



**HAL**  
open science

# Etude de la fonction de l'ARN non codant H19 et de la régulation de l'Empreinte Parentale au locus H19/lgf2 dans des modèles murins transgéniques

Anne Gabory

► **To cite this version:**

Anne Gabory. Etude de la fonction de l'ARN non codant H19 et de la régulation de l'Empreinte Parentale au locus H19/lgf2 dans des modèles murins transgéniques. Biologie du développement. Université Paris VII - Denis Diderot, 2008. Français. NNT : 2008PA077106 . tel-03271000

**HAL Id: tel-03271000**

**<https://theses.hal.science/tel-03271000v1>**

Submitted on 25 Jun 2021

**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 PARIS VII - DENIS DIDEROT**

**THESE**

En vue de l'obtention du grade de

**DOCTEUR DE L'UNIVERSITE PARIS VII**

Discipline : Epigénétique

Formation Doctorale : Biologie Moléculaire et Cellulaire du Développement

Ecole Doctorale : La Logique du Vivant

Présentée et soutenue publiquement  
par

**Anne GABORY**

le vendredi 26 septembre 2008

**Etude de la fonction de l'ARN non codant *H19* et de la  
régulation de l'Empreinte Parentale au locus *H19/Igf2*  
dans des modèles murins transgéniques.**

Directrice de thèse : Luisa DANDOLO

**Jury**

Pr. Jonathan WEITZMAN  
Dr. Carole CHARLIER  
Dr. Matthieu GÉRARD  
Dr. Benoît ARCANGIOLI  
Dr. Laurent JOURNOT  
Dr. Luisa DANDOLO

Président  
Rapporteur  
Rapporteur  
Examineur  
Examineur  
Directrice de thèse

UNIVERSITE PARIS VII - DENIS DIDEROT

## THESE

En vue de l'obtention du grade de

**DOCTEUR DE L'UNIVERSITE PARIS VII**

Discipline : Epigénétique

Formation Doctorale : Biologie Moléculaire et Cellulaire du Développement

Ecole Doctorale : La Logique du Vivant

Présentée et soutenue publiquement  
par

**Anne GABORY**

le vendredi 26 septembre 2008

**Etude de la fonction de l'ARN non codant *H19* et de la  
régulation de l'Empreinte Parentale au locus *H19/Igf2*  
dans des modèles murins transgéniques.**

Directrice de thèse : Luisa DANDOLO

### Jury

Pr. Jonathan WEITZMAN  
Dr. Carole CHARLIER  
Dr. Matthieu GÉRARD  
Dr. Benoît ARCANGIOLI  
Dr. Laurent JOURNOT  
Dr. Luisa DANDOLO

Président  
Rapporteur  
Rapporteur  
Examineur  
Examineur  
Directrice de thèse

**Table des matières**

Liste des abréviations .....	III
Index des figures .....	IV
Index des tableaux .....	V
<b>Introduction</b> .....	<b>1</b>
<b>Chapitre I - Epigénétique : Empreinte Parentale et ARN non codants</b> .....	<b>5</b>
1) Epigénétique : définition et mécanismes .....	6
1-1) Historique et définition actuelle .....	6
1-2) La méthylation de l'ADN .....	8
1-2-1) Le "méthylome" .....	8
1-2-2) Les ADN méthyltransférases (Dnmt) .....	10
a) La méthylation de maintenance : Dnmt1 .....	11
b) La méthylation de novo : Dnmt3a et Dnmt3b .....	11
1-2-3) Les protéines de liaison à l'ADN méthylé (MBP) .....	13
1-3) Les modifications des histones .....	15
1-3-1) Les modifications post-traductionnelles des histones .....	15
1-3-2) L'écriture et l'effacement des marques des histones .....	18
1-3-3) La lecture des marques des histones .....	18
1-3-4) Polycomb et Trithorax .....	19
1-3-5) Les autres modifications de la chromatine .....	28
1-4) Marques épigénétiques et hétérochromatine .....	21
1-4-1) Dialogue entre les marques épigénétiques .....	21
1-4-2) L'inactivation du chromosome X, un exemple de chronologie aboutissant à la répression transcriptionnelle .....	22
2) Les ARN non codants .....	24
2-1) Introduction .....	24
2-2) Empreinte Parentale et ARN non codants .....	25
2-3) siRNA et interférence ARN .....	25
2-4) miRNA et répression traductionnelle .....	26
2-5) piRNAs et protéines Piwi .....	27
3) L'Empreinte Génomique Parentale .....	28
3-1) Historique et définition .....	28
3-2) Caractéristiques .....	30
3-3) Le cycle de l'Empreinte parentale .....	31
3-4) Régulation de l'expression des gènes .....	33
3-4-1) Régulation par les insulateurs .....	33
a) La protéine CTCF .....	33
b) Régulation de l'expression des gènes au locus <i>H19/Igf2</i> .....	34
3-4-2) Régulation par les ARN non codants .....	37
3-5) Empreinte Parentale et Pathologie humaine .....	38
<b>Chapitre II - Croissance et Myogenèse</b> .....	<b>39</b>
1) Croissance embryonnaire et post-natale .....	40
1-1) Introduction .....	40
1-2) L'axe Insulin-like Growth-Factor : une analyse génétique poussée .....	40
1-3) Croissance embryonnaire et croissance post-natale .....	42
1-4) Croissance embryonnaire et Empreinte Parentale .....	43
2) Myogenèse .....	45
2-1) Myogenèse embryonnaire .....	45
2-2) Régénération musculaire .....	47
2-3) Myogenèse et Epigénétique .....	49

<b>Chapitre III - <i>H19</i> et <i>Igf2</i> sont deux gènes étroitement liés</b>	<b>51</b>
1) Le produit du gène <i>H19</i> est un ARN non codant	52
1-1) La découverte du gène <i>H19</i>	52
1-2) Le gène <i>H19</i> est lié au gène <i>Igf2</i>	52
1-2-1) Structure des gènes <i>H19</i> et <i>Igf2</i>	52
1-2-2) L'expression des gènes <i>H19</i> et <i>Igf2</i> est étroitement coordonnée	53
1-3) Fonctions proposées pour le gène et l'ARN <i>H19</i>	53
1-3-1) Le produit du gène <i>H19</i> est un ARN non traduit	53
1-3-2) Modèles murins d'inactivation du gène <i>H19</i>	54
1-3-3) Souris parthénogénétiques et locus <i>H19/Igf2</i>	55
1-3-4) Le gène <i>H19</i> : un gène suppresseur de tumeur ?	57
2) Pathologies humaines associées aux gènes <i>H19/Igf2</i>	58
2-1) Le syndrome de Beckwith-Wiedemann	58
2-2) Le syndrome de Silver-Russell	60
<b>Article 1 - The <i>H19</i> gene: regulation and function of a non-coding RNA</b>	<b>61</b>
Introduction	62
Article	63
<b>Résultats</b>	<b>69</b>
<b>Article 2 - The <i>H19</i> locus acts <i>in vivo</i> as a tumour suppressor</b>	<b>70</b>
Introduction	71
Article	73
Conclusions	79
<b>Article 3 - The <i>H19</i> gene acts as a <i>trans</i> regulator of the imprinted gene network controlling growth in mice</b>	<b>81</b>
Introduction	82
Article	84
Conclusions	104
<b>Article 4 - <i>H19</i> and <i>Igf2</i> imprinting relaxation as a marker for adult stem cells</b>	<b>105</b>
Introduction	106
Article	108
Conclusions	123
<b>Discussion</b>	<b>125</b>
1) Un rôle de régulateur d'un réseau pour le gène <i>H19</i>	126
1-1) Implications de ces résultats	129
1-1-1) Confirmation de l'existence d'un réseau fonctionnel	129
1-1-2) <i>H19</i> , réseau et parthénogenèse	130
1-1-3) Implications pour la pathologie humaine	130
1-2) Mécanismes d'action de l'ARN <i>H19</i>	132
1-2-1) Mécanisme associé à miR-675	132
1-2-2) Mécanisme pour l'ARN pleine longueur	133
1-2-3) ARN <i>H19</i> et méthylation de l'ADN ?	135
1-2-4) Mécanisme d'action indirecte	135
2) Empreinte et cellules souches adultes	137
2-1) Implication de ces résultats pour l'étude des cellules satellites	139
2-2) Cellules souches et thérapie cellulaire	140
3) Conclusions générales et perspectives	141
<b>Bibliographie</b>	<b>142</b>

**Liste des abréviations utilisées**

Air	Antisense of Igf2R
ARNnc	ARN non codant
BWS	Syndrome de Beckwith-Wiedemann
Cdkn1c / p57KIP2	Kinase cycline dépendante inhibitrice du cycle cellulaire
CpG ou meCpG	suite de nucléotides cytosine et guanine sur la séquence de l'ADN, me indique la méthylation
CTCF	CCCTC-binding Factor
Dlk1	Delta-like 1
DMR	Région de Méthylation Différentielle
Dnmt	ADN méthyltransférase
Fg	fully grown oocyte, ovocyte mature
GH	Growth Hormone, hormone de croissance
Gnas	G $\alpha$ s subunit, sous unité $\alpha$ s des protéines G
Grb10	Growth factor receptor-bound protein 10
Gtl2	Genet rap locus 2 / Meg 3 (maternally expressed gene 3)
HDAC	Histone DésACétylase
HSC	Cellules Souches Hématopoïétiques
ICR, IC ou ICE	Imprinting Control Region, centre d'Empreinte
Igf, IgfR	insulin-like growth-factor, Igf Receptor
IGN	Imprinted Gene Network, réseau de gènes soumis à l'empreinte
IMP	Igf2 mRNA-binding Protein
KMT	Lysine MéthylTransférase
Kvlqt1 ou Kcnql	Canal potassium impliqué dans le QT long cardiaque
Lit1	Long intronic transcript
LOI	Loss of Imprinting, Perte d'Empreinte
MAR	Région d'Attachement à la Matrice nucléaire
MBD	MeCpG-Binding Domain, famille de protéine de liaison à l'ADN méthylé
MBP	Protéine de liaison à l'ADN méthylé
MeCP2	MeCpG-binding protein
Mest	Mesoderm specific transcript
miRNA	micro RNA
MRF	Myogenic Regulatory Factor
ng	non growing oocyte, ovocyte immature
Peg	Paternally expressed gene
PGC	Cellules Germinales Primordiales
piRNA	Piwi-interactin RNA
PRC	Polycomb Group Complex
RBP	RNA-Binding Protein
RISC	RNA-induced silencing complex
siRNA	small interfering RNA
SRS	Syndrome de Silver-Russell
UPD	UniParental Disomy
UTR	UnTranslated Region, région non traduite de l'ARNm
Xist	X-inactive specific transcript
YAC	Yeast Artificial Chromosome, chromosome artificiel de levure
Zac1	Zinc finger transcription factor inducing Apoptosis and cell Cycle arrest

**Index des figures****Introduction**

1. Epigénétique et destin cellulaire. ....	6
2. Compaction de la chromatine au sein du noyau. ....	7
3. La méthylation de l'ADN.....	8
4. Evolution du profil de méthylation au cours du Développement.....	9
5. Le méthylome. ....	10
6. La famille des DNA méthyltransférases (Dnmt). ....	10
7. Les principales protéines de liaison à l'ADN méthylé (MBP). ....	13
8. Le nucléosome. ....	14
9. Les modifications des histones. ....	15
10. Modifications des histones et transcription.....	16
11. Dialogue entre les marques des histones ....	16
12. Technique d'Immunoprécipitation de la chromatine (ChIP). ....	17
13. Les Complexes Polycomb et Trithorax.....	19
14. Chronologie des événements au cours de l'inactivation du chromosome X. ....	22
15. Le monde des ARN. ....	24
16. Modèle de mécanisme pour les siRNA.....	25
17. Modèle de mécanisme pour les miRNA ....	26
18. Modèle de mécanisme pour les piRNA. ....	27
19. Expériences de transfert nucléaire. ....	28
20. Carte des gènes soumis à l'Empreinte Parentale chez la souris.....	29
21. Transmission de la mutation et du phénotype des souris KO <i>Igf2</i> .....	29
22. Le cycle de l'Empreinte Parentale.....	31
23. Le locus <i>H19/Igf2</i> . ....	34
24. Technique 3C (Chromosome Conformation Capture). ....	36
25. Architecture de la chromatine au locus <i>H19/Igf2</i> . ....	36
26. Régulation de l'Empreinte par les ARN non codants de grande taille. ....	37
27. Trois mécanismes associés aux pathologies d'Empreinte. ....	38
28. Les acteurs de la voie Insulin-like Growth-Factor (IGF). ....	40
29. Co-évolution de l'Empreinte et de la placentation. ....	43
30. Réseau de gènes soumis à l'Empreinte.....	44
31. La myogenèse embryonnaire. ....	45
32. Cellules satellites et régénération musculaire. ....	47
33. Structure du gène <i>H19</i> murin.....	52
34. Structure du gène <i>Igf2</i> murin. ....	52
35. Modèles murins d'inactivation du gène <i>H19</i> . ....	54
36. Production expérimentale d'embryons parthénogénétiques. ....	55
37. Résumé des travaux sur les embryons bi-maternels ....	56

**Article 1**

1. Common enhancers of the <i>H19</i> and <i>Igf2</i> genes. ....	64
---	----

**Article 2**

1. Expression of the <i>H19-Igf2</i> locus. ....	74
2. Teratocarcinoma model. ....	75
3. Colorectal cancer model ....	76
4. Experimental liver carcinogenesis model ....	76

**Article 3**

1. Expression of the <i>H19</i> gene in the transgenic lines ....	88
2. Growth curves of transgenic animals.....	89
3. RT-sqPCR analysis of the <i>Igf2</i> gene expression in transgenic Tg line ....	90
4. RT-qPCR analysis of the <i>IGN</i> gene expression.....	90

**Article 4**

1. <i>Igf2</i> expression in activated <i>Igf2</i> <sup>+/-</sup> satellite cells. ....	112
2. <i>H19</i> expression in quiescent in <i>H19</i> <sup>+/-</sup> satellite cells.....	112
3. <i>Igf2</i> expression in quiescent and cultured SD7 satellite cells .....	113
4. Methylation status of the <i>H19/Igf2</i> ICR in satellite cells .....	114

**Discussion**

37. Résumé des travaux sur les embryons bi-maternels. ....	130
38. Mécanismes potentiels d'action de l'ARN <i>H19</i> .....	133
39. Locus <i>Dlk1/Gtl2/Dio3</i> .....	134
40. Résumé des résultats et questions sur les cellules souches adultes.....	138

**Index des Tableaux****Introduction**

1. Les domaines de liaisons aux histones.....	18
2. Les loci soumis à l'Empreinte Parentale. ....	33
3. Mutations nulles de la voie Insulin-like Growth-Factor (IGF). ....	41
4. Gènes soumis à l'Empreinte et leurs effets. ....	43
5. Expression embryonnaire coordonnée des gènes <i>H19</i> et <i>Igf2</i> .....	53

**Article 3**

1. Bias in transmission of the transgenic genotype. ....	91
2. Effects of the IGN genes knock-out on growth regulation .....	93
3. Primer used for PCR analysis .....	96
S1. RT-qPCR analysis of the IGN gene expression.....	100

**Discussion**

6. Effet de l'invalidation de gènes soumis à l'Empreinte sur la croissance. ....	128
--	-----

# *Introduction*

**"H19 was the H row, 19th spot. We had made a liver cDNA library and we were screening through them. H19 was incredibly abundant and we showed that it was developmentally regulated. I kept plugging away on it, thinking there must be something here. Given its abundance and its tight regulation, this can't be a garbage RNA.**

**If I have one regret about my career, it is that I had to stop before I figured out what H19 does! Unfortunately this little filter array name has stuck on it."**

Extraits de : The making of a President:  
an interview with Shirley Tilghman.  
Interview by Jane Gitschier.  
(Tilghman, 2006)

Voici donc comment a commencé l'histoire du gène *H19*. Le premier article décrivant le clonage de ce gène a 20 ans (Pachnis *et al.*, 1988) et aujourd'hui le rôle de ce gène n'est toujours pas élucidé. Mon travail de thèse a, en partie, porté sur l'étude de la fonction de l'ARN *H19* et nous a permis de proposer qu'il participe à la régulation d'un réseau de gènes soumis à l'Empreinte impliqué dans le contrôle de la croissance embryonnaire.

*H19* a deux particularités principales : celle d'être soumis à l'Empreinte Parentale et celle de coder un ARN non traduit, deux caractéristiques épigénétiques. Le **premier chapitre** de l'**introduction bibliographique** porte donc sur l'**Epigénétique**. La première partie décrit les différents acteurs épigénétiques que sont la méthylation de l'ADN et les modifications post-traductionnelles des histones et leurs liens avec la transcription des gènes. Une seconde partie résume brièvement les données existantes sur les **ARN non codants**. La troisième partie de ce premier chapitre fait le point sur les connaissances concernant l'**Empreinte Parentale** et en particulier sur la régulation de l'expression au locus *H19/Igf2*.

Le gène *H19* est très fortement exprimé au cours de l'embryogenèse et le phénotype des animaux invalidés pour ce gène est une surcroissance somatique. Par ailleurs, il est rapidement réprimé après la naissance dans tous les tissus à l'exception du muscle squelettique. Aussi, l'ARN *H19* est pressenti pour jouer un rôle dans le contrôle de la croissance fœtale et au cours de la myogenèse. Le **deuxième chapitre** de l'**introduction** porte sur deux aspects. La **croissance embryonnaire et post natale** contrôlée particulièrement par le système IGF fait l'objet de la première partie de ce chapitre. La seconde partie est consacrée à la **myogenèse** et à la **régénération musculaire**, phénomène qui fait intervenir les cellules souches musculaires adultes appelées cellules satellites.

Le **troisième chapitre** de l'**introduction** introduit le gène *H19* et son voisin chromosomique indissociable *Igf2*. Une première partie reprend l'historique de la découverte de l'ARN *H19*, les données concernant le profil d'expression d'*H19* et *Igf2*, les pistes quant à sa fonction. La seconde partie est consacrée aux syndromes de Beckwith-Wiedemann et de Silver-Russell. Ces deux pathologies présentent des phénotypes respectivement de surcroissance et de retard de croissance. Elles sont toutes les deux liées au locus *H19/Igf2*.

La revue de synthèse portant sur la régulation et la fonction du gène *H19*, que nous avons publiée en 2006, clôture cette introduction (Gabory *et al.*, 2006).

Le but de ce travail de thèse a été d'élucider le rôle du gène *H19*. Une controverse existe quant à la **fonction du gène au cours de la tumorigenèse**. Oncogène ou suppresseur de tumeur ? J'ai tout d'abord participé à une étude où nous avons montré dans trois modèles murins que le locus *H19* joue le rôle de suppresseur de tumeur (Yoshimizu *et al.*, 2008). Le premier modèle est l'induction de tératocarcinomes par greffe d'embryon sous la capsule rénale de souris. Les tumeurs induites par les embryons *H19*<sup>-/-</sup> sont plus grosses que les tumeurs induites par des embryons sauvages. Le deuxième modèle est un modèle de tumeurs intestinales dues à une mutation du gène *Apc*. Les souris *Apc*<sup>-/-</sup>*H19*<sup>-/-</sup> développent plus de tumeurs que les souris *Apc*<sup>-/-</sup> et les tumeurs de petite taille sont plus nombreuses. Le dernier modèle est l'induction d'hépatocarcinomes. Les souris n'exprimant pas *H19* développent des tumeurs plus rapidement que les souris sauvages. Ces trois modèles impliquent un rôle de l'ARN *H19* dans l'apparition et la croissance tumorale, montrant ainsi un rôle de suppresseur de tumeur chez la souris.

Dans le but d'étudier la fonction de l'ARN *H19*, des modèles d'inactivation du gène ont été développés, en particulier dans notre laboratoire avec le modèle *H19*<sup>Δ3</sup> (Ripoche *et al.*, 1997). Ces modèles présentent un phénotype de surcroissance somatique liée à la réactivation de l'allèle maternel du gène du facteur de croissance *Igf2* par une modification de la structure de la chromatine. Suite à ces travaux d'inactivation du gène par recombinaison homologe, des modèles de souris transgéniques surexprimant le gène ont été développés au laboratoire peu avant mon arrivée. Ainsi, mon travail a été d'**analyser ces lignées transgéniques murines** dans le but de comprendre la fonction de l'ARN non codant *H19*. Les transgènes, placés sur fond *H19*<sup>-/-</sup>, sont capables de sauver le phénotype et de restaurer une taille et un poids comparables à ceux de souris sauvages. Une analyse de l'expression des gènes d'un réseau de gènes soumis à l'Empreinte récemment décrit (Varrault *et al.*, 2006), qui pourrait

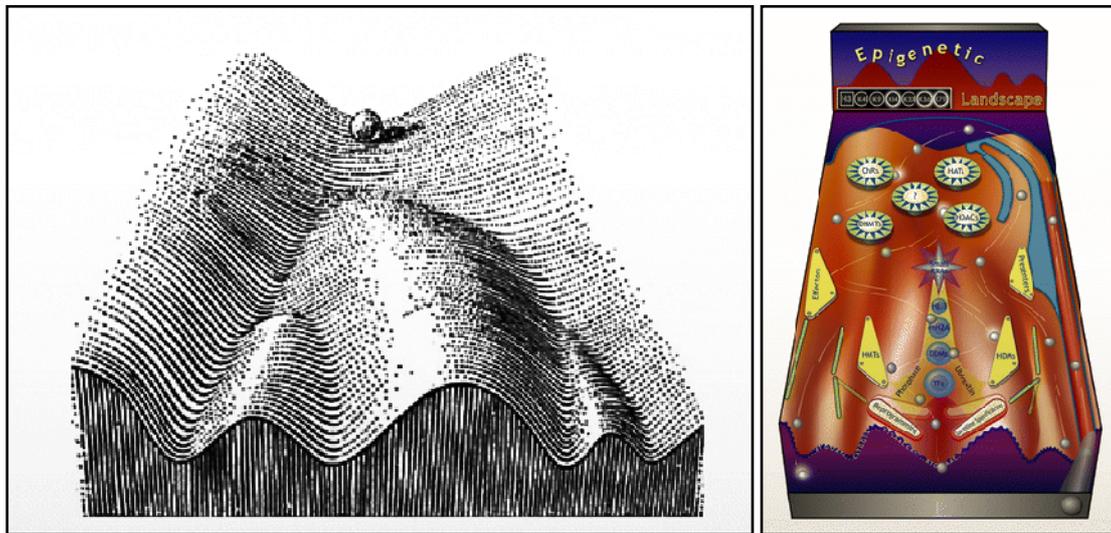
contrôler la croissance embryonnaire et post-natale, a montré que l'ARN *H19* est un *trans* régulateur de ce réseau. Le manuscrit de ce travail a été soumis pour publication.

