kelsey, G. (2012). De novo DNA methylation: a germ cell perspective. *Trends in genetics* : *TIG* *28*, 33-42.
- Smallwood, S.A., Tomizawa, S., Krueger, F., Ruf, N., Carli, N., Segonds-Pichon, A., Sato, S., Hata, K., Andrews, S.R., and Kelsey, G. (2011). Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nature genetics* *43*, 811-814.
- Smith, F.M., Holt, L.J., Garfield, A.S., Charalambous, M., Koumanov, F., Perry, M., Bazzani, R., Sheardown, S.A., Hegarty, B.D., Lyons, R.J., *et al.* (2007). Mice with a disruption of the imprinted *Grb10* gene exhibit altered body composition, glucose homeostasis, and insulin signaling during postnatal life. *Molecular and cellular biology* *27*, 5871-5886.
- Smits, G., and Kelsey, G. (2006). Imprinting weaves its web. *Developmental cell* *11*, 598-599.
- Song, J., Rechkoblit, O., Bestor, T.H., and Patel, D.J. (2011). Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation. *Science* *331*, 1036-1040.
- Spencer, H.G., and Clark, A.G. (2014). Non-conflict theories for the evolution of genomic imprinting. *Heredity* *113*, 112-118.
- Spotswood, H.T., and Turner, B.M. (2002). An increasingly complex code. *The Journal of clinical investigation* *110*, 577-582.
- Steinhoff, C., Paulsen, M., Kielbasa, S., Walter, J., and Vingron, M. (2009). Expression profile and transcription factor binding site exploration of imprinted genes in human and mouse. *BMC genomics* *10*, 144.
- Stelzer, Y., Bar, S., Bartok, O., Afik, S., Ronen, D., Kadener, S., and Benvenisty, N. (2015). Differentiation of human parthenogenetic pluripotent stem cells reveals multiple tissue- and isoform-specific imprinted transcripts. *Cell reports* *11*, 308-320.
- Stoll, C., Alembik, Y., Dott, B., and Feingold, J. (1995). No evidence for genomic imprinting in liveborn Down syndrome patients. *Annales de genetique* *38*, 13-18.
- Stouder, C., Somm, E., and Paoloni-Giacobino, A. (2011). Prenatal exposure to ethanol: a specific effect on the H19 gene in sperm. *Reproductive toxicology* *31*, 507-512.

Strichman-Almashanu, L.Z., Lee, R.S., Onyango, P.O., Perlman, E., Flam, F., Frieman, M.B., and Feinberg, A.P. (2002). A genome-wide screen for normally methylated human CpG islands that can identify novel imprinted genes. *Genome research* 12, 543-554.

Stricker, S.H., Steenpass, L., Pauler, F.M., Santoro, F., Latos, P.A., Huang, R., Koerner, M.V., Sloane, M.A., Warczok, K.E., and Barlow, D.P. (2008). Silencing and transcriptional properties of the imprinted Airn ncRNA are independent of the endogenous promoter. *The EMBO journal* 27, 3116-3128.

Surani, M.A., and Barton, S.C. (1983). Development of gynogenetic eggs in the mouse: implications for parthenogenetic embryos. *Science* 222, 1034-1036.

Surani, M.A., Barton, S.C., and Norris, M.L. (1984). Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* 308, 548-550.

Susiarjo, M., Sasson, I., Mesaros, C., and Bartolomei, M.S. (2013). Bisphenol a exposure disrupts genomic imprinting in the mouse. *PLoS genetics* 9, e1003401.

Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. (2004). Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* 116, 51-61.

Takai, D., and Jones, P.A. (2002). Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proceedings of the National Academy of Sciences of the United States of America* 99, 3740-3745.

Tan, M., Luo, H., Lee, S., Jin, F., Yang, J.S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., *et al.* (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 146, 1016-1028.

Tan, M., Peng, C., Anderson, K.A., Chhoy, P., Xie, Z., Dai, L., Park, J., Chen, Y., Huang, H., Zhang, Y., *et al.* (2014). Lysine glutarylation is a protein posttranslational modification regulated by SIRT5. *Cell metabolism* 19, 605-617.

Tanaka, M., Gertsenstein, M., Rossant, J., and Nagy, A. (1997). Mash2 acts cell autonomously in mouse spongiotrophoblast development. *Developmental biology* 190, 55-65.

Tazi, J., and Bird, A. (1990). Alternative chromatin structure at CpG islands. *Cell* 60, 909-920.

Thorvaldsen, J.L., Duran, K.L., and Bartolomei, M.S. (1998). Deletion of the H19 differentially methylated domain results in loss of imprinted expression of H19 and Igf2. *Genes & development* 12, 3693-3702.

Tobi, E.W., Lumey, L.H., Talens, R.P., Kremer, D., Putter, H., Stein, A.D., Slagboom, P.E., and Heijmans, B.T. (2009). DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Human molecular genetics* 18, 4046-4053.

Tomizawa, S., Kobayashi, H., Watanabe, T., Andrews, S., Hata, K., Kelsey, G., and Sasaki, H. (2011). Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. *Development* 138, 811-820.

Tu, T., Zhang, C., Yan, H., Luo, Y., Kong, R., Wen, P., Ye, Z., Chen, J., Feng, J., Liu, F., *et al.* (2015). CD146 acts as a novel receptor for netrin-1 in promoting angiogenesis and vascular development. *Cell research* 25, 275-287.

Turner, B.M. (2002). Cellular memory and the histone code. *Cell* 111, 285-291.

Tycko, B. (2006). Imprinted genes in placental growth and obstetric disorders. *Cytogenetic and genome research* 113, 271-278.

Tyler, C.R., and Allan, A.M. (2014). Prenatal alcohol exposure alters expression of neurogenesis-related genes in an ex vivo cell culture model. *Alcohol* 48, 483-492.

Umlauf, D., Goto, Y., and Feil, R. (2004). Site-specific analysis of histone methylation and acetylation. *Methods Mol Biol* 287, 99-120.

- van Dijk, M., Mulders, J., Konst, A., Janssens, B., van Roy, F., Blankenstein, M., and Oudejans, C. (2004). Differential downregulation of alphaT-catenin expression in placenta: trophoblast cell type-dependent imprinting of the CTNNA3 gene. *Gene expression patterns : GEP* 5, 61-65.
- Varmuza, S., and Mann, M. (1994). Genomic imprinting--defusing the ovarian time bomb. *Trends in genetics : TIG* 10, 118-123.
- Varmuza, S., and Miri, K. (2015). What does genetics tell us about imprinting and the placenta connection? *Cellular and molecular life sciences : CMLS* 72, 51-72.
- Varrault, A., Gueydan, C., Delalbre, A., Bellmann, A., Houssami, S., Aknin, C., Severac, D., Chotard, L., Kahli, M., Le Digarcher, A., *et al.* (2006). Zac1 regulates an imprinted gene network critically involved in the control of embryonic growth. *Developmental cell* 11, 711-722.
- Verreault, A., Kaufman, P.D., Kobayashi, R., and Stillman, B. (1996). Nucleosome assembly by a complex of CAF-1 and acetylated histones H3/H4. *Cell* 87, 95-104.
- Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J.M., *et al.* (2006). The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* 439, 871-874.
- Wagschal, A., and Feil, R. (2006). Genomic imprinting in the placenta. *Cytogenetic and genome research* 113, 90-98.
- Wagschal, A., Sutherland, H.G., Woodfine, K., Henckel, A., Chebli, K., Schulz, R., Oakey, R.J., Bickmore, W.A., and Feil, R. (2008). G9a histone methyltransferase contributes to imprinting in the mouse placenta. *Molecular and cellular biology* 28, 1104-1113.
- Wake, N., Takagi, N., and Sasaki, M. (1978). Androgenesis as a cause of hydatidiform mole. *Journal of the National Cancer Institute* 60, 51-57.
- Walter, J., Hutter, B., Khare, T., and Paulsen, M. (2006). Repetitive elements in imprinted genes. *Cytogenetic and genome research* 113, 109-115.
- Wang, X., and Clark, A.G. (2014). Using next-generation RNA sequencing to identify imprinted genes. *Heredity* 113, 156-166.
- Wang, X., Soloway, P.D., and Clark, A.G. (2011). A survey for novel imprinted genes in the mouse placenta by mRNA-seq. *Genetics* 189, 109-122.
- Webster, K.E., O'Bryan, M.K., Fletcher, S., Crewther, P.E., Aapola, U., Craig, J., Harrison, D.K., Aung, H., Phutikanit, N., Lyle, R., *et al.* (2005). Meiotic and epigenetic defects in Dnmt3L-knockout mouse spermatogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 102, 4068-4073.
- Weissbach, A. (1993). A chronicle of DNA methylation (1948-1975). *Exs* 64, 1-10.
- Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T., *et al.* (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* 451, 796-801.
- Williamson, C.M., Turner, M.D., Ball, S.T., Nottingham, W.T., Glenister, P., Fray, M., Tymowska-Lalanne, Z., Plagge, A., Powles-Glover, N., Kelsey, G., *et al.* (2006). Identification of an imprinting control region affecting the expression of all transcripts in the Gnas cluster. *Nature genetics* 38, 350-355.
- Wilson, B.D., Li, M., Park, K.W., Suli, A., Sorensen, L.K., Larrieu-Lahargue, F., Urness, L.D., Suh, W., Asai, J., Kock, G.A., *et al.* (2006). Netrins promote developmental and therapeutic angiogenesis. *Science* 313, 640-644.
- Wolf, J.B., and Hager, R. (2006). A maternal-offspring coadaptation theory for the evolution of genomic imprinting. *PLoS biology* 4, e380.