Après la naissance le gène *H19* est réprimé dans tous les organes à l'exception du muscle squelettique. Nous nous sommes donc intéressées à cet organe particulier. Une étude a montré que plusieurs gènes soumis à l'Empreinte, dont *H19* et *Igf2*, sont induits lors du processus de régénération musculaire, qui fait intervenir les cellules satellites, considérées comme les cellules souches adulte de cet organe. Par ailleurs, il avait été décrit une perte d'Empreinte du gène *Igf2* dans les cellules souches hématopoïétiques. La dernière partie de mon travail a consisté en l'étude de la **régulation de l'Empreinte Parentale du locus *H19/Igf2*** dans les cellules souches adultes, avec comme modèle la cellule satellite. Nous avons mis en culture ou trié les cellules satellites de souris *Igf2*<sup>+/-</sup>, *H19*<sup>+/-</sup> et de souris *H19Igf2*<sup>domesticus/spretus</sup>, portant un polymorphisme permettant de distinguer les deux allèles. Nous avons montré une expression biallélique des gènes *Igf2* et *H19* dans les cellules satellites, en absence de modifications des marques de méthylation d'Empreinte. Ces résultats montrent un relâchement de l'Empreinte dans ces cellules. *Igf2* est également exprimé de manière biallélique dans les cellules souches hématopoïétiques, ce qui suggère que l'absence de lecture de l'Empreinte pourrait être une caractéristique des cellules souches adultes. Le manuscrit de ce travail est en cours de préparation.

La **discussion** abordera les résultats obtenus au cours de cette thèse qui montrent *H19* comme un ARN régulateur de la croissance embryonnaire *via* le contrôle d'un réseau de gènes soumis à l'Empreinte. Il a été montré que certains de ces gènes sont réprimés après la naissance de façon concomitante. Leur réactivation à l'âge adulte est observée au cours de la tumorigenèse. Aussi le rôle de suppresseur de tumeur d'*H19* pourrait s'expliquer par le contrôle de ce réseau pour limiter la prolifération. Les mécanismes possibles par lesquels l'ARN *H19* peut contrôler l'expression d'autres ARNm seront détaillés. Enfin, la discussion abordera le problème de l'Empreinte dans les cellules souches : l'implication de ces résultats en ce qui concerne la réflexion autour de l'origine de ces cellules, la fonction de ces gènes dans l'homéostasie musculaire et la connaissance de la biologie de ces cellules qui peuvent être utilisée en thérapie cellulaire.

# Chapitre I

## *Epigénétique : Empreinte Parentale et ARN non codants*



**Figure 1 : Epigénétique et destin cellulaire.**

Waddington représentait sa vision de l'Epigénétique comme une balle, la cellule, qui a le choix au cours de la différenciation cellulaire pendant l'Embryogenèse, entre deux parcours : les destins cellulaires (à gauche). De nos jours, plusieurs acteurs épigénétiques sont connus et il apparait que les différentes voies de méthylation de l'ADN, de modifications des histones, de remodelage de la chromatine et de mémoire épigénétique agissent de concert pour mener les cellules filles vers les différents destins cellulaires (à droite). (Adapté de Goldberg 2007).

## 1) Epigénétique : définition et mécanismes

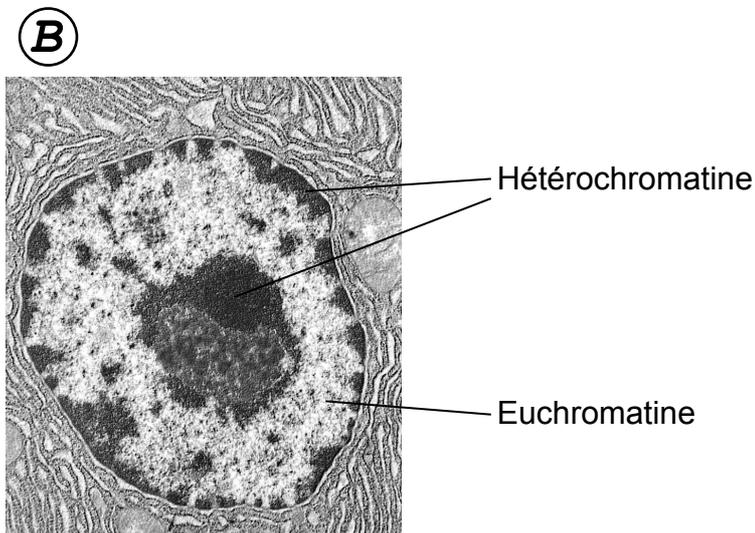
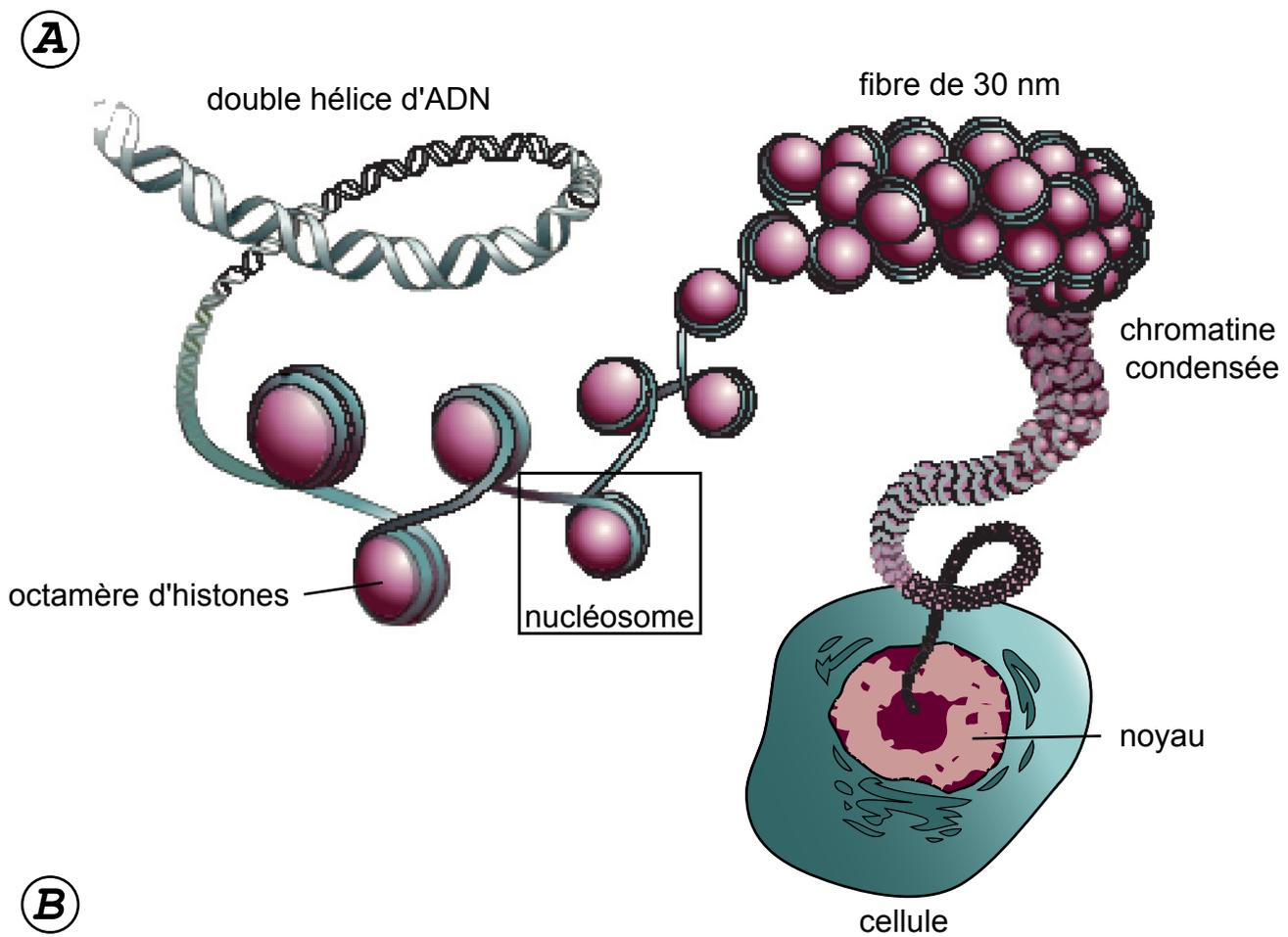
### 1-1) Historique et définition actuelle

C'est à Conrad Waddington qu'est prêtée l'invention du mot épigénétique qu'il définissait comme la discipline de la Biologie étudiant les relations de cause à effet entre les gènes et leurs produits, faisant apparaître le phénotype. En 1957, afin de réconcilier le monde de la Génétique et celui du Développement, Waddington représente l'Epigénétique comme un paysage dans lequel une bille, la cellule, peut prendre différentes trajectoires représentant les différents destins cellulaires ([Figure 1 gauche](#)). A partir d'un même génome, deux cellules filles pourront avoir deux programmes d'expression génique différents, déterminant ainsi le devenir de ces cellules. Dans cette optique, la définition de l'épigénétique serait **l'étude des changements stables et héréditaires de l'expression des gènes sans modification de la séquence primaire de l'ADN.**

La notion d'Epigénétique est aujourd'hui moins restreinte au Développement et comprend entre autres la réparation de l'ADN, les voies de régulation post-transcriptionnelle et traductionnelle des petits ARN non codants aussi bien que la transmission d'un programme d'expression génique aux cellules filles. Aussi, certains chercheurs comme Denise Barlow admettent que **"L'épigénétique a toujours été l'ensemble de ces choses bizarres et merveilleuses que la génétique ne sait pas expliquer."**

La recherche actuelle dans le domaine de l'Epigénétique s'attache particulièrement à étudier les modifications covalentes de l'ADN et post-traductionnelles des histones et les mécanismes par lesquels celles-ci influencent la structure de la chromatine et l'expression des gènes (Goldberg *et al.*, 2007). Depuis Waddington, la méthylation de l'ADN et différentes modifications post-traductionnelles des histones ont été découvertes ainsi que de nombreux "écrivains" et "lecteurs" de ces marques épigénétiques. L'étude du destin cellulaire au cours du développement peut aujourd'hui se faire sous une autre approche ; les différentes voies de méthylation de l'ADN, de modifications des histones et des activateurs et répresseurs transcriptionnels interagissent entre eux pour aboutir au patron d'expression adéquat pour la cellule ([Figure 1 droite](#)).

Dans les cellules eucaryotes, le matériel génétique est localisé dans le compartiment nucléaire et organisé en une structure complexe constituée d'ADN et de protéines : la



**Figure 2 : Compaction de la chromatine au sein du noyau.**

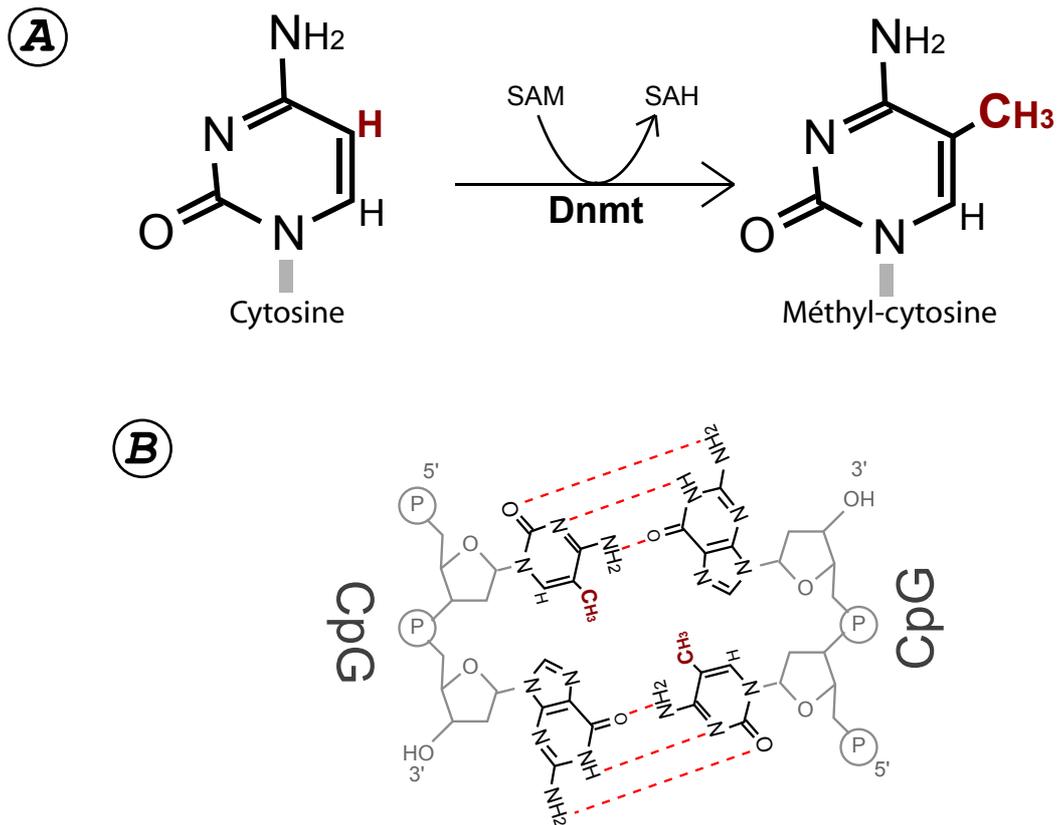
**A.** La double hélice d'ADN s'enroule autour des histones pour former les nucléosomes, les unités fondamentales de la chromatine. Ces nucléosomes peuvent interagir entre eux pour se compacter en fibre de 30nm qui peut être encore compactée en une structure plus dense au sein de l'hétérochromatine.

**B.** La structure du noyau en microscopie électronique fait apparaître deux formes pour l'ADN : l'euchromatine, claire et l'hétérochromatine, dense aux électrons.

chromatine. La longueur du brin d'ADN est d'environ 2 mètres mais le noyau est un compartiment qui fait quelques  $\mu\text{m}$  de diamètre seulement. Il est donc nécessaire de compacter le matériel génétique afin de pouvoir le stocker dans ce petit organelle (Figure 2A) tout en le gardant disponible à tout moment pour la transcription génique, la réplication ou la réparation de l'ADN en réponse aux signaux cellulaires. Ces fonctions de compaction et de mise à disposition sont réalisées par remodelage de la chromatine.

Le mot chromatine vient du grec "kroma", la couleur, du fait de son aspect granuleux et sa capacité à fixer les colorants. Les images de noyaux en microscopie électronique permettent de distinguer deux types de chromatine selon son aspect plus ou moins dense (Figure 2B). L'**hétérochromatine** forme des structures denses aux électrons, situées à la périphérie du noyau et au sein de l'euchromatine. L'hétérochromatine est considérée comme la forme condensée, transcriptionnellement inactive de la chromatine. L'**euchromatine** forme une structure plus claire et est considérée comme la forme relâchée, active transcriptionnellement de la chromatine. Ces structures sont dynamiques, ainsi le matériel génétique de l'hétérochromatine peut être "remodelé" en euchromatine en réponse aux besoins de la cellule et *vice versa*.

L'accessibilité ou non des gènes aux facteurs de transcription est régulée par deux principaux acteurs : la méthylation de l'ADN et les modifications des histones. Dans les deux prochaines parties, ces voies seront brièvement détaillées puis une dernière partie décrira les interactions entre ces acteurs qui aboutissent à la structure de la chromatine appropriée pour une région génomique donnée.



**Figure 3 : La méthylation de l'ADN.**

**A.** La substitution du groupement méthyle (CH<sub>3</sub>) sur la cytosine est catalysée par une ADN méthyltransférase (Dnmt) en utilisant la S-adenosyl-méthionine (SAM) comme substrat.

**B.** La méthylation des cytosines sur les dinucléotides CpG ne modifie pas les appariements Watson-Crick de l'ADN.

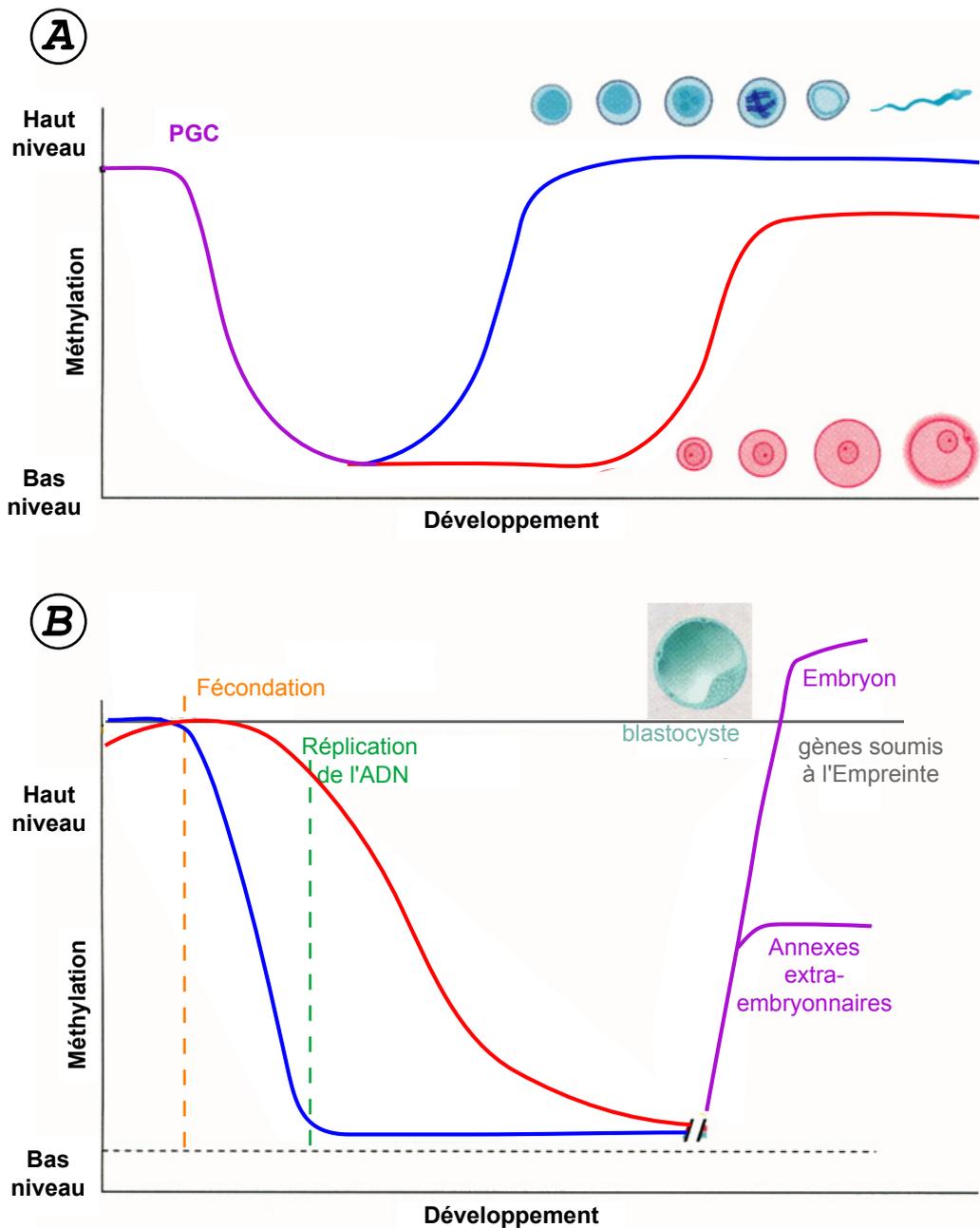
## 1-2) La méthylation de l'ADN

La méthylation de l'ADN est la substitution covalente d'un groupement méthyle (CH<sub>3</sub>) sur les acides nucléiques (Figure 3). Chez les procaryotes, la méthylation a été rapportée sur les adénines et cytosines. Chez les eucaryotes, la méthylation de l'ADN est associée uniquement aux cytosines et plus particulièrement sur les cytosines suivies de guanines (dinucléotide CpG). La méthylation est notamment présente chez les plantes, les mammifères, certaines levures et certains insectes. Toutefois, la *Drosophila* présente seulement un faible niveau de méthylation de l'ADN et la littérature rapporte que la levure *S. cerevisiae* ou le nématode *C. elegans* par exemple en sont dépourvus (Bird, 2002).

Cette réaction est catalysée par une famille d'enzymes appelées Dnmt (DNA méthyltransférase) à partir du substrat S-adenosyl-méthionine (SAM) comme donneur de groupement méthyle. Cette modification de l'ADN n'altère pas les appariements Watson Crick de l'ADN. La méthylation de l'ADN est généralement associée à la répression transcriptionnelle (Klose and Bird, 2006). Le génome des mammifères est généralement pauvre en dinucléotides CpG par rapport à la probabilité attendue, excepté dans les îlots CpG. Ces séquences sont définies comme faisant plus de 500 paires de bases avec un contenu en guanine et cytosine de plus de 55% et un ratio de dinucléotides CpG observé/attendu supérieur à 0,65 (Jones and Takai, 2001; Takai and Jones, 2003). Environ 40% des gènes de mammifères possèdent un îlot CpG majoritairement au niveau de leur promoteur et du premier exon.

### 1-2-1) Le "méthylome"

L'étude de la méthylation de l'ensemble des CpG d'un génome (ou méthylome) a été rendue possible par l'avènement de techniques haut débit comme les microarrays. Les techniques d'immunoprécipitation d'ADN méthylé (MeDIP pour Methyl-DNA ImmunoPrecipitation) associées aux puces à ADN ont permis d'établir une carte des domaines constants de méthylation chez les mammifères (Weber *et al.*, 2005; Weber and Schubeler, 2007). Ainsi, les îlots CpG sont généralement faiblement méthylés. En revanche, les exons et les séquences géniques hors promoteurs présentent des plus hauts degrés de méthylation. Enfin, les séquences répétées et les séquences mobiles des rétrotransposons sont en général méthylées.



**Figure 4 : Evolution du profil de méthylation au cours du Développement.**

Deux grandes vagues de déméthylation/reméthylation ont lieu.