- Wu, H., Coskun, V., Tao, J., Xie, W., Ge, W., Yoshikawa, K., Li, E., Zhang, Y., and Sun, Y.E. (2010). Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. *Science* 329, 444-448.
- Wutz, A., Smrzka, O.W., Schweifer, N., Schellander, K., Wagner, E.F., and Barlow, D.P. (1997). Imprinted expression of the Igf2r gene depends on an intronic CpG island. *Nature* 389, 745-749.
- Xie, H., Zou, L., Zhu, J., and Yang, Y. (2011). Effects of netrin-1 and netrin-1 knockdown on human umbilical vein endothelial cells and angiogenesis of rat placenta. *Placenta* 32, 546-553.
- Xie, Z., Dai, J., Dai, L., Tan, M., Cheng, Z., Wu, Y., Boeke, J.D., and Zhao, Y. (2012). Lysine succinylation and lysine malonylation in histones. *Molecular & cellular proteomics : MCP* 11, 100-107.
- Yamakawa, K., Huot, Y.K., Haendelt, M.A., Hubert, R., Chen, X.N., Lyons, G.E., and Korenberg, J.R. (1998). DSCAM: a novel member of the immunoglobulin superfamily maps in a Down syndrome region and is involved in the development of the nervous system. *Human molecular genetics* 7, 227-237.
- Yang, T., Adamson, T.E., Resnick, J.L., Leff, S., Wevrick, R., Francke, U., Jenkins, N.A., Copeland, N.G., and Brannan, C.I. (1998). A mouse model for Prader-Willi syndrome imprinting-centre mutations. *Nature genetics* 19, 25-31.
- Yokomine, T., Kuroiwa, A., Tanaka, K., Tsudzuki, M., Matsuda, Y., and Sasaki, H. (2001). Sequence polymorphisms, allelic expression status and chromosome locations of the chicken IGF2 and MPR1 genes. *Cytogenetics and cell genetics* 93, 109-113.
- Yokomine, T., Shirohzu, H., Purbowasito, W., Toyoda, A., Iwama, H., Ikeo, K., Hori, T., Mizuno, S., Tsudzuki, M., Matsuda, Y., *et al.* (2005). Structural and functional analysis of a 0.5-Mb chicken region orthologous to the imprinted mammalian *Ascl2/Mash2-Igf2-H19* region. *Genome research* 15, 154-165.
- Yu, L., Chen, M., Zhao, D., Yi, P., Lu, L., Han, J., Zheng, X., Zhou, Y., and Li, L. (2009). The H19 gene imprinting in normal pregnancy and pre-eclampsia. *Placenta* 30, 443-447.
- Zhang, L., Huang, Y., Chen, J.Y., Ding, Y.Q., and Song, N.N. (2015). DSCAM and DSCAML1 regulate the radial migration and callosal projection in developing cerebral cortex. *Brain research* 1594, 61-70.
- Zhang, Y., Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., Rathert, P., Brandt, O., Reinhardt, R., Fischle, W., *et al.* (2010). Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of the ADD domain with the histone H3 tail. *Nucleic acids research* 38, 4246-4253.
- Zhao, Z., Tavoosidana, G., Sjolinder, M., Gondor, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K.S., Singh, U., *et al.* (2006). Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nature genetics* 38, 1341-1347.