**A.** La première déméthylation a lieu au cours de la gamétogenèse dans les cellules germinales primordiales (PGC) et touche à la fois le génome maternel et paternel (violet). Le nouveau profil de méthylation est apposé plus tard sur le génome maternel (rouge) au cours de l'ovogenèse que sur le génome paternel (bleu) au cours de la spermatogenèse.

**B.** La seconde vague a lieu juste après la fécondation. La déméthylation est rapide et active pour le génome paternel et passive pour le génome maternel. Les gènes soumis à l'Empreinte sont protégés de cette déméthylation (gris). Le génome de l'embryon (violet) est ensuite reméthylié dans les lignages somatiques et extraembryonnaires.

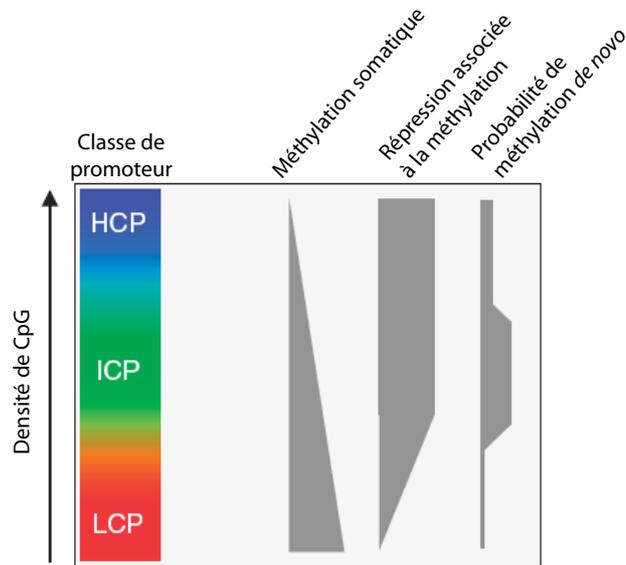
(Adapté de Reik *et al.*, 2001)

Par ailleurs, la méthylation évolue au cours de la vie d'un individu. Il y a deux grandes vagues de reprogrammation de la méthylation : au cours de la gamétogenèse et au cours des premières divisions après fécondation (Figure 4).

**Dans les cellules germinales primordiales** (PGC pour Primordial Germ Cells), le génome subit une forte déméthylation. Au cours de la spermatogenèse, l'apposition d'un nouveau profil de méthylation a lieu avant les divisions méiotiques. Chez la femelle, la reméthylation a lieu après les divisions méiotiques au moment de la maturation ovocytaire (Reik *et al.*, 2001). Cette phase de déméthylation/reméthylation serait importante pour effacer le programme d'expression et mettre en place un programme spécifique aux lignages des gamètes. De plus, ce phénomène est particulièrement important pour le cycle de l'Empreinte Parentale, comme nous le verrons par la suite. Les gamètes matures n'ont pas le même profil : le génome des spermatozoïdes est fortement méthylé et le génome de l'ovocyte au moment de l'ovulation l'est beaucoup moins (Monk *et al.*, 1987).

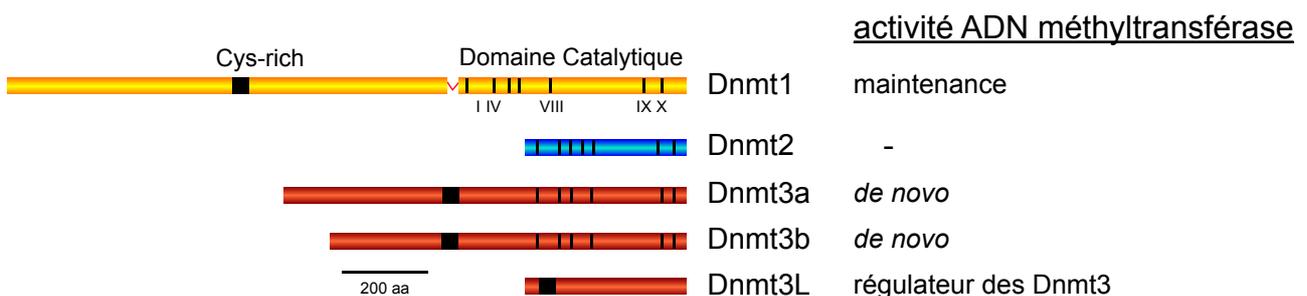
**L'autre phase importante de déméthylation a lieu immédiatement après la fécondation** de l'ovocyte par le spermatozoïde. Le **génom paternel** fortement condensé subit alors des changements majeurs dont l'un des plus frappant est le remplacement des protamines par les histones au niveau des nucléosomes de l'ADN. Il subit également une déméthylation rapide qui est achevée avant la répllication du génome zygotique (Mayer *et al.*, 2000). La vitesse à laquelle elle intervient suggère que cette déméthylation est active. Deux voies de déméthylation sont proposées : l'action d'une déméthylase et l'excision/réparation des méthyl-cytosines. A l'heure actuelle, aucune étude n'a permis de mettre en évidence d'enzyme catalysant la déméthylation de l'ADN.

Le **génom maternel** subit lui aussi une perte de méthylation mais avec une cinétique différente. Les résultats de Mayer *et al.* suggèrent que le génome maternel est tout d'abord méthylé pour atteindre un niveau comparable à celui du génome du spermatozoïde. Ensuite au moment de la répllication de l'ADN, il est déméthylé de façon lente, suggérant un mécanisme passif, en lien avec la localisation cytoplasmique de l'ADN méthyltransférase de maintenance Dnmt1 dans l'embryon préimplantatoire (Hirasawa *et al.*, 2008).



**Figure 5 : Le méthylome.**

L'analyse haut débit de la méthylation de l'ADN a permis de mettre en évidence des comportements différents pour trois classes de séquences. Les promoteurs à faible contenu en CpG (**LCP**) sont méthylés dans les cellules souches et différenciées et la méthylation n'est pas liée à l'activité du promoteur. Peu de promoteurs à haut contenu en CpG (**HCP**) sont méthylés dans les cellules somatiques mais lorsqu'ils le sont, leur activité est plutôt réprimée. Les promoteurs à contenu en CpG intermédiaire (**ICP**) sont sujets à la méthylation *de novo*, présentent plus fréquemment des différences de méthylation entre cellules souches et cellules différenciées et entre les différents types de cellules différenciées. La méthylation de ces promoteurs est souvent en rapport avec une répression transcriptionnelle.



**Figure 6 : La famille des DNA méthyltransférases (Dnmt).**

Cinq Dnmt ont été caractérisées et présentent des régions riches en cystéines (Cys-rich) et un domaine catalytique conservés. Dnmt1 est une méthyltransférase de maintenance, elle copie la méthylation du brin parental sur le brin néosynthétisé au cours de la réplication. Dnmt2 n'a pas d'activité méthyltransférase pour l'ADN mais pour l'ARNt. Les protéines Dnmt3a et 3b ont une activité *de novo*, elles apposent la méthylation sur un brin vierge. Enfin, Dnmt3L n'a pas d'activité méthyltransférase mais régule Dnmt3a et 3b, notamment au cours de la mise en place de l'Empreinte pendant la gamétogenèse. (Adapté de Bourc'his *et al.*, 2001)

Au stade 32 cellules, les deux génomes ont un niveau bas de méthylation. Il est à noter que les gènes soumis à l'Empreinte sont protégés de cette vague de déméthylation (Reik *et al.*, 2001), ce qui est important pour le cycle de l'Empreinte comme nous le verrons par la suite.

Dès le stade blastocyste, un nouveau profil de méthylation est mis en place, ce qui est interprété comme la mise en place de nouveaux programmes d'expression spécifiques des différents lignages cellulaires (Reik *et al.*, 2001). Mais la méthylation des gènes au cours de la différenciation des gènes n'est pas claire et systématique. Une analyse haut débit de la méthylation dans des fibroblastes (cellules somatiques) en culture primaire versus des cellules germinales a permis d'établir 3 classes de promoteurs selon leur faible, moyenne ou forte densité en CpG (Weber *et al.*, 2007). Seulement 4% des promoteurs à îlots CpG sont méthylés dans les fibroblastes et non dans les cellules germinales et représentent en général des facteurs de transcription spécifiques de lignages cellulaires. Les promoteurs à faible densité en CpG sont en général méthylés et leur profil n'est pas modifié entre fibroblastes et cellules germinales. La méthylation *de novo* est en revanche plus significative pour les gènes à densité moyenne de CpG entre les cellules germinales et les fibroblastes. De plus, les gènes spécifiques des lignages des cellules germinales sont plus fréquemment méthylés dans les cellules somatiques différenciées (Figure 5). Cette étude concorde avec les résultats publiés sur des gènes individuels ou sur des chromosomes particuliers (Weber and Schubeler, 2007).

### **1-2-2) Les ADN méthyltransférases (Dnmt)**

La première ADN méthyltransférase, Dnmt1, a été décrite et le gène a été cloné en 1988 (Bestor *et al.*, 1988). Par analogie de séquence, quatre autres gènes murins ont été identifiés : Dnmt2, Dnmt3a, Dnmt3b et Dnmt3L (Altschul *et al.*, 1990; Aapola *et al.*, 2000) (Figure 6). Ces enzymes ont un domaine catalytique qui présente des motifs de reconnaissance des substrats SAM et ADN et d'activation de la cytosine cible et un domaine riche en cystéine (Cys-rich). Cependant Dnmt2 ne possède pas d'activité ADN méthyltransférase *in vivo* mais serait impliquée dans la méthylation d'ARNt (Okano *et al.*, 1998; Goll *et al.*, 2006). Dnmt3L ne possède pas les motifs caractéristiques du domaine catalytique mais est un régulateur de Dnmt3a comme nous allons le voir par la suite (Bourc'his *et al.*, 2001).

Les ADN méthyltransférases ont deux fonctions : une **activité de méthylation *de novo***, pour établir la modification à partir d'une séquence qui en est dépourvue et une **activité de méthylation de maintenance**, qui permet de recopier les marques pour transmettre l'information aux cellules filles lors de la division cellulaire.

### a) La méthylation de maintenance : Dnmt1

L'ADN méthyltransférase de maintenance permet de recopier le patron de méthylation de la cellule mère et de le transmettre aux cellules filles au cours de la mitose. Dnmt1 a une forte affinité pour les dinucléotides CpG hémiméthylés. Au cours de la phase S du cycle cellulaire, Dnmt1 est localisé au niveau des fourches de réplication et méthylerait ainsi l'ADN au cours de sa synthèse (Leonhardt *et al.*, 1992).

Le maintien de l'information détenue par la méthylation de l'ADN semble très important pour la viabilité des individus. En effet, l'inactivation de Dnmt1 chez la souris conduit à une létalité embryonnaire précoce peu après l'implantation (Lei *et al.*, 1996). Les embryons Dnmt1<sup>-/-</sup> dégèrent dès E9.5 (jour embryonnaire 9,5) au stade 8 somites. Les cellules ES dérivées de ces mutants sont cependant viables et présentent une réduction très forte mais non totale de la méthylation. L'infection de ces cellules par un virus a permis de montrer qu'une activité de méthylation *de novo* existe dans ces cellules. Dans l'ensemble, l'étude de ce mutant a permis de montrer que le maintien de la méthylation par Dnmt1 au travers des divisions mitotiques était crucial pour le développement et que la méthylation *de novo* était assurée par une ou plusieurs autres enzymes.

Par ailleurs, comme nous le verrons par la suite, Dnmt1 est nécessaire et suffisant pour maintenir le profil de méthylation différentiel des gènes soumis à l'Empreinte dans l'embryon préimplantatoire (Hirasawa *et al.*, 2008).

### b) La méthylation de novo : Dnmt3a et Dnmt3b

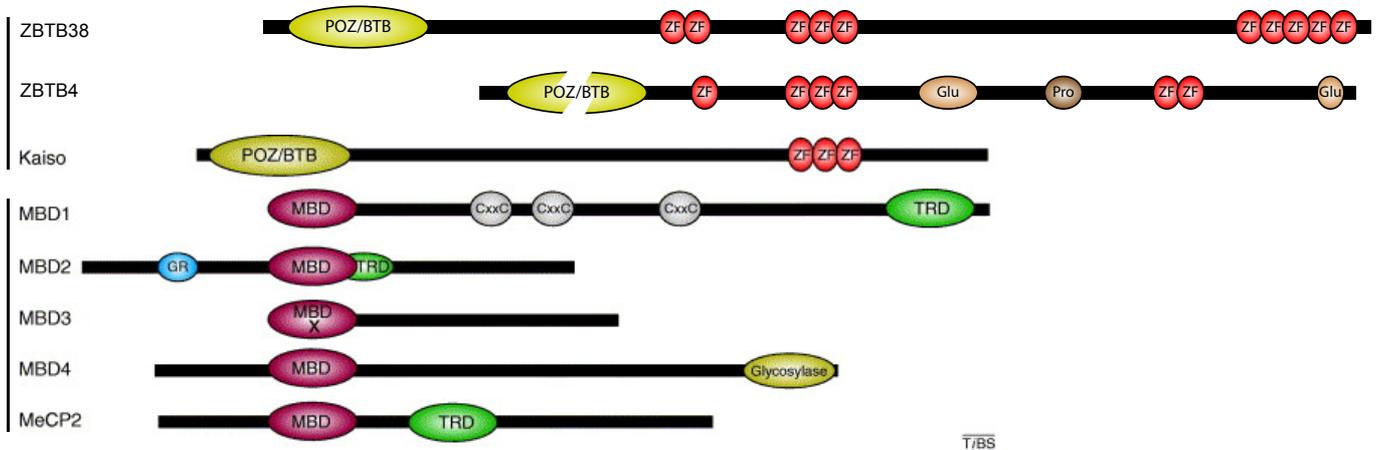
Au cours du développement, il est nécessaire d'établir un nouveau patron de méthylation spécifique des lignages cellulaires après la déméthylation globale des génomes parentaux. Cette fonction est assurée par les Dnmt3a et 3b. L'étude des mutants invalidés pour

chacun de ces gènes a montré que ces deux enzymes sont indispensables pour la viabilité des souris et qu'elles ont chacune un rôle bien défini (Okano *et al.*, 1999).

Les souris  $Dnmt3a^{-/-}$  apparaissent normales à la naissance mais meurent dans les 4 premières semaines de vie alors que les embryons  $Dnmt3b^{-/-}$  meurent *in utero* entre E9.5 et la naissance. Cela suggère que ces enzymes ont des fonctions distinctes. Les doubles mutants  $Dnmt3a3b^{-/-}$  présentent des défauts majeurs dès E8.5 et meurent aux environs de E11.5. L'aggravation du phénotype par rapport aux simples mutants montre qu'il existe également une redondance fonctionnelle entre ces enzymes au début du développement embryonnaire (Okano *et al.*, 1999).

Parmi les fonctions propres à  $Dnmt3b$ , la méthylation des séquences répétées satellites péri-centromériques du génome serait particulièrement importante. En effet, la mutation de cette enzyme est responsable du syndrome humain ICF (Immunodeficiency, Centromeric instability and Facial abnormalities), caractérisé par l'hypométhylation de ces satellites et une instabilité de l'hétérochromatine centromérique associée (Miniou *et al.*, 1994).

Un rôle particulier pour  $Dnmt3a$  a été montré grâce à sa capacité de liaison à  $Dnmt3L$ . Cette protéine 3L est dépourvue du domaine catalytique méthyltransférase et pourtant, l'inactivation du gène chez la souris entraîne des défauts de méthylation. En effet, les mâles  $Dnmt3L^{-/-}$  sont stériles et les femelles sont fertiles mais leur descendance meurt *in utero* (Bourc'his *et al.*, 2001). Chez ces embryons, la méthylation globale du génome n'est pas affectée mais les séquences spécifiques des gènes soumis à empreinte sont en revanche très atteintes. C'est l'absence de marques associées à l'Empreinte, normalement appliquées au cours de l'ovogenèse, qui serait responsable de la létalité embryonnaire. Pendant la spermatogenèse, des défauts d'acquisition de la méthylation aux loci soumis à Empreinte ont également été observés en plus de défauts de méthylation des séquences hétérochromatiques et des répétitions non péri-centriques. De plus, au moment de la méiose, les chromosomes ne s'apparient pas, ce qui aboutit à un arrêt de la spermatogenèse, et entraîne une infertilité chez les mâles  $Dnmt3L^{-/-}$  (Webster *et al.*, 2005).  $Dnmt3L$  pourrait interagir physiquement avec  $Dnmt3a$  et  $Dnmt3b$  pour potentialiser l'activité méthyltransférase de ces enzymes (Suetake *et al.*, 2004).



**Figure 7 : Les principales protéines de liaison à l'ADN méthylé (MBP).**

Deux grandes familles de MBP sont répertoriées : la famille MBD et la famille Kaiso. Les protéines de la famille Kaiso ont un domaine de répression transcriptionnelle POZ/BTB (poxvirus and zinc finger/broad complex, tramtrack and bric à brac) et un domaine de liaison à l'ADN en doigt de zinc (ZF). Les membres de la famille MBD possèdent un domaine de liaison aux méthyl CpG (MBD methylCpG Binding Domain) et les trois principaux régulateurs de la transcription ont un domaine TRD (transcription repression domain).

(Adapté de Klose and Bird, 2006 et Filion *et al.*, 2006)

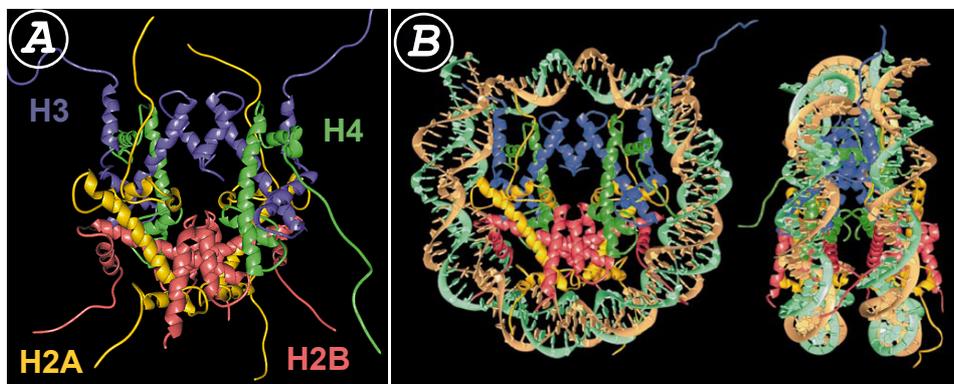
Ainsi, la méthylation *de novo* est assurée par les protéines Dnmt3. Dnmt3a et Dnmt3L formeraient un complexe pour réguler la méthylation des gènes soumis à l'Empreinte au cours de la gamétogenèse. Dnmt3b serait plus particulièrement impliquée dans l'hétérochromatine péricentrique. De plus, Dnmt3a et 3b pourraient intervenir sur d'autres séquences comme les séquences transposables du génome.

### **1-2-3) Les protéines de liaison à l'ADN méthylé (MBP)**

Deux mécanismes non exclusifs sont proposés pour expliquer l'action de répression transcriptionnelle de la méthylation : une impossibilité pour certaines protéines de se lier à l'ADN méthylé et/ou le recrutement de protéines liant spécifiquement l'ADN méthylé. Dans la première hypothèse, la méthylation empêcherait l'accès des facteurs de transcription ou d'autres protéines régulatrices à leurs sites de reconnaissance spécifiques. C'est par exemple le cas de la protéine CTCF (Bell and Felsenfeld, 2000), importante pour la régulation des gènes *H19* et *Igf2* et dont nous reparlerons donc ultérieurement. Le second mécanisme propose que les <sup>me</sup>CpG puissent recruter spécifiquement des protéines et par cet intermédiaire participer à la répression transcriptionnelle. Effectivement, deux grandes familles de protéines liant spécifiquement l'ADN méthylé (MBP pour Methyl Binding Proteins) ont été décrites (Figure 7) : les protéines à domaine MBD et les protéines à domaine POZ/BTB.

La première protéine de la **famille MBD** à avoir été décrite est MeCP2 (Methyl CpG-binding Protein) (Lewis *et al.*, 1992). Cette protéine possède un domaine de liaison à l'ADN méthylé (MBD pour methyl binding domain). Trois autres membres ont été caractérisés : MBD1, MBD2 et MBD4 possédant tous ce domaine MBD et pouvant lier l'ADN méthylé. MeCP2, MBD1 et MBD2 possèdent également un domaine de répression de la transcription (TRD). L'invalidation chez la souris de ces trois protéines engendre des phénotypes neurologiques plus ou moins importants. Ces protéines seraient impliquées dans la répression transcriptionnelle via l'association avec des protéines de remodelage de la chromatine. MBD4 possède un domaine glycosylase impliqué dans la réparation de l'ADN (Klose and Bird, 2006; Miranda and Jones, 2007).

Une autre **famille de protéines à doigts de zinc** liant l'ADN méthylé a été découverte à partir de son membre fondateur Kaiso. Kaiso est capable, grâce à un domaine à trois doigts



**Figure 8 : Le nucléosome.**

- A.** Interaction entre les deux dimères d'histones H2A-H2B et deux dimères d'histones H3-H4 pour former l'octamère d'histone.
- B.** Autour de l'octamère d'histone, l'ADN s'enroule en un peu moins de 2 tours pour former le nucléosome.
- (Adapté de Luger *et al.*, 1997)

de Zinc (ou ZF pour Zinc-Finger), de lier une séquence contenant deux CpG symétriquement méthylés (Prokhortchouk *et al.*, 2001). Le même domaine peut également lier une séquence consensus KBS non méthylée. Par analogie de séquence, deux autres protéines ont été identifiées possédant les mêmes motifs à trois ZF : ZBTB4 et ZBTB38 (Filion *et al.*, 2006). Ces deux protéines peuvent lier un CpG méthylé et ZBTB4 est également capable de lier la séquence KBS. Ces trois protéines sont connues pour être des répresseurs transcriptionnels.

La littérature fait mention d'une autre protéine capable de lier plutôt l'ADN hémiméthylé via un domaine SRA (SET and RING Associated) : UHRF1 (ubiquitin-like, containing PHD and RING finger domains 1) (Bostick *et al.*, 2007). Cette protéine serait impliquée dans la maintenance de la méthylation en interagissant avec Dnmt1.