# ANNEXES

## **ANNEXES**

### **Communication 1**

Communication écrite à la Journée « Jeunes Chercheurs » du DHU « Risques et Grossesse », Paris –  
Juin 2014

Prix du meilleur Poster

## INTRODUCTION

Les gènes soumis à l'empreinte parentale (GSE) ont une expression monoallélique parent-dépendante suite à une régulation épigénétique complexe. Ces gènes appartiennent à des familles diverses mais ont en commun une expression préférentielle dans le placenta et le cerveau.

Du point de vue clinique, ces gènes sont impliqués dans les pathologies de la grossesse comme la prééclampsie (PE) et le retard de croissance intra-utérin (RCIU) ainsi que des pathologies neuro-développementales. De nombreux syndromes cliniques associent RCIU et déficience intellectuelle suggérant des voies de signalisation communes dans le développement des tissus placentaire et nerveux. La caractérisation de ces GSE pourrait donc permettre la mise en évidence de ces voies.

Environ 150 GSE ont été identifiés chez la souris et moins de 90 chez l'homme. Ils ont majoritairement été mis en évidence grâce à des réarrangements chromosomiques ou des disomies uniparentales à l'origine de syndromes cliniques (Prader-Willi, Beckwith-Wiedeman, etc.). D'autres approches d'identification sont basées sur des prédictions bio-informatiques mais peu des gènes prédits ont été validés. Récemment, notre équipe a mené un crible systématique, par une approche haut débit, pour identifier sans *a priori* de nouveaux GSE dans le placenta humain. Cette étude pilote a permis dans un premier temps d'identifier 7 nouveaux GSE (Barbaux et al, Epigenetics 2012). Depuis, nous avons poursuivi la validation d'autres gènes candidats aboutissant à la confirmation du statut de GSE pour l'un d'entre eux (*DSCAM*) que nous sommes en train de caractériser.

## METHODES ET RESULTATS

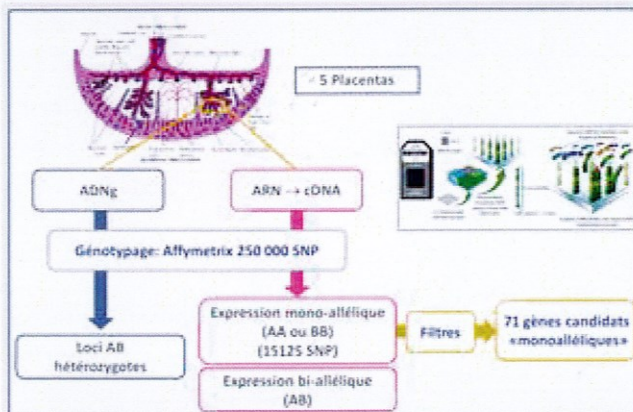


Figure 1 : Approche de crible systématique par génotypage pangénomique, pour l'identification de nouveaux GSE.

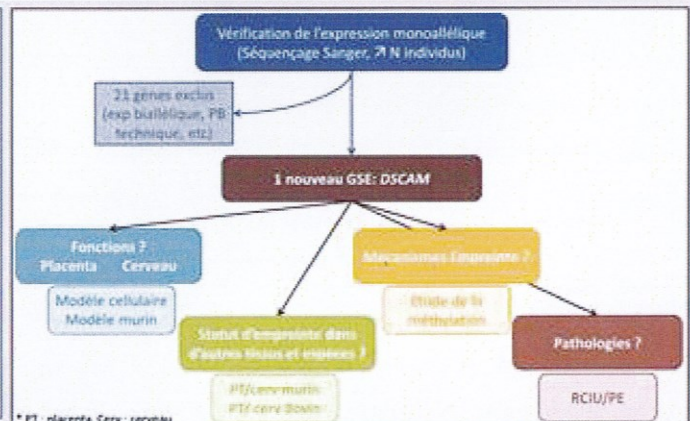


Figure 2 : Processus de validation et de caractérisation d'un nouveau GSE

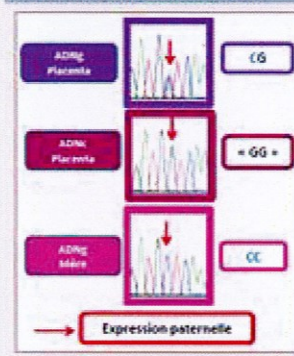


Figure 3 : Génotypage par séquençage d'un SNP du gène *DSCAM*

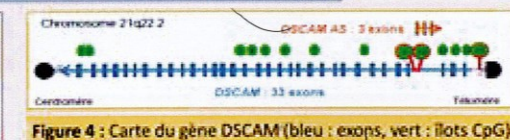


Figure 4 : Carte du gène *DSCAM* (bleu : exons, vert : îlots CpG)

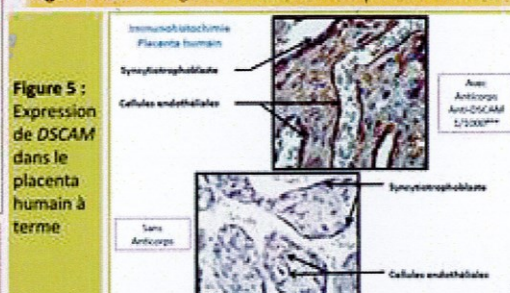


Figure 5 : Expression de *DSCAM* dans le placenta humain à terme

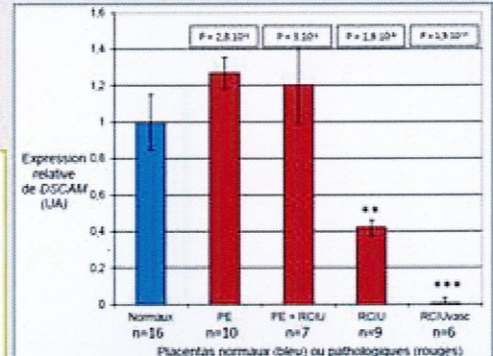


Figure 6 : Expression de *DSCAM* dans des placentas issus de grossesses normales et pathologiques

## DISCUSSION

- Par ce travail, nous avons identifié un nouveau GSE localisé sur le chromosome 21, *DSCAM*. Ce résultat est inédit d'autant plus qu'à ce jour il n'y avait pas de locus connu comme étant soumis à l'empreinte sur ce chromosome.
- Ce gène est connu pour son rôle dans le guidage des neurones lors du développement du système nerveux central. D'autres GSE déjà décrits ont également une fonction neuronale (ex: *UBE3A*, *NNTA*, *NDN*). Le placenta étant dépourvu de cellules nerveuses, cette observation pourrait s'expliquer par des voies de signalisation communes via des processus d'adhésion, prolifération et mobilité cellulaire.
- L'étude par immunohistochimie dans le placenta a montré une expression dans le syncytiotrophoblaste et l'endothélium. Par ailleurs, nous avons observé une baisse du niveau d'expression de *DSCAM* dans les RCIU notamment d'origine vasculaire. *DSCAM* pourrait donc jouer un rôle soit dans l'angiogénèse placentaire, soit dans le développement du placenta même.
- Concernant les mécanismes épigénétiques à l'origine de l'établissement de l'empreinte, nous n'avons pas pu identifier de régions différenciellement méthylées (DMR) caractéristiques des GSE, malgré l'analyse de 3 îlots CpG candidats parmi les 23 présents dans le gène.
- Des études fonctionnelles sont en cours sur des modèles cellulaires et murins pour décrypter la fonction placentaire de *DSCAM*.