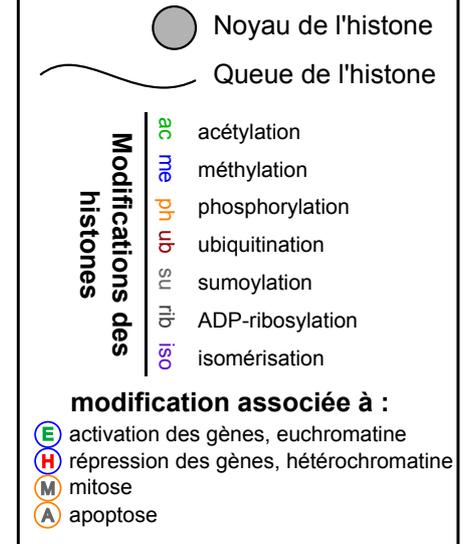
Les mécanismes de répression transcriptionnelle des MBP seront discutés plus tard, en lien avec les protéines de modification de la chromatine.

**A**

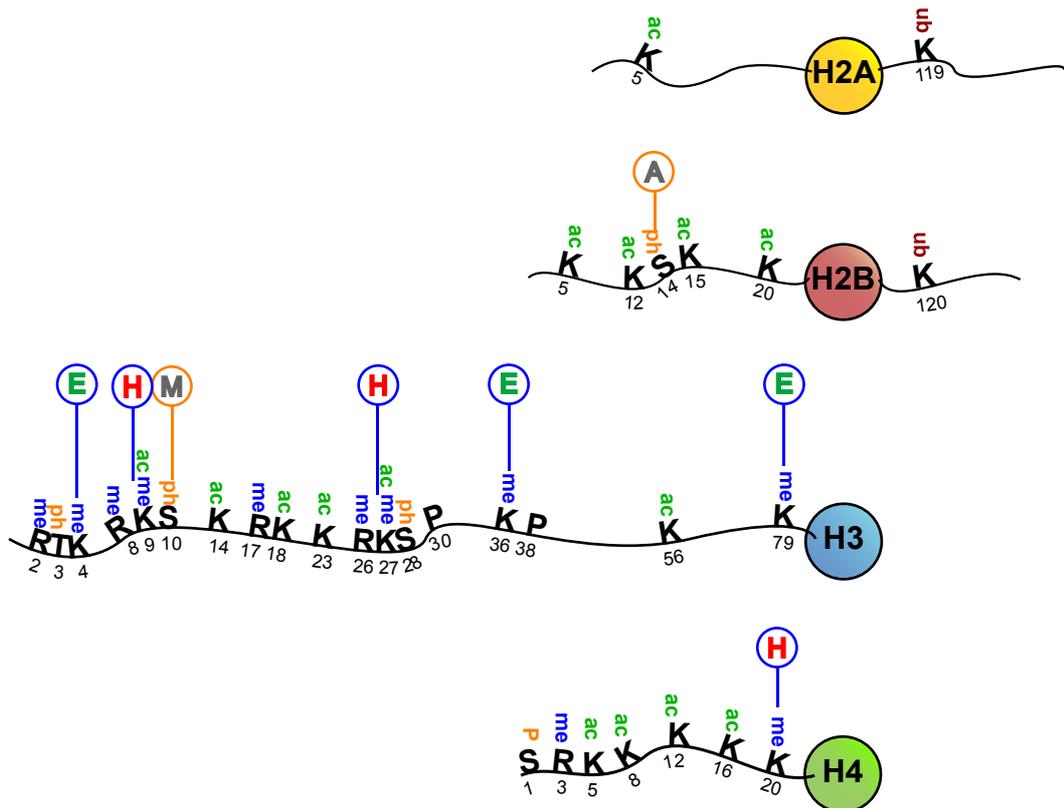
### Principales modifications associées aux différents acides aminés des queues des histones

Lysine	<b>K</b>	me	ac	ub	su
Sérine	<b>S</b>	ph			
Arginine	<b>R</b>	me			
Thréonine	<b>T</b>	ph			
Proline	<b>P</b>	iso			
Acide Glutamique	<b>E</b>	rib			

### Légende

**B**

### Principales modifications des acides aminés des queues des histones et signaux associés



**Figure 9 : Les modifications des histones.**

**A.** Les principales modifications et les principaux acides aminés modifiés.

**B.** Sur les queues N et C terminales des histones, différents résidus peuvent être modifiés. Ces modifications peuvent être spécifiques d'un état de la chromatine : hétéro (H3K9me, H3K27me, H4K20me) ou euchromatine (H3K4me, H3K36me, H3K79me) ; ou d'un processus cellulaire : apoptose (H2BS14ph), compaction et ségrégation des chromosomes mitotiques (H3S10ph). (Adapté de Kouzarides, 2007).

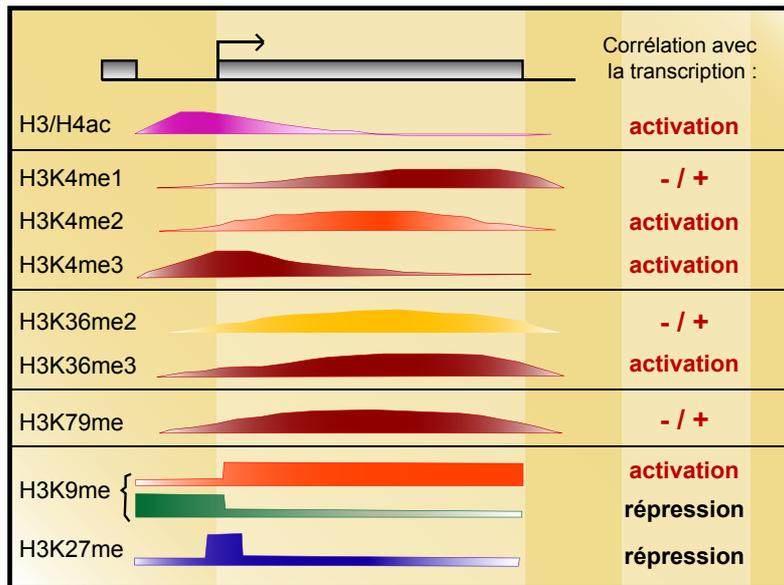
### 1-3) Les modifications des histones

L'unité fonctionnelle de la chromatine est le nucléosome (Figure 8). Il s'agit d'un octamère des quatre protéines histones H2A, H2B, H3 et H4 autour duquel s'enroulent approximativement deux tours de la double hélice d'ADN, ce qui correspond à 147 paires de bases (Luger *et al.*, 1997). Ces protéines histones sont globulaires avec les extrémités N et C terminales libres. Les parties globulaires des histones du nucléosome interagissent entre elles pour former le noyau (core en anglais) autour duquel s'enroule l'ADN. Les extrémités libres ou queues d'histones sont accessibles et peuvent être modifiées par des enzymes de remodelage de la chromatine.

Aux quatre histones du nucléosome s'ajoute une histone H1 impliquée dans les formes compactes de la chromatine. De plus, il existe des **variants d'histones** (notamment H3.2, H3.3, macroH2A, H2AZ, H2ABbd, H2A.X) qui divergent des histones canoniques H2A et H3 au niveau des queues d'histones ou encore au niveau des acides aminés clefs. Ces variants sont impliqués dans l'activation ou la répression transcriptionnelle, l'inactivation du chromosome X ou la réparation des dommages de l'ADN (Li *et al.*, 2007).

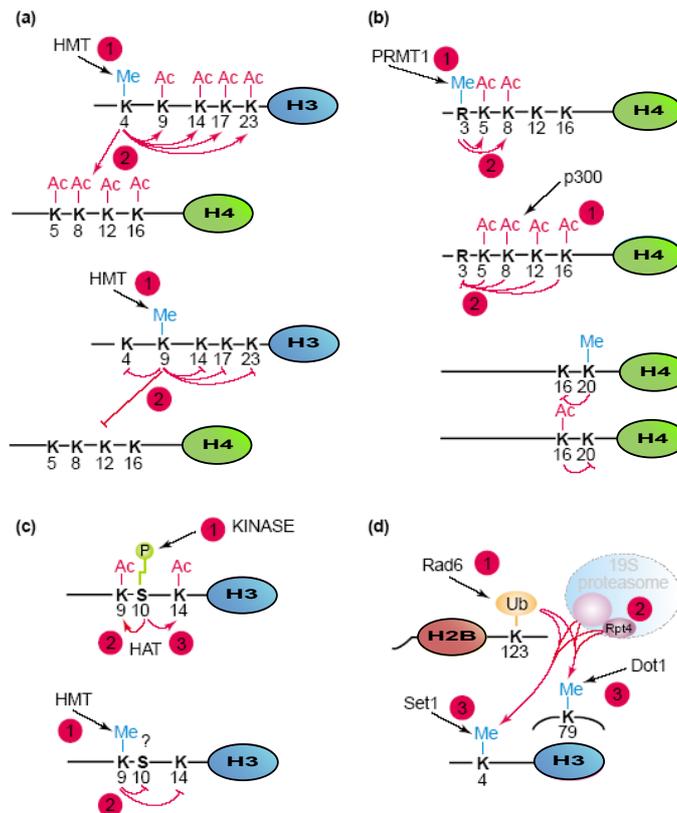
#### 1-3-1) Les modifications post-traductionnelles des histones

Il y a plus de 70 sites de modifications post-traductionnelles recensées sur les histones des nucléosomes et au moins huit classes de liaisons covalentes différentes ont été détectées sur une dizaine d'acides aminés distincts. Ces modifications sont l'acétylation de la lysine, la méthylation des lysines et arginines, la phosphorylation des sérines et thréonines, l'ubiquitination et la sumoylation de la lysine, l'ADP-polyribosylation de l'acide glutamique, la déamination de l'arginine et l'isomérisation de la proline (Kouzarides, 2007). Les principales modifications sont reportées sur la Figure 9. Ces modifications ont notamment pu être identifiées grâce aux techniques de spectrométrie de masse et à l'utilisation d'anticorps spécifiques des différentes modifications des histones. La diversité des marques recensées s'accroît rapidement et le nombre de modifications détectées devrait encore augmenter dans les prochaines années.



**Figure 10 : Modifications des histones et transcription.**

Grâce à l'analyse haut débit des modifications des histones, des profils ont pu être dessinés en rapport avec la structure d'un gène. Les indications +/- indiquent que la modification post-traductionnelle n'est pas toujours en relation avec l'activation ou la répression transcriptionnelle. (Adapté de Li *et al.*, 2007)



**Figure 11 : Dialogue entre les marques des histones.**

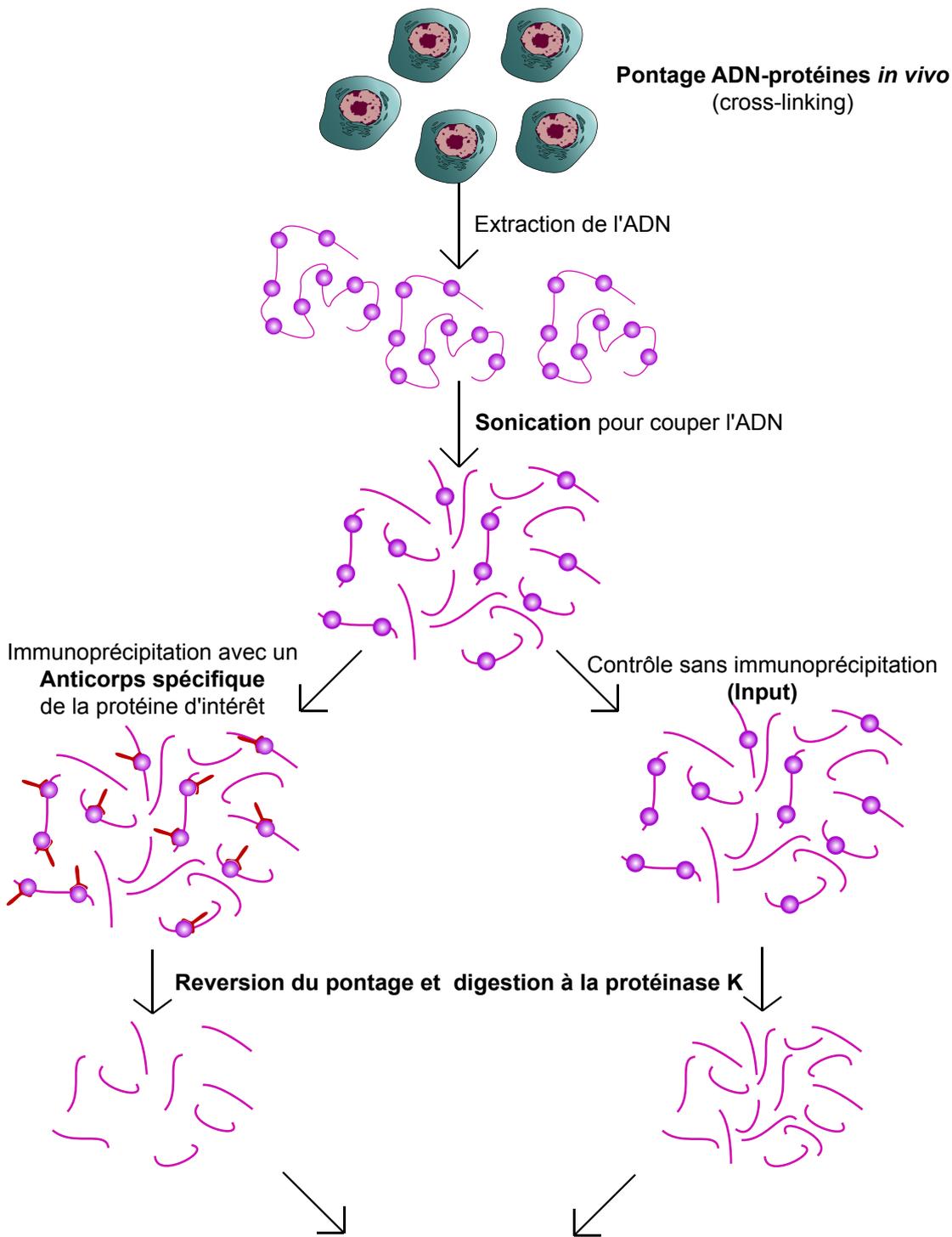
Les marques des histones s'influencent les unes les autres. (a) Ainsi, H3K4me favorise l'acétylation de H3 et H4 alors que H3K9me la défavorise. (b) H4R3me favorise l'acétylation des lysines adjacentes alors que l'acétylation de H4 inhibe H4R3me. (c) H3S10ph favorise l'acétylation des lysines adjacentes alors que H3K9me les inhibe. (d) L'ubiquitination de H2B favorise la méthylation de H3K9 et H3K79. Les ronds rouges indiquent un ordre chronologique supposé. (Adapté de Margueron *et al.*, 2005).

La phosphorylation de la sérine 10 de l'histone H3 (H3S10ph)<sup>1</sup> est impliquée dans la mitose et la ségrégation des chromosomes. L'acétylation des histones H3 et H4 caractérise les domaines d'activité transcriptionnelle de l'euchromatine (Grunstein, 1997). La méthylation des histones est plus complexe, les lysines peuvent être monométhylées, diméthylées ou triméthylées et les arginines peuvent être monométhylées ou diméthylées. Les réponses sont fonction du résidu impliqué et de sa position sur la queue de l'histone (Li *et al.*, 2007). Par des techniques haut débit de ChIP-on-chip ou ChIP-Seq (Figure 12) avec des anticorps spécifiques de chacune de ces modifications, un patron de distribution sur le génome et au niveau des gènes a pu être établi. Ainsi les modifications H3K4me, H3K36me et H3K79me sont associées à l'euchromatine alors que H3K9me, H3K27me et H4K20Me sont hétérochromatiques. Plus précisément, H3K4me<sub>3</sub> est présente au niveau des promoteurs des gènes à îlots CpG et le taux de méthylation est corrélé à l'expression du gène (Figure10). Cette marque est également concomitante avec l'acétylation des histones H3 et H4. H3K36me<sub>3</sub> est présente au niveau de la séquence transcrite du gène et serait impliquée dans l'élongation de la transcription. H3K9me<sub>3</sub> et H4K20me<sub>3</sub> marquent spécifiquement les séquences répétées du génome (télomères, satellites, LTR) et seraient donc impliquées dans l'hétérochromatine constitutive. H3K27me<sub>3</sub> est plutôt associée à l'hétérochromatine facultative ou la répression transcriptionnelle dans les régions euchromatiques (Kouzarides, 2007; Li *et al.*, 2007; Mikkelsen *et al.*, 2007; Wang *et al.*, 2007a).

Dans les cellules souches embryonnaires (ES), des promoteurs "bivalents" portent les marques opposées H3K4me<sub>3</sub> et H3K27me<sub>3</sub>. Ces promoteurs bivalents "en équilibre" pourraient permettre à des gènes importants pour le développement et la différenciation, d'être dirigés vers un état actif ou réprimé selon le destin de la cellule fille (Azucara *et al.*, 2006; Bernstein *et al.*, 2006; Mikkelsen *et al.*, 2007). Un exemple frappant est celui du gène du facteur de transcription *Olig1*, spécifique des neurones, qui est bivalent dans les cellules ES et conserve la marque répressive H3K27me<sub>3</sub> dans les fibroblastes alors qu'il garde la marque active H3K4me<sub>3</sub> dans les progéniteurs neuronaux.

---

<sup>1</sup> Par convention, les marques de histones sont notés de la façon suivante : nom de l'histone suivi de la lettre de l'acide aminé et de son numéro à partir du N-ter de la protéine suivi de la modification (ac pour acétylation, me pour méthylation, ph pour phosphorylation) et du nombre de liaisons covalentes formées (1, 2 ou 3 ; si ce n'est pas précisé, cela indique l'ensemble de ces trois modifications). Ainsi, par exemple la triméthylation de la lysine 27 de l'histone H3 sera notée H3K27me<sub>3</sub>.



**Mesure de l'enrichissement dans la fraction immunoprécipité vs Input**

- par PCR pour des loci spécifiques
- par puce (ChIP-on-chip) pour une analyse précise d'un locus donné (tiling arrays) ou globale d'un génome.
- par séquençage (ChIP-Seq) pour une analyse globale d'un génome.

**Figure 12 : Technique d'Immunoprécipitation de la chromatine (ChIP).**

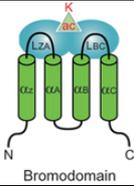
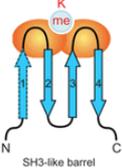
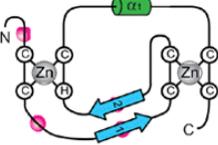
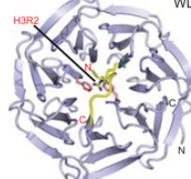
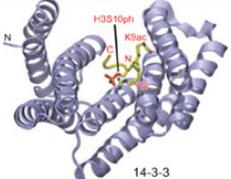
Cette technique permet d'étudier la fixation d'une protéine à ses séquences cibles de l'ADN ou d'analyser les modifications des histones et leur répartition sur le génome. La comparaison de la fraction immunoprécipitée et de la fraction non-immunoprécipitée (input) permet de connaître l'enrichissement qui mesure la fixation d'une protéine ou la présence des marques des histones sur une séquence donnée. Cette technique peut être utilisée en combinaison avec les analyses haut débit par puces (ChIP-on-chip) ou séquençage (ChIP-Seq).

Il existe un **dialogue entre les modifications post-traductionnelles** des histones, certaines étant mutuellement exclusives et d'autres favorisant des modifications ultérieures (Figure 11). De cette façon, les modifications d'un même acide aminé, par exemple méthylation et acétylation des lysines, ne peuvent être présentes en même temps sur le même résidu. La modification peut également empêcher la substitution sur un acide aminé proche, ainsi l'acétylation de H4K16 et la méthylation de H4K20 peuvent se réprimer l'une l'autre. De même H3K9me inhibe la phosphorylation de H3S10. En revanche H3S10ph favoriserait l'acétylation de H3K9 et H3K14. Il a également été rapporté que l'ubiquitination de l'histone H2B était requise pour la méthylation de H3K4 (revue par (Margueron *et al.*, 2005)).

De nombreuses études tentent d'établir le lien entre les modifications des histones et l'architecture de la chromatine. Deux mécanismes sont décrits : l'action directe des modifications et le recrutement de protéines effectrices. Ces deux modèles ne sont pas mutuellement exclusifs et il est tout à fait probable qu'ils coexistent pour modérer la chromatine de façon appropriée à un locus donné.

La possibilité que les modifications post-traductionnelles des queues d'histones agissent directement sur la structure de la chromatine a été décrite pour l'acétylation en particulier. Les lysines possèdent une chaîne latérale ionisable basique chargée positivement à pH physiologique. Cette charge pourrait jouer un rôle important pour les interactions entre histone et ADN (chargé négativement en raison des nombreux groupements phosphate) ou encore entre les histones des nucléosomes adjacents. L'ajout covalent du groupement acétyle (COCH<sub>3</sub>) annule la charge de l'acide aminé et pourrait ainsi réduire la force de la liaison avec l'ADN et les nucléosomes adjacents, ce qui permettrait d'ouvrir la chromatine. Cette hypothèse a récemment été appuyée par la démonstration que l'acétylation de la lysine 16 de l'histone H4 (H4K19ac) inhibe la formation des structures compactes de la chromatine comme la fibre de 30nm (Shogren-Knaak *et al.*, 2006).

Le second mécanisme suppose que les modifications post-traductionnelles de la chromatine servent de plateformes sur lesquelles sont recrutées des protéines effectrices impliquées dans le remodelage de la chromatine et l'activation ou la répression transcriptionnelle. Les modifications post-traductionnelles formeraient ainsi un **"code des histones"** apposées par des "écrivains" (writers), gommées par des "effaceurs" (erasers) et lues par des "lecteurs" (readers). En fonction du contexte cellulaire, des modifications sont apposées, ces modifications sont ensuite verrouillées et lues pour ouvrir ou fermer certains loci à la transcription (Jenuwein and Allis, 2001).

Domaines	Modifications cibles	
 <p>Bromodomain</p>	Kac	
 <p>Superfamille Royale</p>	Chromodomaine Double chromodomaine Double Tudor Tandem Tudor Répétitions MBT	H3K9me2/3, H3K27me2/3 H3K4me1/2/3 H3K4me3 H3K20me1/2 Kme1/2
 <p>PHD finger family</p>	PHD finger	H3K4me2/3 H3K4me0
 <p>WD40</p>	Répétition WD40	H3R2me0
 <p>14-3-3</p>	14-3-3	H3S10P

**Tableau 1 : Les domaines de liaison aux histones.**

Les protéines à bromodomaine sont capables de lier les histones acétylées. Les sites de méthylation des histones sont reconnus par les protéines à chromodomaine, à domaine Tudor, à répétitions MBT ou WD40 ou à PHD finger. La phosphorylation est quand à elle reconnue par les protéines à domaines 14-3-3. (Adapté de Taverna *et al.*, 2007)

### **1-3-2) L'écriture et l'effacement des marques des histones**

De nombreuses protéines ont été décrites pour leur capacité à catalyser les réactions de modifications post-traductionnelles des histones. Ces enzymes ont été étudiées dans des systèmes modèles très variés comme les levures *S. pombe* et *S. cerevisiae*, la Drosophile, la souris ou encore l'humain et sont pour la plupart conservées entre ces espèces. Elles peuvent porter des noms différents en fonction du modèle d'étude et j'utiliserai par la suite la nouvelle nomenclature proposée en 2007 (Allis *et al.*, 2007).

Pour apposer ces marques aux histones, il a été recensé plus d'une vingtaine d'histones acétyltransférases (KAT pour lysineAcetylTransferase), près de trente lysine-méthyltransférases (KMT), cinq arginine-méthyltransférases (CARM et PRMT) et des Serine/thréonine phosphatases ainsi que des ubiquitilases. Les marques sont effacées par des histones désacétylases (HDAC), des lysines déméthylases (KDM) (Kouzarides, 2007). Les KAT, KMT, HDAC et KDM ciblent les histones de manière très spécifique, ainsi H3K4 et H3K36 ne sont pas méthylées par les mêmes enzymes. Cela permet certainement une plus grande finesse dans la réponse aux signaux cellulaires.

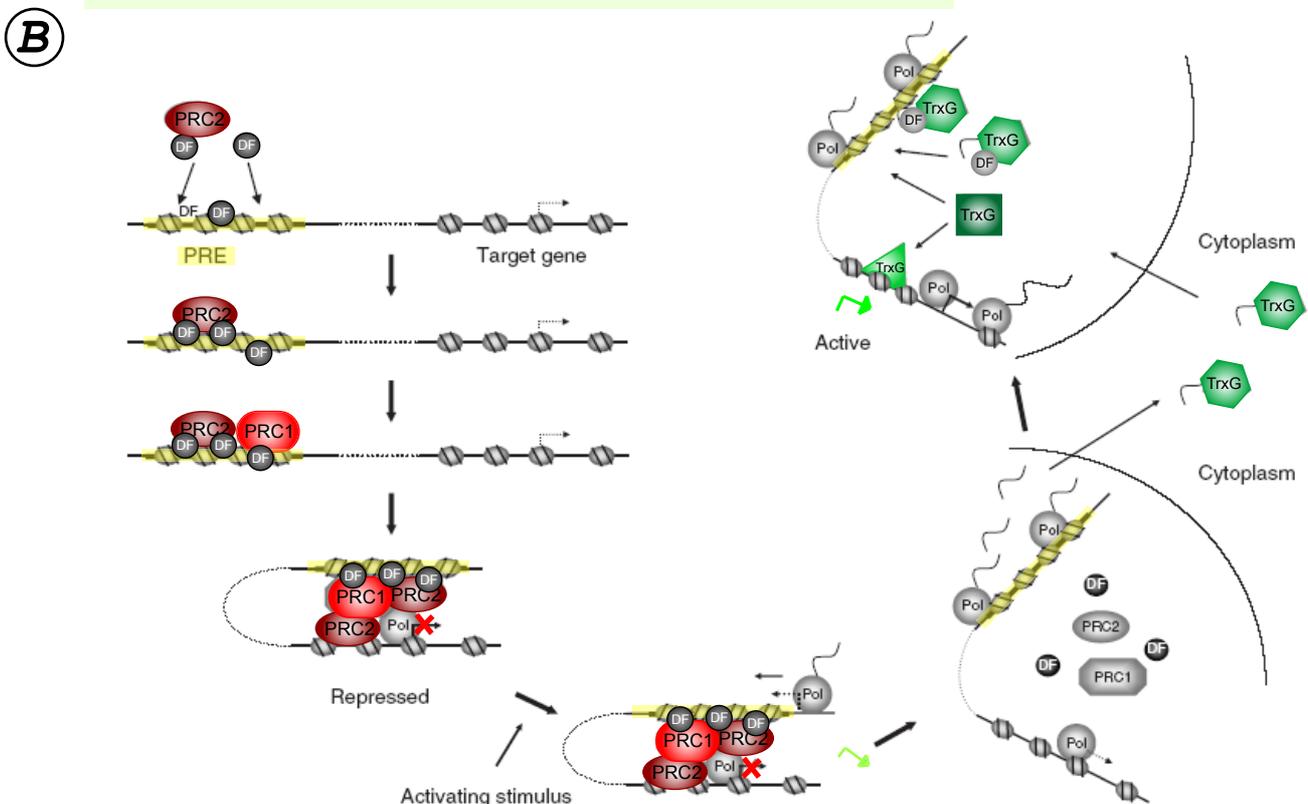
### **1-3-3) La lecture des marques des histones**

Dans la suite de la théorie du code des histones, de nombreuses protéines ont été décrites comme liant spécifiquement à une modification post-traductionnelle précise d'une histone. Par homologie de séquence, certains domaines protéiques ont ainsi pu être mis en évidence ([Tableau 1](#)).

Ainsi, l'acétylation des histones ne semble pas jouer uniquement un rôle de par la modification de charge associée mais peut également servir de plate-forme pour les protéines à bromodomaine. La méthylation des lysines est également spécifiquement reconnue par deux familles de domaine à "cage aromatique" : la superfamille Royal et la famille PHD-finger. Le motif de répétition WD40 est quand à lui spécifique des arginines non méthylées. Enfin la marque des phosphosérines recrute les protéines 14-3-3 (Taverna *et al.*, 2007).

**A**

PRC2		PRC1
KMT6	→ H3K27me3 →	HPC1-3
EED		HPH1-3
SUZ12		BMI1
RpAp46/48		SCMH1-3
		RING1A
		TBP associated factors
MLL1-3		SWI/SNF
KTM2A-C	→ H3K4me	BRM
WDR5		BAF250
ASH2L		BAF170
RbBP5		BAF47
CFP1		



**Figure 13 : Les complexes Polycomb et Trithorax.**

**A.** Composition de deux complexes de protéines du groupe Polycomb (PcG en rouge) : PRC1 et PRC2 (PcG Repressive complex). KMT6 méthyle H3K27, cette modification permet le recrutement des protéines HPC1-3. Composition de deux complexes de protéines du groupe Trithorax (TrxG en vert) : MLL1-3 et SWI/SNF.

**B.** Modèle de contrôle transcriptionnel par les protéines PcG et TrxG. En l'absence de transcription à travers les PRE (PcG response element en jaune), les complexes PRC2 puis PRC1 sont recrutés et répriment l'expression du gène (croix rouge). Un stimulus activateur entraîne la transcription à travers les PRE, libère les complexes PRC et recrute les complexes TrxG qui vont permettre l'activation du gène (flèche verte). (Adapté de Breiling *et al.*, 2007).

### 1-3-4) Polycomb et Trithorax

Les protéines Polycomb et Trithorax ont été à l'origine décrites comme des régulateurs de l'expression des gènes homéotiques Hox chez la Drosophile. Finalement il a été montré que ces protéines, organisées en complexes, sont en fait impliquées dans l'héritage épigénétique du patron d'expression des gènes Hox au cours du développement. Ces complexes sont plus largement capables de contrôler la répression ou l'activation de gènes et de garder la mémoire de cette régulation au cours des divisions cellulaires (Schuettengruber *et al.*, 2007). Ainsi, les protéines du groupe Polycomb (PcG pour Polycomb Group) sont impliquées dans la répression transcriptionnelle et les protéines du groupe Trithorax (TrxG) dans l'activation des gènes.

Plusieurs complexes de **protéines PcG** ont été identifiés grâce à des expériences de co-immunoprécipitation. Le complexe PRC1 (Polycomb Repressive Complex 1) comprend les protéines BMI1, HPC1-3, RING1, HPH1-3<sup>2</sup> et est associé à des facteurs de transcription ou des HDAC chez la Drosophile. Les protéines HPC possèdent un chromodomaine permettant le recrutement du complexe au niveau des H3K27me3. Les protéines principales du complexe PRC2 sont EED, KMT6 (EZH2), SUZ12. KMT6 permet la méthylation de H3K27 ([Figure 13A](#)). Plusieurs **complexes TrxG** ont été rapportés. Les deux principaux sont BRM (SWI/SNF) constitué de BRM, BAF47, BAF170 et BAF250 et MLL1-3 composé de KTM2A-C (MLL1-3), WDR5, ASH2L, RbBP5 et CFP1. Les KTM2 peuvent méthyler H3K4.

Chez la Drosophile, les protéines PcG et TrxG possèdent une séquence cible au niveau de l'ADN : les PRE (Polycomb Response Elements). Ces séquences n'ont pas pour l'heure été identifiées chez les mammifères. Aucune protéine du groupe Polycomb n'a été caractérisée comme capable de se fixer à l'ADN. Le recrutement aux PRE se ferait via l'intermédiaire de protéines accessoires comme Zeste ou Pho (Ringrose and Paro, 2004; Breiling *et al.*, 2007).

La [figure 13B](#) décrit un modèle simplifié d'action de répression ou d'activation des gènes par les complexes Polycomb et Trithorax impliquant l'interaction directe avec la machinerie transcriptionnelle, l'apposition de marques spécifiques sur les histones et des ARN non codants. Les protéines sont recrutées au locus grâce aux protéines accessoires. En

---

<sup>2</sup> J'ai choisi de nommer par défaut les protéines avec leur dénomination décrite chez les mammifères. Ainsi, chez la Drosophile, PRC1 comprend entre autres les protéines PC, PH, PSC, RING.

l'absence de transcription au niveau des PRE, le complexe PRC2 est recruté et peut recruter à son tour PRC1 via entre autres la méthylation de H3K27. Cela conduit à la formation de boucles et à la répression du gène cible. L'ARN polymérase II est localisée au niveau du promoteur mais l'élongation de la transcription est bloquée. En réponse à des signaux cellulaires, la transcription pourrait être possible au niveau des PRE, ce qui pourrait conduire à la libération des complexes PRC. Les ARN non codants produits recrutent alors un complexe Trx (Ash) qui peut recruter d'autres complexes Trx au niveau du gène cible et stabiliser ainsi l'activation transcriptionnelle du locus (via H3K4me ou acétylation) (Ringrose and Paro, 2004; Breiling *et al.*, 2007; Schuettengruber *et al.*, 2007).

### **1-3-5) Les autres modifications de la chromatine**

La structure de la chromatine fait également intervenir des **ATPases de remodelage de la chromatine**. Ces enzymes, hautement conservées depuis la levure jusqu'à l'humain, utilisent l'énergie de l'hydrolyse de l'ATP pour rompre le contact histone/ADN afin de déplacer les nucléosomes, retirer les histones sur le brin d'ADN ou permettre l'échange des variants d'histones. Quatre familles d'ATPases de remodelage de la chromatine sont identifiées : SWI/SNF, ISWI, NuRD/Mi2/CHD et ISO80 (Wang *et al.*, 2007b). Les domaines ATPases sont similaires mais ces enzymes ont des domaines d'interaction avec la chromatine différents et forment des complexes multiprotéiques qui assurent leur spécificité d'action.

## 1-4) Marques épigénétiques et hétérochromatine

Depuis la découverte de la méthylation de l'ADN et du "code des histones", une question importante est la hiérarchie de ces marques. C'est un peu l'histoire de l'œuf et de la poule. De nombreux rapports montrent que les MBP sont capables de recruter des protéines de modification des histones. D'un autre côté, il semble que les enzymes de remodelage de la chromatine puissent recruter les ADN méthyltransférases.

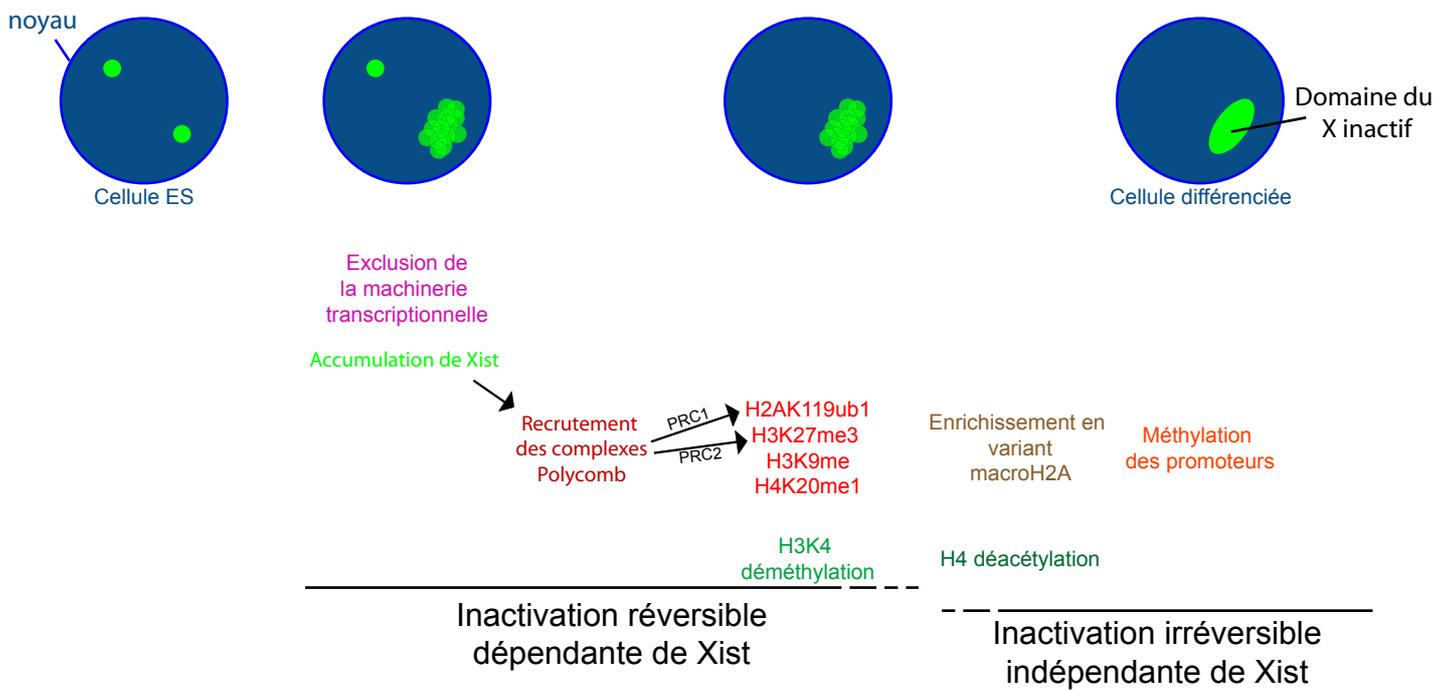
Dans un premier paragraphe, quelques exemples de dialogues seront brièvement exposés. Puis, l'exemple de l'inactivation du chromosome X, un phénomène faisant intervenir trois acteurs principaux de l'épigénétique (un ARN non codant, les histones et la méthylation de l'ADN), sera détaillé.

### 1-4-1) Dialogue entre les marques épigénétiques

De nombreuses études ont tenté de faire le lien entre les deux grandes marques épigénétiques. Des évidences d'un dialogue entre les marques de méthylation de l'ADN et de modifications des histones existent mais l'ordre hiérarchique de l'écriture et l'effacement de ces marques est encore discuté (Vaissiere *et al.*, 2008).

Ainsi certaines études placent la méthylation de l'ADN comme premier évènement de la répression transcriptionnelle au cours de la différenciation, du processus cancéreux ou comme mémoire épigénétique pendant la réplication. Les protéines de liaison à l'ADN méthylé en particulier joueraient un rôle important dans le recrutement des protéines de modification des histones. Il a notamment été montré que la protéine MeCP2 interagit avec des HDAC et qu'elle pourrait également intervenir dans la méthylation des histones pour réprimer la transcription (Nan *et al.*, 1998; Fuks *et al.*, 2003b). Par ailleurs, Dnmt1, qui assure la méthylation de maintenance de l'ADN au cours de la réplication, est associée à des HDAC au niveau des fourches de réplication, permettant ainsi la transmission de marques répressives des histones aux cellules filles avec la méthylation de l'ADN (Rountree *et al.*, 2000).

D'autres travaux placent les modifications des histones en amont de la méthylation de l'ADN. Il a ainsi été montré que, dans des cellules ES délétées pour KMT1 (lysine méthyltransférase des histones), l'hypométhylation de H3K9 ainsi induite était associée à l'hypométhylation des séquences satellites du génome. Le dialogue pourrait ici s'expliquer par



### Figure 14 : Chronologie des évènements au cours de l'inactivation du chromosome X.

L'inactivation du chromosome X fait intervenir de nombreuses marques épigénétiques selon un programme précis au cours de la différenciation des cellules ES. Dans le noyau, l'ARN *Xist* est tout d'abord transcrit à partir des deux allèles puis s'accumule sur l'un des chromosome et s'éteint sur l'autre. Il forme enfin un domaine sur le chromosome X inactif. Parallèlement, les marques actives des histones (vert) disparaissent et les marques inactives (rouge) sont enrichies sur le chromosome X inactif. L'inactivation est tout d'abord réversible et dépendante de *Xist* puis devient irréversible.

l'association de Dnmt3b en complexe avec KMT1 et HP1 (Fuks *et al.*, 2003a; Lehnertz *et al.*, 2003). Par ailleurs, l'utilisation d'inhibiteurs des HDAC sur des cellules en culture induit une hyperacétylation associée à une hypométhylation globale du génome et à une déméthylation spécifique de séquences particulières dans les promoteurs de certains gènes (Ou *et al.*, 2007).

Ces exemples illustrent bien le dialogue entre les marques. Il est probable que ces interactions ne soient pas unilatérales mais s'entretiennent plutôt les unes les autres, dans une boucle de rétrocontrôle, afin d'aboutir à une structure fermée et répressive de la chromatine (Fuks, 2005).

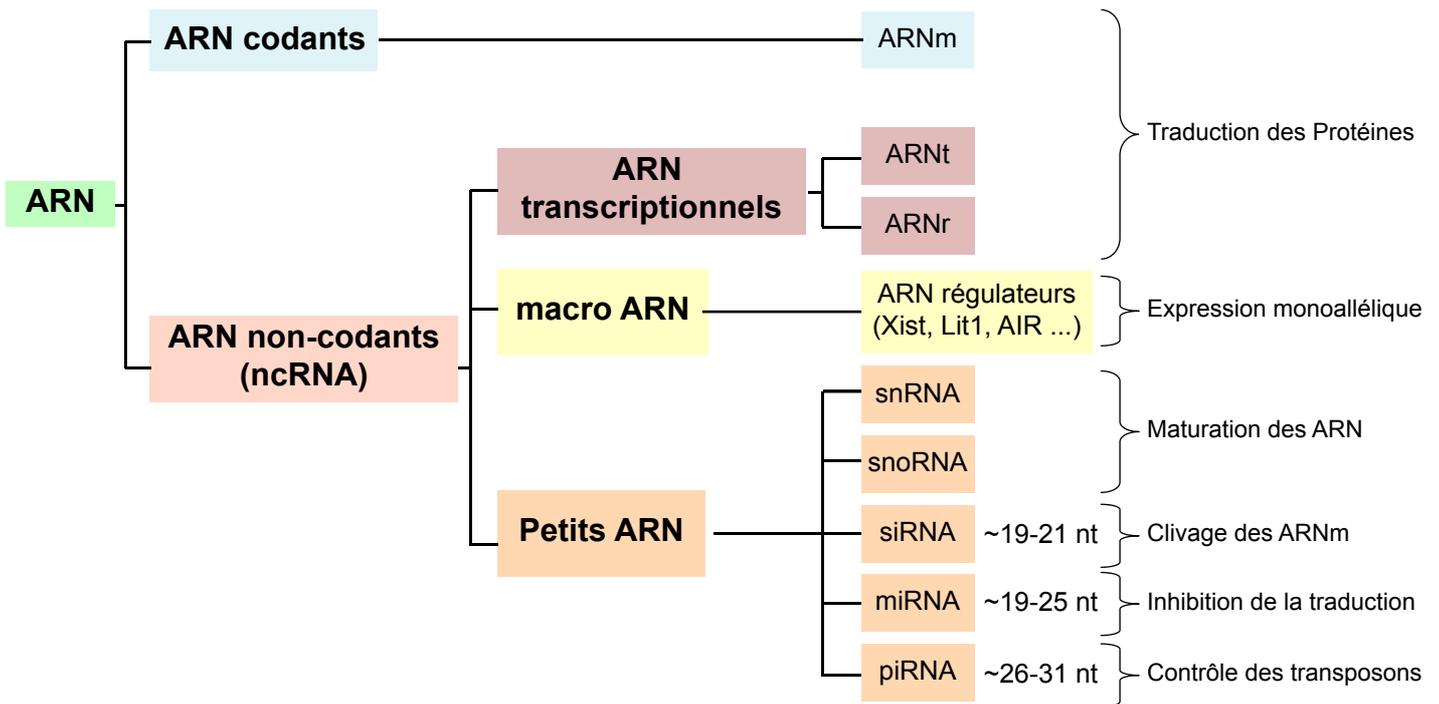
### **1-4-2) L'inactivation du chromosome X, un exemple de chronologie aboutissant à la répression transcriptionnelle.**

Chez les mammifères, la détermination du sexe se fait par les chromosomes sexuels, la femelle étant XX et le mâle XY. La compensation de dosage du double X chez la femelle contre un chez le mâle se fait par l'inactivation des gènes sur un des chromosomes de la paire de X. Ce processus a été postulé en 1961 par Mary Lyon.

Chez la souris, l'inactivation du X est aléatoire dans les cellules somatiques et soumise à l'Empreinte dans les annexes extra embryonnaires, le X paternel étant préférentiellement inactivé. Pour étudier les processus nécessaires à l'extinction des gènes du chromosome X, les cellules souches embryonnaires (ES) sont isolées à partir d'un blastocyste et la différenciation peut être induite en culture notamment à l'aide d'acide rétinoïque (Heard *et al.*, 1997). A partir de ce modèle *in vitro*, un programme d'événements a pu être mis en évidence (Figure 14).