## **Communication 2**

Communication type Poster-flash aux 20<sup>èmes</sup> Journées du Collège des Histologistes, Embryologistes et Cytogénéticiens, Paris – Mars 2015

## INTRODUCTION

Les gènes soumis à l'empreinte parentale (GSE) ont une expression monoallélique parent-dépendante suite à une régulation épigénétique complexe. Ces gènes appartiennent à des familles diverses mais ont en commun une expression préférentielle dans le placenta et le cerveau.

Du point de vue clinique, ces gènes sont impliqués dans les pathologies de la grossesse comme la prééclampsie et le retard de croissance intra-utérin (RCIU) ainsi que des pathologies neuro-développementales. De nombreux syndromes cliniques associent RCIU et déficience intellectuelle suggérant des voies de signalisation communes dans le développement des tissus placentaire et nerveux. La caractérisation de ces GSE pourrait donc permettre la mise en évidence de ces voies.

Environ 80 GSE ont été identifiés chez l'homme. Certains ont été mis en évidence grâce à des réarrangements chromosomiques ou des disomies uniparentales à l'origine de syndromes cliniques (Prader-Willi, Beckwith-Wiedeman, etc). D'autres approches d'identification sont basées sur des prédictions bio-informatiques mais peu des gènes prédits ont été validés. Récemment, notre équipe a mené un crible systématique, par une approche haut débit, pour identifier sans *a priori* de nouveaux GSE dans le placenta humain. Cette étude pilote a permis dans un premier temps d'identifier 7 nouveaux GSE (Barbaux et al, Epigenetics 2012). Depuis, nous avons confirmé le statut de GSE pour deux nouveaux gènes. Ce travail décrit la caractérisation de l'un d'entre eux, *DSCAM* (Down Syndrome Cell Adhesion Molecule).

## METHODES ET RESULTATS

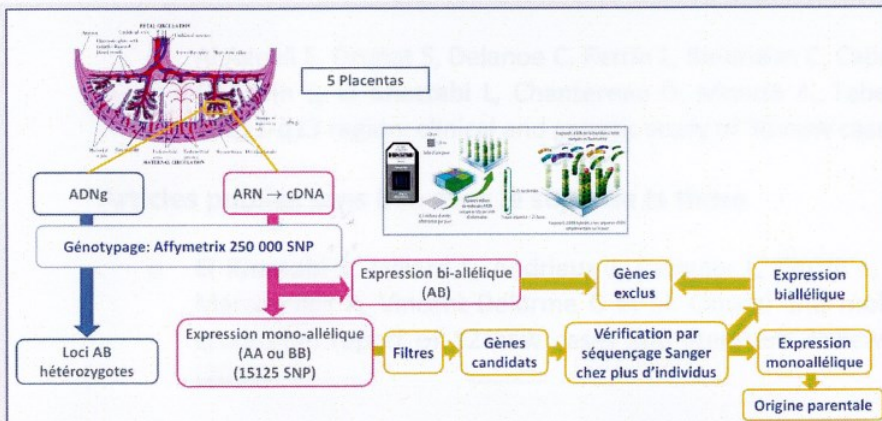


Figure 1 : Approche systématique d'identification de nouveaux GSE par génotypage pangénomique.

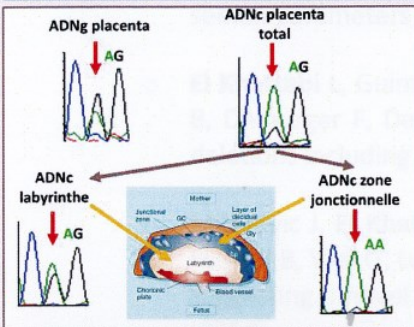


Figure 3 : Génotypage par séquençage Sanger du SNP rs46432878 du gène *DSCAM* dans le placenta murin. L'étude du SNP montre une expression monoallélique limitée à la zone jonctionnelle. L'étude de l'ADN maternel montre que l'allèle exprimé est d'origine maternelle.

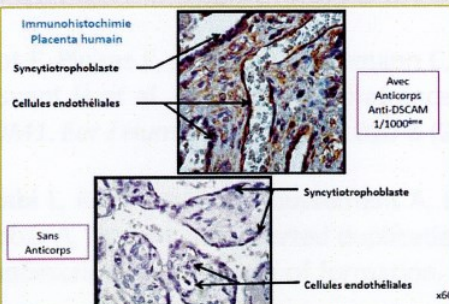


Figure 4 : Expression de *DSCAM* dans le placenta humain à terme. Le marquage montre un marquage au niveau du syncytiotrophoblaste et des cellules endothéliales.

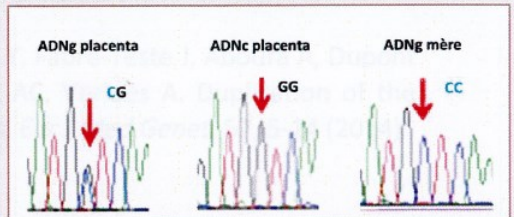
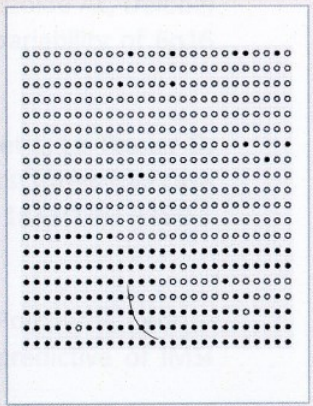


Figure 2 : Génotypage par séquençage Sanger du SNP rs2297270 du gène *DSCAM* dans le placenta humain. L'étude du SNP montre une hétérozygotie dans l'ADNg pour cet échantillon, une « homozygotie » de l'ADNc qui reflète une expression monoallélique. L'étude de l'ADN maternel montre que l'allèle exprimé est d'origine paternelle.



Figure 5 : Carte du gène *DSCAM* (rectangles bleus : exons, cercles verts : îlots CpG)

Figure 6 : Analyse de l'état de méthylation de l'îlot CpG23 après clonage. Chaque ligne correspond à un clone. Chaque cercle correspond à une cytosine, en blanc les C non méthylés et en noir les C méthylés. Ces résultats préliminaires sont en faveur d'une méthylation différentielle des allèles parentaux dans cette région.



## DISCUSSION

- Par ce travail, nous avons identifié deux nouveaux GSE localisés sur le chromosome 21, *DSCAM* et *NCAM2*. Ce résultat est inédit puisqu'à ce jour il n'y avait pas de locus connu comme étant soumis à l'empreinte sur ce chromosome.
- *DSCAM* est connu pour son rôle dans le guidage des neurones lors du développement du système nerveux central. D'autres GSE déjà décrits ont également une fonction neuronale (ex: *UBE3A*, *NNTA*, *NDN*). L'étude par immunohistochimie dans le placenta a montré une expression dans le syncytiotrophoblaste et l'endothélium. Le placenta étant un organe transitoire et dépourvu de cellules nerveuses, la présence de *DSCAM* pourrait s'expliquer par des voies de signalisation communes via des processus d'adhésion, de prolifération et de mobilité cellulaire. Ces voies de signalisation sont également de plus en plus évoquées pour les phénomènes d'angiogenèse. *DSCAM* pourrait donc jouer un rôle soit dans l'angiogenèse placentaire, soit dans le développement du trophoblaste. Nous réalisons actuellement des études fonctionnelles sur des modèles cellulaires de trophoblastes et de cellules endothéliales.
- Nous avons identifié une région différentiellement méthylée (DMR) caractéristique des GSE. Il a été montré que la mise en place de l'empreinte passe par des DMR dans des GSE largement étudiés comme le locus *H19/IGF2* où la présence ou non des marques de méthylation sur les cytosines régule la fixation de l'insulateur CTCF ce qui induit des modifications épigénétiques (modification des histones et remodelage de la chromatine) qui gouvernent la transcription.