Dans la cellule, un comptage du nombre de chromosomes a tout d'abord lieu de sorte à laisser un seul X actif par lot d'autosomes (Heard and Distèche, 2006). Ensuite, un des chromosomes est inactivé de manière aléatoire et forme une structure hétérochromatinienne compacte, appelée corpuscule de Barr. Un des premiers événements est l'expression d'un ARN non codant nucléaire de grande taille (15 kb) : *Xist* (*X-inactive specific transcript*). Cet ARN est exprimé à partir des deux allèles du chromosome puis s'accumule sur l'un des deux chromosomes X tandis qu'il s'éteint sur le second. *Xist* va recouvrir le chromosome à partir

duquel il est produit. Cet ARN est indispensable à l'initiation du processus d'inactivation qui est réversible à ce stade (Wutz and Jaenisch, 2000). Un autre évènement précoce est l'exclusion de la machinerie de transcription du territoire chromosomique du futur X inactif (Chaumeil *et al.*, 2006). *Xist* recruterait les complexes Polycomb PRC1 et PRC2 et ceux-ci permettraient la triméthylation de H3K27 et l'ubiquitination de H2A. Ces complexes pourraient jouer un rôle dans la mémoire cellulaire de l'X inactivé, indépendante de *Xist* (Kohlmaier *et al.*, 2004; Schoeftner *et al.*, 2006). D'autres marques d'hétérochromatine sont enrichies sur le chromosome X au cours de l'inactivation, comme H4K20me et H3K9me tandis que les histones sont déacétylés et la marque active H3K4me est effacée (Rougeulle *et al.*, 2004; Chaumeil *et al.*, 2006). Ces étapes ont lieu pendant une fenêtre de temps où l'inactivation est réversible et dépendante de *Xist*. Ensuite, l'inactivation est irréversible et indépendante de *Xist* et est transmise au cours des divisions cellulaires. Le chromosome X est enrichi en variant d'histone MacroH2A. La méthylation de l'ADN des promoteurs du chromosome X inactif est un évènement tardif et pourrait intervenir pour verrouiller l'état inactif des gènes, notamment via le recrutement de MBP (Heard *et al.*, 1997; Wutz and Jaenisch, 2000; Wutz and Gribnau, 2007).



**Figure 15 : Le monde des ARN.**

Les ARN non codants représentent une classe d'ARN variée de par leur taille et leur fonction dans la cellule. Les macros ARN régulateurs ont été impliqués dans l'inactivation du chromosome X et la régulation de l'Empreinte Parentale. Une classe de petits ARN incluant les small nuclear RNA (snRNA), les small nucléolar RNA (snoRNA) sont nécessaires à la maturation des ARN. L'interférence ARN utilise de petits ARN interférents (siRNA) pour réguler de façon post transcriptionnelle l'expression de gènes cibles. Les micro ARN (miRNA) sont plutôt responsables d'une répression traductionnelle. Et les ARN interagissant avec les protéines Piwi (piRNA) seraient responsables du contrôle des éléments transposables au cours de la gamétogenèse.

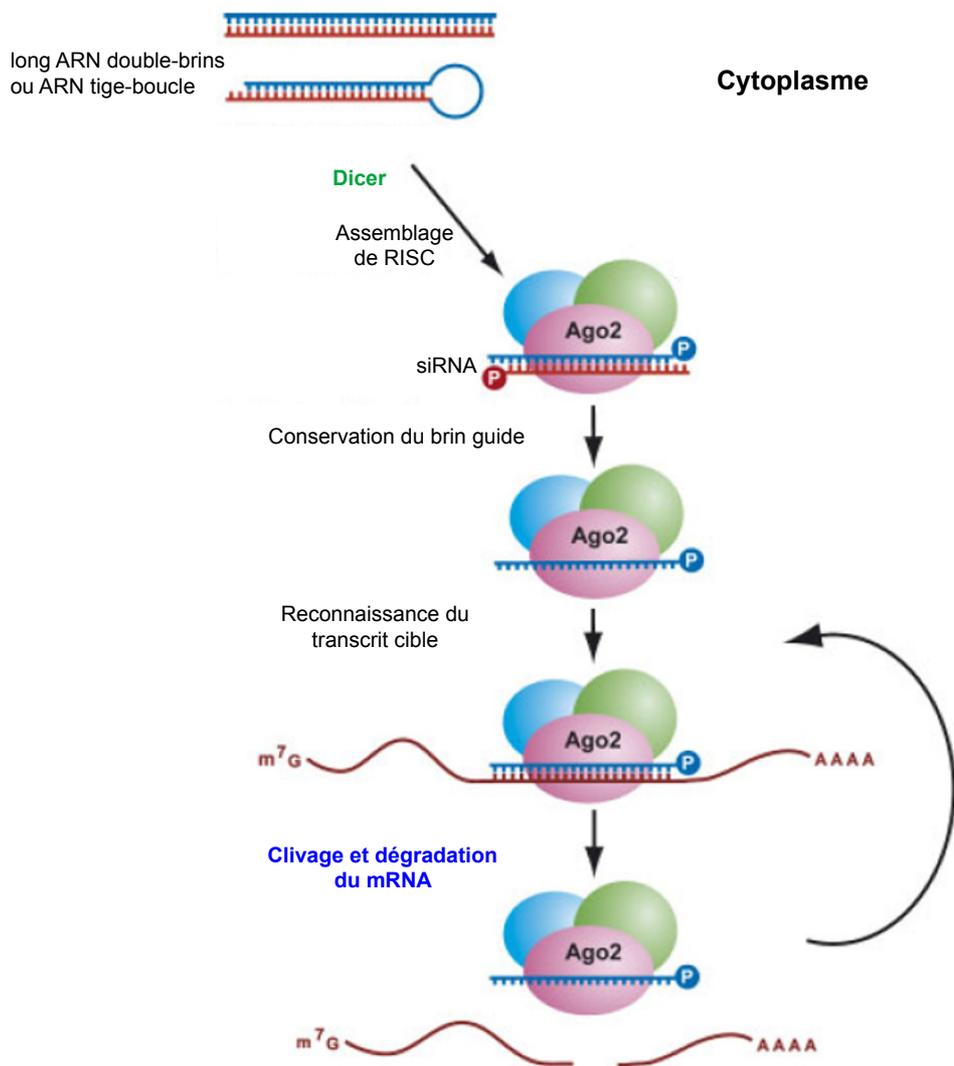
## 2) Les ARN non codants

### 2-1) Introduction

Une définition d'un ARN non codant (ARNnc) pourrait être : un ARN dont la fonction biologique n'est pas de servir de messenger pour la synthèse de protéine (Costa, 2007). Parmi les ARN non codants (Figure 15), les ARN impliqués dans la traduction sont connus depuis longtemps : **ARNr** et **ARNt** (r pour ribosomique et t pour transfert). D'autres ARNnc sont produits à partir des séquences parasites du génome tels les séquences SINE et Alu. Il existe de grands **ARN régulateurs** comme l'ARN *Xist* et les ARN soumis à l'Empreinte, que nous verrons dans la partie 3 de ce chapitre. D'autres ARN sont impliqués dans la maturation des ARN : les petits ARN nucléaires (**SnRNA** pour small nuclear RNA) impliqués dans l'épissage des messagers et les petits ARN nucléolaires (**snoRNA** pour small nucleolar RNA), localisés dans le nucléole comme leur nom l'indique et qui permettent la maturation des ARNr. *H19* a été le premier ARNnc décrit (autre que ARNr, ARNt et snoRNA) (Brannan *et al.*, 1990).

Il existe une autre catégorie d'ARN non codants découverte à partir de la fin des années 90: les petits ARN (Chu and Rana, 2007). Les premiers ont été les petits ARN interférants (ou **siRNA** pour small-interfering RNA), impliqués dans la dégradation des messagers par interférence ARN (Fire *et al.*, 1998). Ensuite les micros ARN (**miRNA** pour micro RNA) ont été décrits pour être impliqués dans la régulation traductionnelle. Enfin, récemment, les ARN interagissant avec les protéines Piwi (**piRNA** pour piwi-interacting RNA) impliqués dans la dégradation de l'ARN rétroviral ont été caractérisés. Les voies de régulation de ces petits ARN, qui seront détaillées dans les prochains paragraphes, ont en commun certaines protéines comme Dicer ou les protéines Argonautes.

La découverte des petits ARN est relativement récente, aussi les fonctions qu'ils assurent dans la physiologie de la cellule ne sont pas encore clairement définies. Les ARNnc semblent jouer des rôles variés dans diverses fonctions cellulaires comme la différenciation neuronale ou musculaire, l'apoptose et la survie cellulaire, la réponse au stress, l'activité transcriptionnelle, etc ... (Costa, 2007).



**Figure 16 : Modèle de mécanisme pour les siRNA.**

Des ARN en duplex double-brins ou en structure tige-boucles sont clivés par Dicer et intégrés au complexe RISC. Une reconnaissance parfaite de l'ARNm cible entraîne son clivage puis sa dégradation. (Adapté de Chu and Rana, 2007).

## 2-2) Empreinte Parentale et ARN non codants.

Le mécanisme d'Empreinte Parentale, qui sera décrit au prochain chapitre, fait intervenir de grands ARN régulateurs, comme *Lit1* (Long intronic transcript 1) ou *Air* (Antisense of *Igf2R*), qui permettent la répression transcriptionnelle en cis des gènes de leur domaines soumis à l'Empreinte (Cf p37).

Des centaines de petits ARNnc ont également été décrits dans les loci soumis à l'Empreinte. La plupart est localisée dans les loci *Dlk1* et *Snrpn* et leur expression est soumise à l'Empreinte (Royo *et al.*, 2006; Royo and Cavaille, 2008).

De nombreux snoRNA ont été décrits au locus *Snrpn* mais leurs cibles n'ont pas encore été découvertes (Cavaille *et al.*, 2000). Ils sont exprimés uniquement à partir de l'allèle paternel. Certains de ces ARN pourraient être impliqués dans le syndrome de Prader-Willi (Nicholls and Knepper, 2001). Un cluster de piRNA a également été caractérisé dans ce locus.

Au niveau du locus *Dlk1/Gtl2*, de nombreux ARNnc sont produits : *Gtl2*, qui est polyadénylé et épissé, deux ARNnc antisens, deux domaines de snoRNA et deux domaines de miRNA. Ils sont tous exprimés de façon monoallélique maternelle (Cavaille *et al.*, 2002; Seitz *et al.*, 2004). La dérégulation de l'expression des gènes de cette région est associée à la mutation Callipyge chez le mouton (Charlier *et al.*, 2001; Freking *et al.*, 2002).

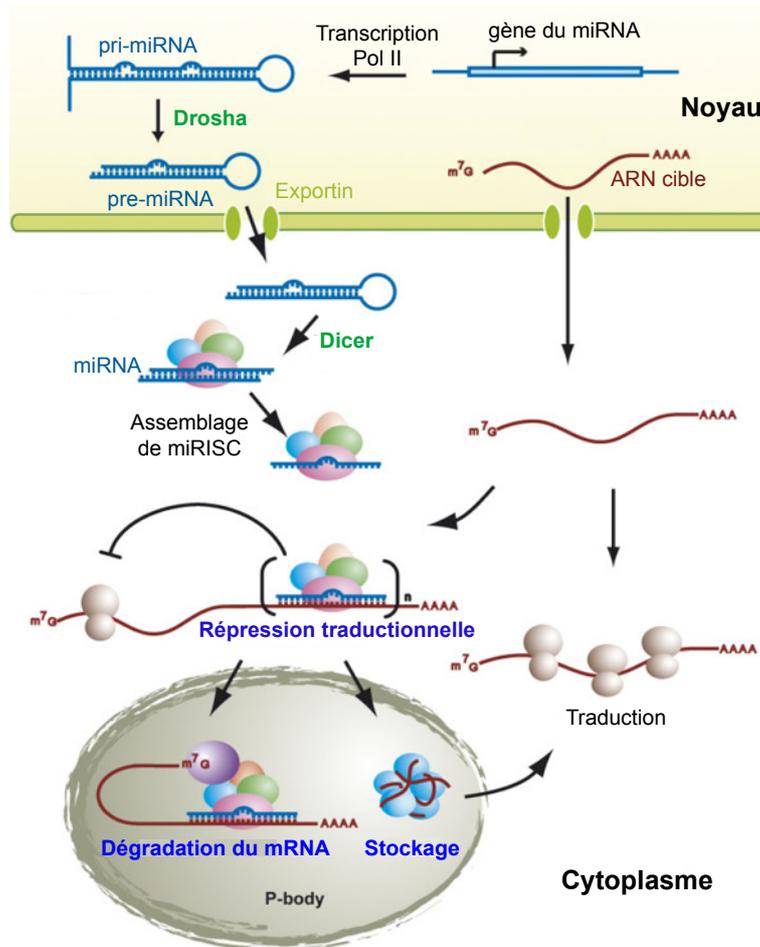
## 2-3) siRNA et interférence ARN

Les small-interfering RNA ou siRNA sont des petits ARN de 21 à 23 nucléotides impliqués dans l'interférence ARN. Ce mécanisme entraîne la dégradation des ARNm dans le cytoplasme (Figure 16).

Un ARN de longue taille est transcrit dans le noyau et forme des structures en épingle à cheveux (ou tige-boucle) ou des duplex double brins. Ces séquences sont reconnues et clivées par la protéine Dicer dans le cytoplasme de la cellule (Bernstein *et al.*, 2001). Les siRNA ainsi produits sont pris en charge par la protéine Argonaute<sup>3</sup> au niveau du complexe RISC (RNA-induced silencing complex). Seul un des deux brins dit "brin guide" du siRNA

---

<sup>3</sup> Il existe 8 protéines Argonautes chez l'Homme (AGO1 à 4 et PIWIL1 à 4). Au niveau des complexes RISC et miRISC chez les mammifères, les petits ARN peuvent s'associer à AGO1 à 4 mais seule la protéine Ago2 a une activité de clivage des ARN (Sasaki *et al.*, 2003; Meister *et al.*, 2004).



**Figure 17 : Modèle de mécanisme pour les miRNA.**

Les gènes des miRNA sont transcrits par la Pol II. Les pri-miRNA produits sont clivés par Drosha et exportés dans le cytoplasme. Les pre-miRNA sont alors clivés par Dicer et les miRNA sont intégrés au complexe miRISC. Ils peuvent alors exercer leurs fonctions : la répression traductionnelle mais aussi la dégradation des mRNA cibles et/ou le stockage dans les p-bodies.

(Adapté de Chu and Rana, 2007).

est conservé et peut s'hybrider avec un ARNm. L'appariement parfait avec le messenger cible entraîne son clivage puis sa dégradation par les exonucléases cellulaires (Meister and Tuschl, 2004; Rana, 2007). Ainsi, l'expression de l'ARNm cible est régulée de façon post-transcriptionnelle par le siRNA.

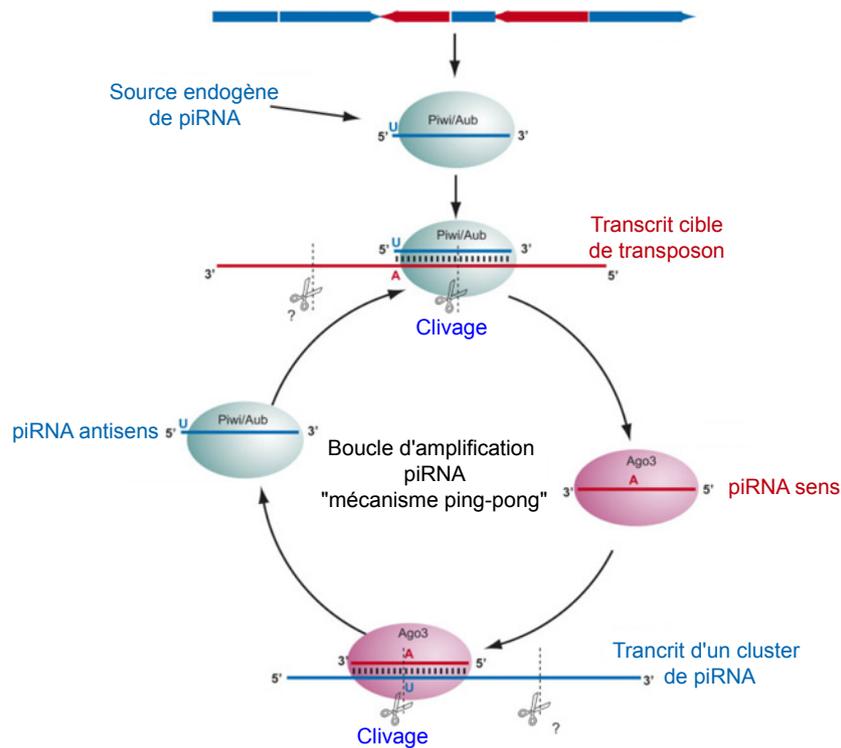
## 2-4) miRNA et répression traductionnelle

Les micros ARN ou miRNA sont des petits ARN d'environ 22 nucléotides (19 à 25 nt) impliqués dans la régulation de l'expression des gènes en inhibant la traduction (Figure 17).

Les miRNA peuvent être produits à partir de gènes codants ou non codants et à partir d'exons ou d'introns. Les pri-miRNA sont transcrits à partir de la polymérase II et sont coiffés et polyadénylés (Cai *et al.*, 2004; Lee *et al.*, 2004). Dans le noyau de la cellule, ces transcrits sont clivés par la protéine Drosha (Lee *et al.*, 2003). Cette étape permet l'obtention d'une structure en épingle à cheveux d'environ 70 nucléotides qui est exportée dans le cytoplasme. Ce pre-miRNA est reconnu et clivé par Dicer (Meister and Tuschl, 2004; Rana, 2007). Le miRNA peut alors être reconnu par le complexe RISC alors appelé miRISC (miRNA-induced silencing complex). De même que pour les siRNA, seul un brin est gardé par le complexe miRISC, le second brin ou "brin\*" (brin star) est éliminé (Khvorova *et al.*, 2003). Le miRNA est pris en charge par la protéine Argonaute et l'ARNm cible est reconnu mais avec un appariement imparfait. Une séquence de 2 à 8 nucléotides en 5' du miRNA guide, appelée séquence "seed", reconnaît spécifiquement l'ARN cible, permettant ainsi à un miRNA de reconnaître plusieurs cibles en général au niveau du 3'UTR.

La fixation du complexe miRISC au niveau de l'ARNm entraîne la répression de la traduction, la dégradation de l'ARNm ou sa séquestration dans des compartiments de stockage et dégradation appelés P-bodies (Meister and Tuschl, 2004; Chu and Rana, 2007; Rana, 2007).

Chez la Drosophile, les voies de dégradation post-transcriptionnelle des siRNA et d'inhibition traductionnelle des miRNA semblent bien distinctes, avec des protéines Dicer et Argonaute différentes. Chez les mammifères, la situation semble plus complexe mais les protéines Dicer et Argonaute des complexes RISC et miRISC sont les mêmes (Kawaji and Hayashizaki, 2008). Dans certaines conditions, les siRNA pourraient fonctionner comme des miRNA (Doench *et al.*, 2003).



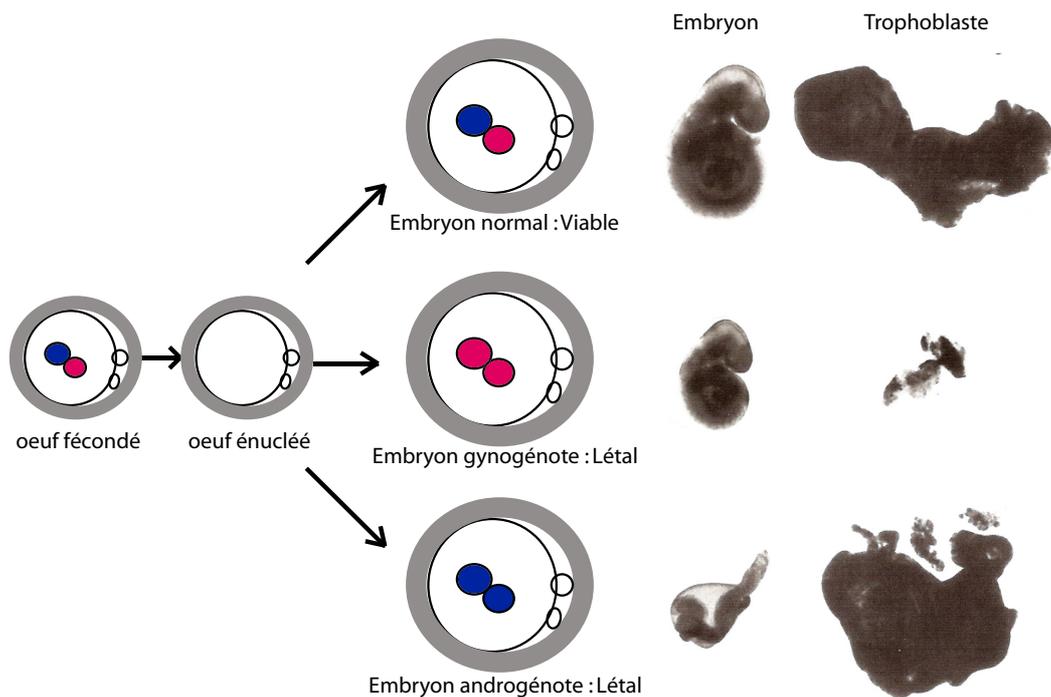
### Figure 18 : Modèle de mécanisme pour les piRNA.

Les piRNA sont pris en charge par les protéines Piwi (Piwi ou Aubergine). La reconnaissance de l'ARN de transposon entraîne son clivage. Une partie de cet ARN est conservée et prise en charge par la protéine Ago3 tandis que le reste est dégradé. Ago3 reconnaît, via son piRNA, l'ARN produit à partir d'un cluster piRNA, ce qui entraîne son clivage. Ce nouveau piRNA est pris en charge par Piwi et Aub et peut alors s'hybrider avec l'ARN de transposon. Ceci produit une boucle d'amplification. Ce mécanisme est appelé "ping-pong". (Adapté de Chu and Rana, 2007).

## 2-5) piRNAs et protéines Piwi

La protéine Piwi est une protéine de la famille Argonaute dont l'expression est restreinte à la lignée germinale chez la drosophile. Deux autres protéines Argonaute présentent cette caractéristique : Aubergine (Aub) et Ago3. Chez les mammifères, cette particularité est retrouvée pour les protéines homologues MIWI, MILI et MIWI2. Comme toutes les protéines Argonaute, les protéines Piwi s'associent à des petits ARN nommés piRNA (piwi-interacting RNAs) (Aravin *et al.*, 2006; Girard *et al.*, 2006).