## Bibliométrie

### Article sur le travail de la thèse (prêt à soumettre)

- **El Khattabi L**, Backer S, Pinard A, Vaiman D, Dandolo L, Bloch-Gallego E, Jammes H, Barbaux S. A genome-wide search for new imprinted genes in the human placenta identifies the first imprinted gene on chromosome 21.

### Articles publiés en lien avec le sujet de la thèse

- Abdelhedi F, **El Khattabi L**, Cuisset L, Tsatsaris V, Viot G, Druart L, Lebbar A, Dupont JM. Neonatal Silver Russel syndrome with maternal uniparental heterodisomy, trisomy 7 mosaicism and dysplasia of the cerebellum. *Am J Clin Pathol.* 142(2) : 248-53 (2014)
- Al Ageeli E, Drunat S, Delanoe C, Perrin L, Baumann C, Capri Y, Fabre-Teste J, Aboura A, Dupont C, Auvin S, **El Khattabi L**, Chantereau D, Moncla A, Tabet AC, Verloes A. Duplication of the 15q11-q13 region: clinical and genetic study of 30 new cases. *Eur J Med Genet.* 57, 5-14 (2014)

### Articles publiés sans lien avec le sujet de la thèse

- **El Khattabi L**, Jaillard S, Andrieux J, Pasquier L, Perrin L, Capri Y, Benmansour A, Toutain A, Marcorelles P, Vincent-Delorme C et al. Clinical and molecular delineation of Tetrasomy 9p syndrome: report of 12 new cases and literature review. *Am J Med Genet.* 167(6):1252-61 (2015)
- Faure C, Dupont C, Chavatte-Palmer P, Gautier B, Levy R; **ALIFERT Collaborative Group**. Are semen parameters related to birth weight? *Fertil Steril.* 103(1):6-10 (2015)
- **El Khattabi L**, Guimiot F, Pipiras E, Andrieux J, Baumann C, Bouquillon S, Delezoide AL, Delobel B, Demurger F, Dessuant H et al. Incomplete penetrance and phenotypic variability of 6q16 deletions including *SIM1*. *Eur J Hum Genet.* 23(8):1010-8 (2015)
- Milosevic J, **El Khattabi L**, Roubergue A, Coussement A, Doummar D, Cuisset L, Le Tessier D, Flageul B, Viot G, Lebbar A, Dupont JM. Inverted duplication with deletion: first interstitial case suggesting a novel undescribed mechanism of formation. *Am J Med Genet.* 1641(12) : 3180-6 (2014)
- Sifer C, **El Khattabi L**, Hugues JN, Dupont C, Sermondade N, Herbemont C, Porcher R, Cedrin-Durnerin I, Poncelet C, Lévy R. Is sperm grade under high magnification predictive of IMSI clinical outcome? *Eur J Obstet Gynecol Reprod Biol.* 181 : 189-94 (2014)
- **El Khattabi L**, Dupont C, Sermondade N, Hugues JN, Poncelet C, Porcher R, Cedrin-Durnerin I, Lévy R, Sifer C. Is morphologically selected sperm injection useful for patients selected for teratozoospermia or for previous implantation failures? *Fertil Steril.* 100, 62-8 (2013)
- Pallier K, Cazes A, **El Khattabi L**, Lecchi C, Desroches M, Danel C, Riquet M, Fabre-Guillevin E, Laurent-Puig P, Blons H. DeltaN TP63 activation, epithelial phenotype maintenance and survival in Lung Squamous Cell Carcinoma. *Tumor Biol.* 33, 41-51 (2012)