Un mécanisme "ping-pong" a été proposé pour réprimer l'expression des transposons de façon post-transcriptionnelle au cours de la spermatogenèse (Figure 18) : Piwi ou Aub s'associent à un piRNA et reconnaissent un ARN de transposon. Cet ARN est clivé en deux endroits et le produit de clivage est pris en charge par Ago3. Ce complexe s'hybride avec le transcrite source de piRNA qui est à son tour clivé pour engendrer le piRNA reconnu par Piwi et Aub, ce qui produit une boucle d'amplification pour dégrader les transcrits des transposons. Un article récent montre que les piRNA ont également un rôle dans la méthylation des séquences géniques des transposons (Kuramochi-Miyagawa *et al.*, 2008).



**Figure 19 : Expériences de transfert nucléaire.**

A partir d'un oeuf qui vient d'être fécondé et énucléé, trois types d'embryons peuvent être reconstitués : des embryons normaux (pronucléus femelle rose et pronucléus mâle bleu), des gynogénotes (deux pronucléi femelles) et des androgénotes (deux pronucléi mâles). Les embryons gynogénotes et androgénotes ne sont pas viables et présentent des phénotypes différents : les gynogénotes ont des annexes extraembryonnaires atrophiées (trophoblaste) alors que les androgénotes présentent un retard de développement de l'embryon. Ces expériences ont montré la non équivalence des génomes parentaux, laissant supposer que chaque génome porte une Empreinte sexe spécifique. (Adapté de Miyoshi *et al.*, 2006).

### 3) L'Empreinte Génomique Parentale

#### 3-1) Historique et définition

La notion d'Empreinte Génomique Parentale est apparue au début des années 1980 suite à des expériences d'embryologie. Après la fécondation, les pronoyaux des embryons sont retirés et échangés. Des embryons gynogénotes (deux pronoyaux femelles) et androgénotes (deux pronoyaux mâles) sont ainsi reconstitués (Figure 19). Les embryons ayant reçu un pronoyau femelle et un mâle se développent jusqu'à terme alors que les embryons gynogénotes et androgénotes dégèrent au milieu de la gestation (McGrath and Solter, 1984; Surani *et al.*, 1984). De plus, les embryons n'ont pas le même phénotype selon la combinaison reçue : les gynogénotes peuvent se développer jusqu'au stade 25 somites mais les annexes extra embryonnaires sont atrophiées, les androgénotes ont un développement retardé et survivent jusqu'au stade 9 somites seulement (Barton *et al.*, 1984). Ces différences suggèrent que les génomes maternel et paternel ne sont pas équivalents ; ils portent donc une "empreinte" spécifique du sexe. Ainsi, chez les mammifères, la présence du génome maternel et paternel est indispensable au développement d'un nouvel individu viable.

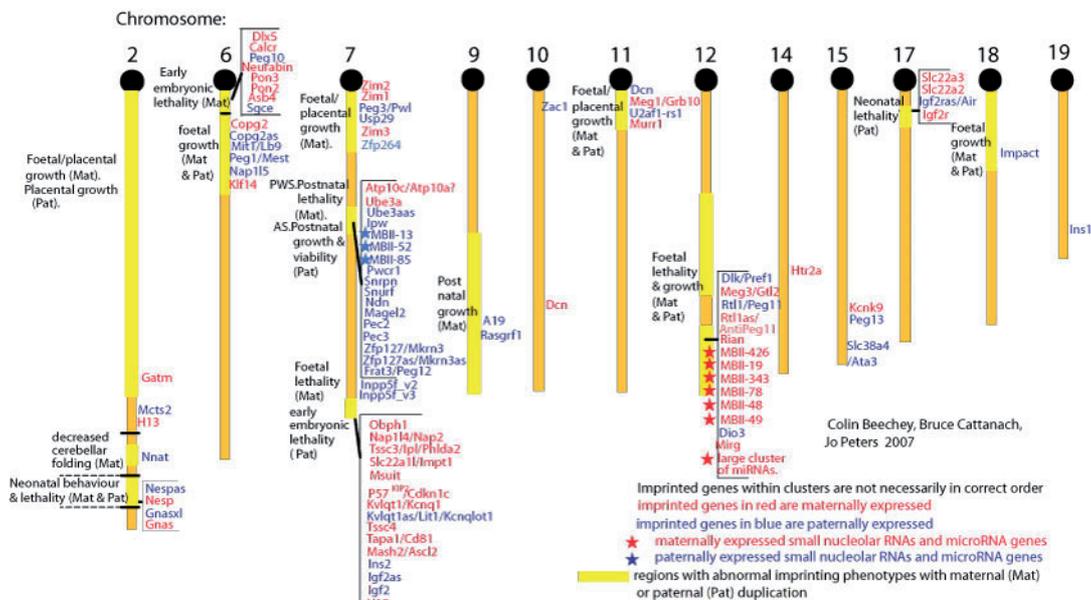
Par ailleurs, Cattanach a initié des travaux sur des lignées murines portant des translocations réciproques qui permettent ainsi de transmettre une région chromosomique en deux copies maternelles ou deux copies paternelles. Un certain nombre de ces régions induisent des phénotypes caractéristiques d'un disomie uniparentale (maternelle ou paternelle) (Cattanach and Kirk, 1985). Ainsi, une carte du génome de la souris a pu être établie indiquant pour chaque chromosome les régions à effet maternel ou à effet paternel (Figure 20) (Cattanach *et al.*, 2006).

En 1991, trois premiers gènes ont été décrits comme soumis à l'Empreinte : *Igf2*, *Igf2R* et *H19* (DeChiara *et al.*, 1991; Barlow *et al.*, 1991; Bartolomei *et al.*, 1991). Dans le cas du gène *Igf2*, la transmission du phénotype dans la lignée invalidée était surprenante. Les souris *Igf2*<sup>-/-</sup> présentent une réduction de taille et de poids de 40% par rapport aux animaux sauvages (DeChiara *et al.*, 1990). Les animaux *Igf2*<sup>+/-</sup><sup>4</sup> qui héritent de la délétion de leur père (hétérozygotes paternels) ont le même phénotype de défaut de croissance que les

---

<sup>4</sup> Dans le cas des gènes soumis à l'Empreinte, le génotype des mutants est indiqué ainsi : gène<sup>allèle maternel/allèle paternel</sup>. Ainsi, les animaux *Igf2*<sup>+/-</sup> portent la délétion sur l'allèle paternel.

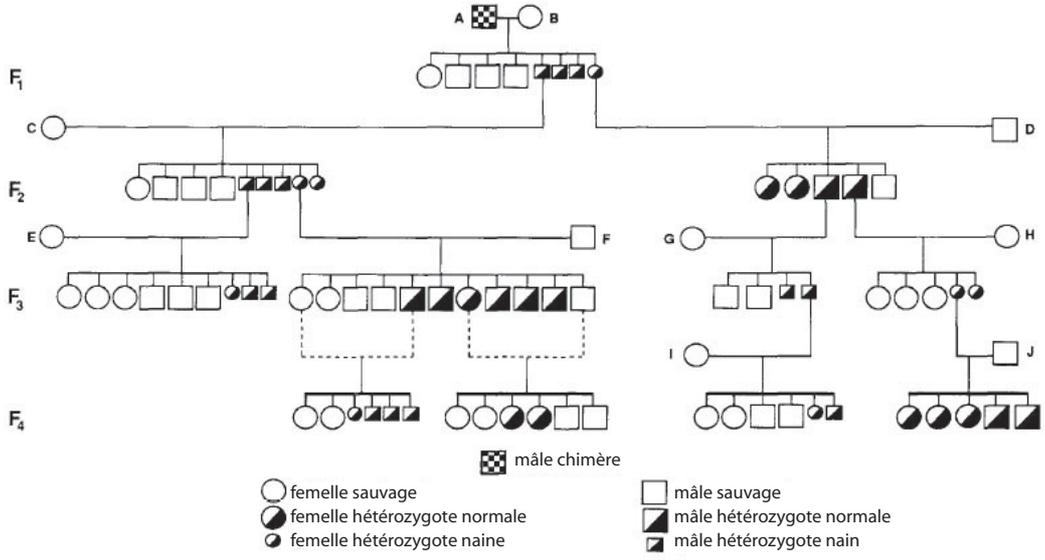
Mouse imprinted genes, regions and phenotypes



**Figure 20 : Carte des gènes soumis à l'Empreinte Parentale chez la souris.**

Cette carte représente les chromosomes murins (orange) présentant des régions influençant le phénotype en fonction de la transmission maternelle ou paternelle (en jaune). Les gènes soumis à l'Empreinte appartiennent presque tous à ces régions et sont pour la plupart regroupés en clusters. Ils sont d'expression monoallélique maternelle (rouge) ou paternelle (bleu). Les gènes codant des petits ARN non codants sont notés par des étoiles.

Source : [http://www.har.mrc.ac.uk/research/genomic\\_imprinting/maps.html](http://www.har.mrc.ac.uk/research/genomic_imprinting/maps.html)



**Figure 21 : Transmission de la mutation et du phénotype des souris KO *Igf2*.**

Le phénotype des animaux *Igf2*<sup>-/-</sup> est une taille réduite (petit rond ou petit carré) comparé aux sauvages (grand rond ou grand carré). La transmission à l'état hétérozygote de l'inactivation (en noir) se fait selon un arbre d'hérédité non mendélienne : la transmission par la mère (hétérozygotes maternels) entraîne un phénotype sauvage alors que la transmission par le père (hétérozygotes paternels) entraîne un phénotype de défaut de croissance. L'explication vient du fait que le gène *Igf2* est d'expression monoallélique paternelle. (DeChiara *et al.*, 1991).

homozygotes  $Igf2^{-/-}$ . En revanche les souris hétérozygotes maternelles  $Igf2^{-/+}$  ne présentent pas de défaut de croissance et sont tout à fait semblables aux souris sauvages (Figure 21). Une analyse de l'expression du gène a montré qu' $Igf2$  s'exprime uniquement à partir de l'allèle paternel. Ainsi, une transmission paternelle de la délétion entraîne une répression totale du gène alors qu'une transmission maternelle ne modifie pas l'expression du gène  $Igf2$  (DeChiara *et al.*, 1991).

Ainsi, l'Empreinte Parentale se définit comme un **mécanisme épigénétique qui conduit à l'expression monoallélique d'un gène en fonction de l'origine parentale de l'allèle**. Bien que les deux allèles soient identiques en terme de séquence, ils sont différents en terme d'expression : l'un est exprimé et l'autre réprimé en fonction de leur héritage maternel ou paternel. La notion d'Empreinte maternelle ou d'Empreinte paternelle est parfois utilisée. Il est alors considéré que l'allèle portant l'Empreinte est réprimé. Cependant, comme nous le verrons par la suite, la discussion sur l'origine parentale de l'Empreinte n'est pas close. En effet, pour certains loci, comme celui du gène  $H19$ , il apparaît difficile de savoir si l'Empreinte est apposé par le père sous forme de méthylation ou par la mère sous forme de protection contre la méthylation, ou par les deux mécanismes. Aussi, il me semble préférable de parler de gène d'expression maternelle et de gène d'expression paternelle (ou gène maternel et gène paternel).

A l'heure actuelle, près d'une centaine de gènes ont été identifiés comme étant soumis à l'Empreinte Parentale ([http://www.har.mrc.ac.uk/research/genomic\\_imprinting/maps.html](http://www.har.mrc.ac.uk/research/genomic_imprinting/maps.html)). Aucun biais n'est observé concernant l'origine de l'allèle exprimé puisqu'il y presque autant de gènes à expression maternelle que paternelle (Figure 20). Des approches *in silico* basées sur les comparaisons de séquences des gènes connus ont par ailleurs permis de prédire un total de 600 gènes soumis à l'Empreinte chez la souris (Luedi *et al.*, 2005). Ce nombre représente 2,5% des gènes annotés, ce qui indique que les gènes soumis à l'Empreinte forment une famille restreinte dans le génome.

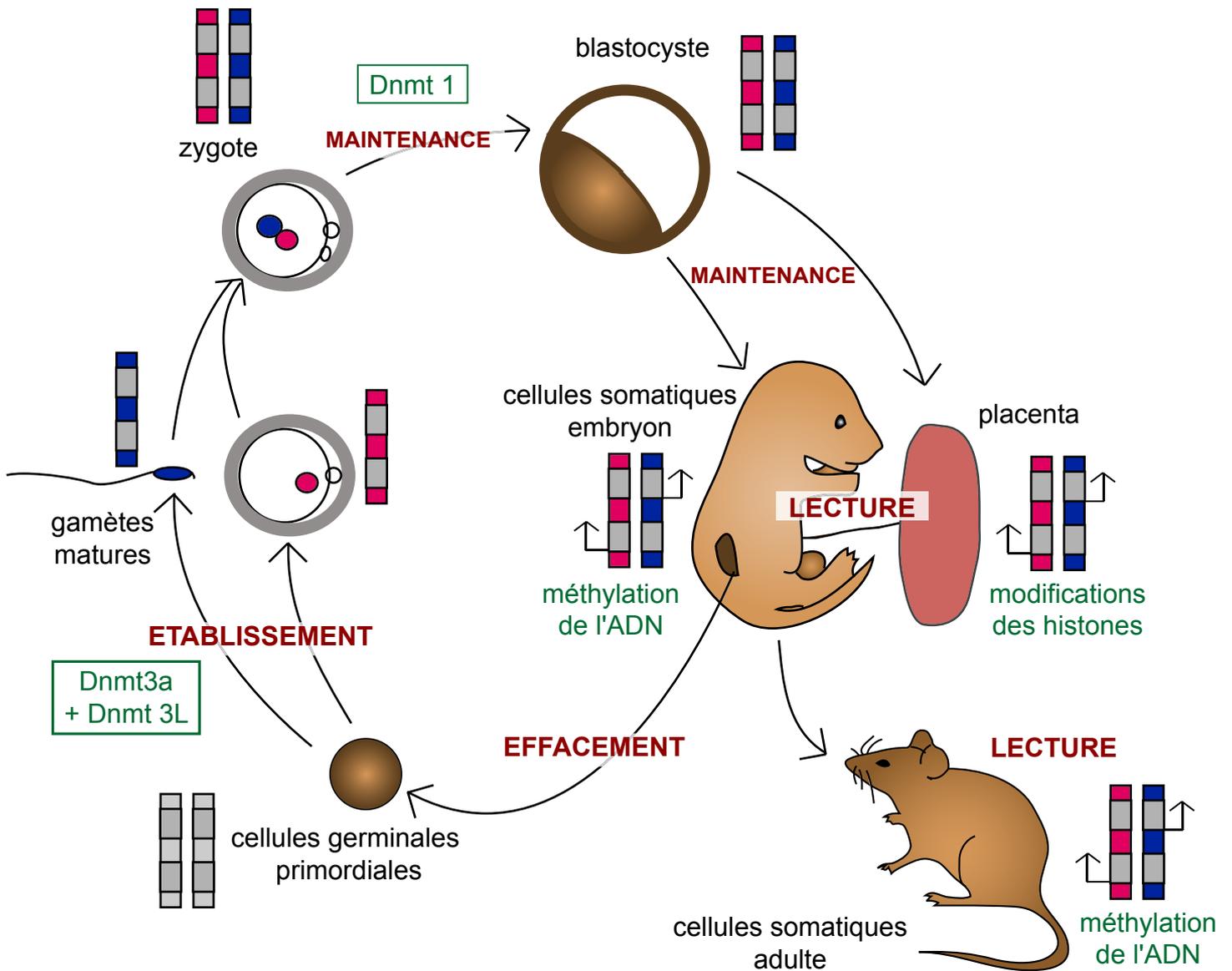
### 3-2) Caractéristiques

Les gènes soumis à l'Empreinte ne présentent pas de spécificité particulière commune en terme de séquence. En revanche, ils partagent un certain nombre de caractéristiques au niveau de leur localisation chromosomique et au niveau épigénétique.

Tout d'abord, les gènes soumis à l'Empreinte sont **regroupés dans de larges domaines au niveau du génome** (ou clusters) (Figure 20). Il est rare qu'un gène soumis à l'Empreinte soit isolé. Le plus souvent les gènes appartiennent à un locus dans lequel certains sont d'expression maternelle et d'autres d'expression paternelle. Les clusters de gènes soumis à l'Empreinte contiennent en général au moins un **ARNnc**, exprimé à partir de l'allèle opposé de celui des gènes codant des protéines (Cattanach and Kirk, 1985; Thorvaldsen and Bartolomei, 2007).

De plus, ces gènes présentent des séquences différenciellement méthylées en fonction de l'allèle paternel ou maternel, appelées **DMR** (Differentially Methylated Regions) (Bartolomei *et al.*, 1993). Elles sont de deux types : les DMR gamétiques sont méthylées au cours de la gamétogenèse et les DMR somatiques sont méthylées au cours de l'embryogenèse et de la vie post-natale. Par exemple, en ce qui concerne le locus *H19/Igf2*, la DMR *H19* est gamétique et les DMR du gène *Igf2* sont somatiques chez la souris. Il existe d'ailleurs une hiérarchie, la DMR *H19* étant nécessaire à la mise en place de la méthylation différentielle des DMR du gène *Igf2* (Lopes *et al.*, 2003).

En plus des DMR, **les marques des histones semblent être différentielles** sur les allèles des gènes soumis à l'Empreinte. Ainsi, ces modifications différentielles ont été observées au locus *H19/Igf2*, comme nous le verrons par la suite (Han *et al.*, 2008). Par ailleurs, les expériences de ChIP-Seq ont permis de les mettre en évidence. Les gènes soumis à l'Empreinte semblaient porter des marques bivalentes H3K4me3/H3K9me3 dans les cellules ES. En fait, l'analyse de cellules ES, produites à partir d'embryons issus de croisement *domesticus/castaneus*, a montré que le gène *Peg3*, d'expression paternelle, porte spécifiquement la marque active H3K4me3 sur l'allèle paternel *castaneus* et la marque répressive H3K9me3 sur l'allèle *domesticus*. Par ailleurs la marque H3K36me3, associée à la transcription, est présente uniquement sur l'allèle exprimé des microARN du locus *Dkl1* et des gènes *Snrpn*, *Grb10* et *Peg3* notamment (Mikkelsen *et al.*, 2007).



**Figure 22 : Le cycle de l'Empreinte parentale.**

L'Empreinte est apposée au cours de la gamétogenèse, alors que les deux génomes parentaux sont séparés, grâce à l'action combinée de Dnmt3a et Dnmt3L. Après la fécondation, l'Empreinte est maintenue par Dnmt1 alors que le reste des génomes parentaux est complètement déméthylé. L'Empreinte est lue dans les cellules somatiques de l'embryon et de l'individu en particulier *via* la méthylation de l'ADN. Dans le placenta, la lecture serait plutôt assurée par les modifications de la chromatine. L'Empreinte est effacée dans les cellules germinales primordiales afin de réapposer une nouvelle Empreinte spécifique du sexe de l'individu pour la transmettre à sa descendance.

Enfin, les loci soumis à l'Empreinte comportent un **centre d'Empreinte (ICR pour Imprinting Control Region<sup>5</sup>)**. Ces séquences, qui peuvent s'étendre sur plusieurs kilobases et sont riches en CpG, régulent l'expression monoallélique de l'ensemble des gènes du locus. Lorsqu'elles sont délétées, les gènes sont exprimés de façon biallélique ou totalement réprimés. Tous les centres d'Empreintes connus sont des DMR gamétiques (Wan and Bartolomei, 2008).

### 3-3) Le cycle de l'Empreinte parentale

L'Empreinte Parentale présente dans un individu et reçue par les parents doit être réinitialisée à chaque génération afin que les gamètes transmettent une Empreinte caractéristique du sexe d'un individu. Cette "remise à zéro" se fait au moment où les génomes femelles et mâles sont séparés, c'est-à-dire au cours de la gamétogenèse (Figure 22). La première marque associée à la régulation de l'expression des gènes soumis à l'Empreinte a été la méthylation de l'ADN ; l'étude des modifications des histones n'était alors pas aussi répandue qu'à l'heure actuelle. La méthylation des DMR de ces gènes a donc particulièrement été étudiée au cours de ce processus. Dans l'embryon au stade E11.5, le génome des cellules germinales primordiales (PGC) subit une vague de déméthylation globale. L'Empreinte est alors effacée. En effet une étude montre que les DMR des gènes *Igf2R*, *H19*, *Igf2* et *Snrpn* sont déméthylées et que ces gènes s'expriment de façon biallélique dans les PGC (Szabo and Mann, 1995a).

Les **DMR gamétiques** sont ensuite spécifiquement méthylées soit uniquement au cours de la spermatogenèse, comme c'est le cas de la DMR *H19*, soit uniquement au cours de la maturation ovocytaire, comme c'est le cas de la KvDMR du locus du gène *KvLQT1*. **Dnmt3a** et **Dnmt3L** sont indispensables à la méthylation des DMR gamétiques. Dans les ovocytes des souris *Dnmt3L<sup>-/-</sup>*, les DMR à méthylation maternelle sont hypométhylées alors que les DMR à méthylation paternelle sont correctement déméthylées (Bourc'his *et al.*, 2001; Hata *et al.*, 2002). En miroir, dans les cellules germinales mâles, les DMR à méthylation paternelle sont hypométhylées alors que les DMR à méthylation maternelle sont correctement déméthylées (Webster *et al.*, 2005). Ainsi, la protéine Dnmt3L, associée à l'ADN

---

<sup>5</sup> Les dénominations IC, pour Imprinting Center, ou ICE, pour Imprinting Control Element, peuvent parfois être utilisées.

méthyltransférase de novo Dnmt3a, joue un rôle crucial dans l'établissement de la méthylation et donc du bon fonctionnement des gènes soumis à l'Empreinte Parentale.

Immédiatement après la fécondation, le génome du zygote subit une vague de déméthylation, à laquelle les DMR des gènes soumis à l'Empreinte échappent (Figure 22). Des études récentes ont confirmé le rôle supposé de **Dnmt1** dans le maintien de la méthylation des DMR à ce stade (Cirio *et al.*, 2008; Hirasawa *et al.*, 2008). Dnmt1 est cytoplasmique dans le zygote au stade 1 cellule puis cytoplasmique et nucléaire dès le stade 2 cellules. L'embryon lui-même n'exprime *Dnmt1* qu'à partir de la fin du stade 2 cellules, la protéine du zygote provenant du stock ovocytaire (Dnmt1o) (Cirio *et al.*, 2008). Par ailleurs, la délétion conditionnelle ovocytaire et zygotique de Dnmt3a et 3b entraîne peu de modification de la méthylation différentielle des gènes soumis à l'Empreinte. En revanche, la perte zygotique de Dnmt1 entraîne une déméthylation partielle et la perte ovocytaire et zygotique entraîne une déméthylation totale des DMRs. Ces résultats indiquent que **Dnmt1 est nécessaire et suffisante au maintien de la méthylation des gènes soumis à l'Empreinte** dans l'embryon préimplantatoire (Hirasawa *et al.*, 2008). D'autre part, il a été montré que la protéine Stella est importante pour ce maintien, les embryons délétés pour *Stella* perdant la méthylation d'au moins cinq DMR, dont celle de l'ICR du locus *H19/Igf2* (Nakamura *et al.*, 2007).

Ensuite l'Empreinte, notamment grâce à ces marques de méthylation est lue dans les cellules somatiques de l'embryon puis de l'adulte et cela conduit à l'expression monoallélique des gènes.

Dans plusieurs loci, des gènes soumis à l'Empreinte uniquement dans le placenta ont été identifiés. Ces gènes sont tous exprimés à partir de l'allèle maternel (Wagschal and Feil, 2006). Différentes études ont montré que la lecture de l'Empreinte dans le placenta n'est pas due à la méthylation de l'ADN. Dans le cas du locus de *KvLQT1*, la méthylation des histones, et non celle de l'ADN, est impliquée dans la répression monoallélique des gènes : les marques répressives H3K27me3, H3K9me2, H3ac0 et H4K4me0 sont observées sur l'allèle paternel tandis que les marques activatrices H3K4me2 et l'acétylation marquent l'allèle maternel. L'établissement de ces marques est toutefois dépendant de l'ICR. Il a également été montré pour ce locus et pour d'autres que l'expression monoallélique nécessite des protéines du groupe Polycomb (Lewis *et al.*, 2004; Umlauf *et al.*, 2004; Monk *et al.*, 2008).

Localisation Souris/Homme	Locus	ARNnc	Autres gènes		DMR gamétique	Mécanisme de régulation
2H4/20q13	<i>Gnas</i>	<i>Nespas, Gnas ex1A</i>	<i>Nesp, GnasXL</i>	Promoteurs <i>GnasXL</i> & <i>Nespas</i> (1) Promoteurs <i>Gnas ex1a</i> (2)		?
6A1/7q21.3	<i>Peg10</i>		<i>Ppp1r9a, Asb4, Sgce</i>	Promoteur de <i>Peg10</i> et de <i>Sgce</i>		?
6A3/7q32.2	<i>Mest (Peg1)</i>	<i>Copg2as</i>	<i>Copg2, Kfl14</i>	Promoteur & exon1 de <i>Mest</i>		?
7A1/19q13	<i>Peg3</i>		<i>Zim1, Zim2, Zim3, Usp29, Zfp264</i>	Promoteur & exon1 de <i>Peg3</i>		?
7B5/15q11.2	<i>Snrpn</i>	<i>Ube3a-AS snoRNA</i>	<i>ATP10a, Ube3A, Snurf, Ndn, Magel2, Mkrn3, Peg12</i>	PWS-IC (ICR)		ARN non codant ?
7F5/11p15.4	<i>Kcnq1</i>	<i>Kcnq1ot1</i>	<i>Cdkn1c, Osbp15, Tnfrsf23, Nap114, Tssc3, Slc22a1, Tssc4, Cd81, Ascl2</i>	KvDMR1 (ICR)		ARN non codant
7F5/11p15.5	<i>Igf2</i>	<i>H19</i>	<i>Ins2</i>	DMR H19 (ICR)		Insulateur
9E3/15q25.1	<i>Rasgrf1</i>	4930524O08Rik		-30kb DMR <i>Rasgrf1</i> (ICR)		Insulateur ?
10A2/6q24.2	<i>Zac1 (Plagl1)</i>	<i>Hymai</i>		<i>Hymai</i> exon1		?
11A1/7p12.2	<i>Grb10</i>		<i>Grb10 brain isoform</i>	<i>Grb10</i> CpG island 2		?
11A3/5p22.2	<i>Zrsr1 (U2af1-rs1)</i>		<i>Commd1</i>	<i>Zrsr1</i> CpG island		?
12F1/14q32.2	<i>Dlk1</i>	<i>Gtl2, antiRtl1 miRNA, Rian snoRNA, Mirg miRNA</i>	<i>Rtl1, Dio3</i>	IG-DMR (ICR)		?
17A1	<i>Igf2R</i>	<i>AIR</i>	<i>Slc22a2, Slc22a3</i>	Promoteur de <i>AIR</i> (ICR)		ARN non codant

**Tableau 2 : Les loci soumis à l'Empreinte Parentale.**

Les gènes soumis à l'Empreinte sont souvent regroupés en clusters et synténiques entre la souris et l'Homme. Les principaux clusters sont représentés dans ce tableau, le gène principal rappelle le nom du locus, les ARN non codants sont indiqués ainsi que les autres gènes codant des protéines (expression des gènes : rose = maternelle, vert = paternelle). Une majorité des DMR gamétiques est méthylée au cours de l'ovogenèse (méthylation des DMR : orange = maternelle, bleu = paternelle). Deux mécanismes de régulation sont caractérisés : ARNnc de grande taille (rouge) ou insulation (jaune), mais pour la plupart des loci, il est encore inconnu (?). (Adapté de Thorvaldsen & Bartolomei, 2007)

### 3-4) Régulation de l'expression des gènes

Les gènes soumis à l'Empreinte sont en général regroupés en clusters sur les chromosomes. L'ICR est capable de réguler l'ensemble des gènes du locus. Une dizaine de loci sont actuellement recensés (Tableau 2) et deux grands mécanismes de régulation se dégagent : la régulation par les insulateurs et la régulation par de grands ARN non codants (Thorvaldsen and Bartolomei, 2007; Wan and Bartolomei, 2008).

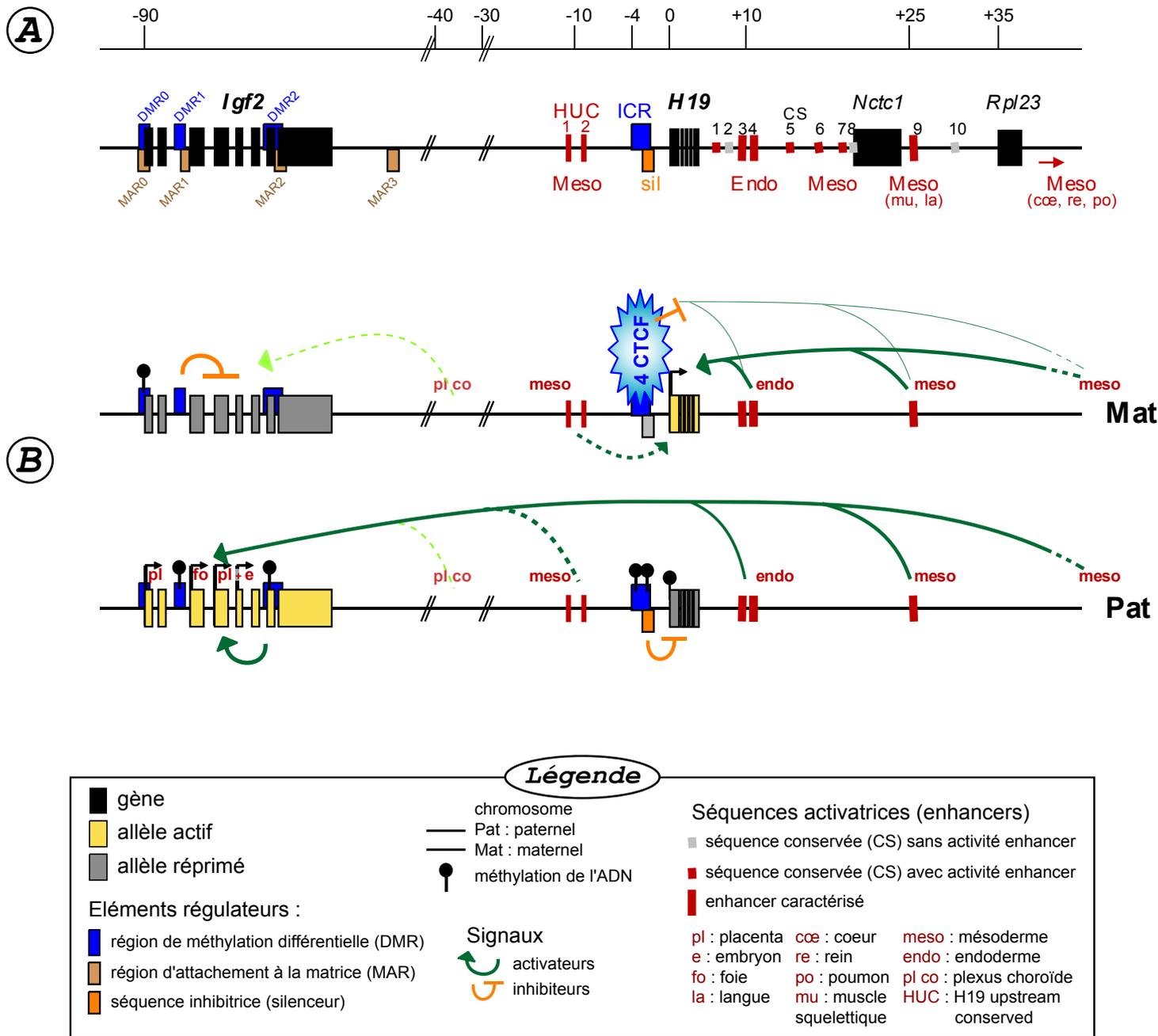
#### 3-4-1) Régulation par les insulateurs

Un insulateur est une séquence d'ADN capable d'isoler les promoteurs des séquences activatrices (enhancers) afin de réprimer l'expression d'un gène. La régulation de l'expression monoallélique des gènes par les insulateurs a été décrite pour deux loci soumis à l'Empreinte (Tableau 2) : *H19/Igf2*, détaillé ci-après, et *Rasgrf1* (Yoon *et al.*, 2005; Wan and Bartolomei, 2008). Cette insulation est médiée par la protéine CTCF.

##### a) La protéine CTCF

La protéine CTCF (CCCTC-binding Factor) a été à l'origine décrite comme un **facteur de transcription** de l'oncogène *Myc*. CTCF est en effet capable de lier l'ADN via son domaine à 11 doigts de zinc (ZnF pour Zinc Fingers). Ce domaine peut lier différentes séquences d'ADN en fonction de la combinaison de doigts de zinc utilisée (Filippova, 2008). En revanche, CTCF est sensible à la méthylation qui empêche la liaison aux séquences cibles (Bell and Felsenfeld, 2000; Hark *et al.*, 2000).

CTCF est non seulement un facteur de transcription mais une **protéine à activité insulatrice**, caractérisée pour la première fois au niveau du locus  $\beta$ -globin (Bell *et al.*, 1999). Lorsque la séquence de liaison de CTCF est située entre une séquence régulatrice (activatrice : enhancer ou inhibitrice : silencer) et le promoteur d'un gène, CTCF peut bloquer les interactions entre les séquences pour modifier l'expression du gène (Filippova, 2008). L'activité insulatrice de CTCF a été impliquée dans différents processus tels que la compaction en hétérochromatine des séquences répétées du génome, l'inactivation du chromosome X et l'Empreinte Parentale, comme nous allons le voir en b) (Filippova, 2008).



**Figure 23 : Le locus *H19/Igf2*.**

**A.** Eléments de contrôle de l'expression des gènes *H19* et *Igf2*.

**B.** Modèle de régulation de l'expression monoallélique des gènes. En se fixant sur l'ICR non méthylé sur l'allèle maternel, la protéine CTCF inhibe les signaux des enhancers qui ne peuvent activer les promoteurs maternels du gène *Igf2*. CTCF ne peut pas se fixer sur l'ICR paternel méthylé et les enhancers peuvent activer l'expression d'*Igf2*. La séquence silencer et la méthylation du promoteur inhibent l'expression paternelle du gène *H19*.

CTCF joue également un **rôle dans l'établissement et le maintien des marques d'Empreinte**. En effet, la liaison de CTCF à l'ICR du locus *H19/Igf2* permettrait de le préserver de la méthylation au cours de l'ovogenèse (Fedoriw *et al.*, 2004) et de conserver la méthylation différentielle après la fécondation (Schoenherr *et al.*, 2003; Pant *et al.*, 2004; Szabo *et al.*, 2004; Engel *et al.*, 2006). Il faut toutefois noter qu'après mutations des sites CTCF, le profil de méthylation de l'ICR est conservé au cours de la gamétogenèse alors que les siRNA dirigés contre CTCF entraînent une méthylation pendant l'ovogenèse. Ce point reste donc encore à éclaircir. De plus, les DMR sont organisées de manière hiérarchique, l'ICR étant nécessaire à la méthylation des DMR du gène *Igf2* (Lopes *et al.*, 2003). Cette hiérarchie pourrait être due à l'architecture de la chromatine dont CTCF pourrait être une des pièces principales comme il sera évoqué plus loin.

Une protéine paralogue CTCF-like (CTCFL ou BORIS pour Brother of the regulator of imprinted sites) est exprimé dans la lignée germinale mâle au moment de l'établissement de l'Empreinte. CTCFL pourrait jouer un rôle dans l'établissement de la méthylation de l'ICR sur le futur allèle paternel (Loukinov *et al.*, 2002; Jelinic *et al.*, 2006).

### **b) Régulation de l'expression des gènes au locus *H19/Igf2***

De nombreux modèles murins ont été développés pour mieux comprendre le fonctionnement des gènes soumis à l'Empreinte, et en particulier au locus *H19/Igf2* (Arney, 2003). De ces études, un modèle "canonique" de la régulation de l'expression de ces gènes a été établi chez la souris (Figure 23).

Ce locus est situé sur la partie distale du chromosome 7 chez la souris et sur la région synténique 11p15.5 chez l'Homme (Zemel *et al.*, 1992). Il comprend les gènes *H19*, d'expression maternelle, *Igf2*, exprimé à partir de l'allèle paternel, et *ins2*, soumis à l'Empreinte uniquement dans le sac vitellin et exprimé paternellement (Bartolomei *et al.*, 1991; DeChiara *et al.*, 1991; Zhang and Tycko, 1992; Giannoukakis *et al.*, 1993; Giddings *et al.*, 1994; Deltour *et al.*, 1995; Moore *et al.*, 2001).

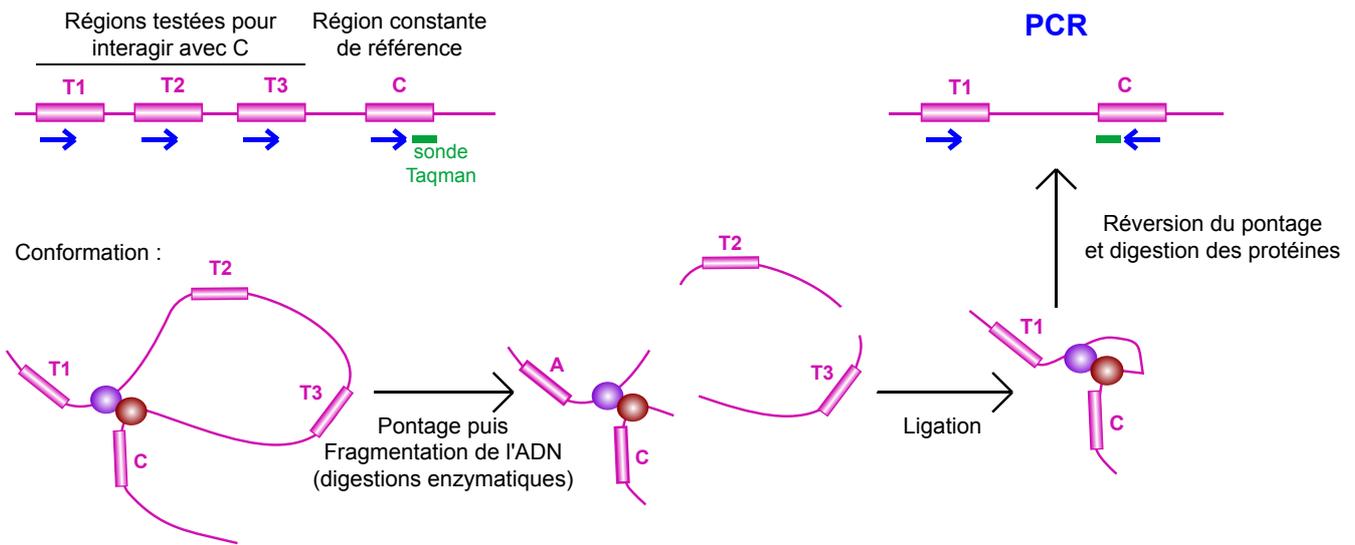
Les gènes *H19* et *Igf2* partagent un certain nombre de séquences régulatrices et notamment une série d'enhancers à spécificité tissulaire (Figure 23). Tout d'abord, deux **enhancers endodermiques** ont été identifiés 10 kb en aval du gène *H19*. La délétion de ces séquences entraîne la perte d'expression de l'ARN *H19* dans les tissus endodermiques lorsqu'elle est transmise maternellement et l'absence d'expression de l'ARNm *Igf2* après

transmission paternelle, ce qui implique que ces enhancers sont communs aux deux gènes (Yoo-Warren *et al.*, 1988; Leighton *et al.*, 1995).

Dans le but de trouver les **enhancers mésodermiques**, dix séquences conservées (CS) ont été caractérisées (Ishihara *et al.*, 2000). De plus, un transgène YAC de 130 kb, s'étendant jusqu'à +35 kb en aval du début de transcription du gène *H19*, permet l'expression appropriée de l'ARN *H19* dans les tissus endodermiques et mésodermiques à l'exception du cœur, des poumons et des reins (Ainscough *et al.*, 1997; Ainscough *et al.*, 2000). Une autre étude a montré que les enhancers mésodermiques musculaires et cardiaques sont contenus dans un transgène BAC s'étendant jusqu'à +120 kb (Kaffer *et al.*, 2000). Enfin, le mutant *minute* a une perte d'expression du gène *Igf2* due à une inversion dont le point de cassure se situe au niveau de la séquence CS9 (Davies *et al.*, 2002). L'ensemble de ces données permet de proposer la séquence CS9 comme enhancer mésodermique spécifique du muscle squelettique et de la langue. Un ou plusieurs autres enhancers spécifiques du cœur, des poumons et des reins se trouvent entre +35 et +120 kb.

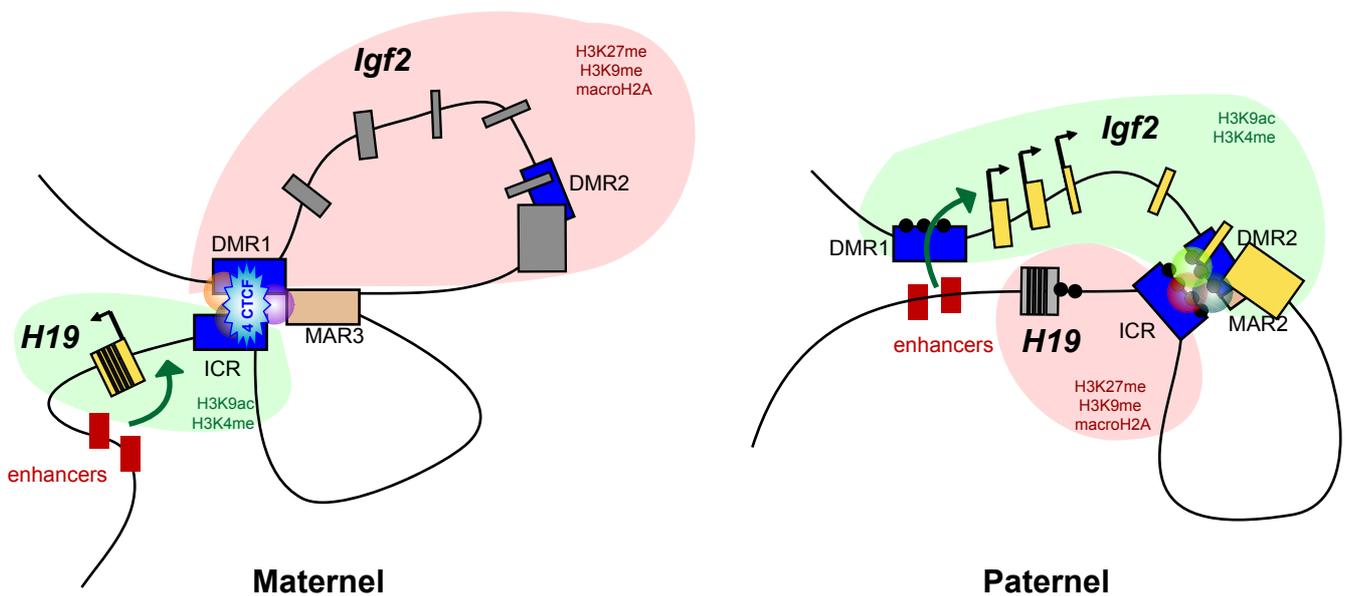
Deux autres séries d'enhancers ont été rapportées : un couple d'enhancers mésodermiques potentiels communs HUC1 et HUC2 (*H19* upstream conserved) en amont du gène *H19* (Drewell *et al.*, 2002a) et une région CDD (centrally conserved domain) qui serait un enhancer mésodermique spécifique du gène *Igf2* (Ainscough *et al.*, 2000; Jones *et al.*, 2001; Charalambous *et al.*, 2004).

Le centre d'Empreinte du locus est situé 2 kilobases en amont du gène *H19* (Tremblay *et al.*, 1995; Tremblay *et al.*, 1997). Il comprend deux régions : une DMR et un silencer. La DMR est complètement méthylée sur l'allèle paternel et complètement déméthylée sur l'allèle maternel. Elle contient quatre sites pour CTCF, qui se fixe donc uniquement sur l'allèle maternel (Bell and Felsenfeld, 2000; Hark *et al.*, 2000; Kaffer *et al.*, 2000; Kanduri *et al.*, 2000; Szabo *et al.*, 2000). Ainsi, sur l'allèle maternel, la protéine CTCF isole les signaux des enhancers qui ne peuvent pas activer les promoteurs maternels du gène *Igf2*. En revanche, sur l'allèle paternel, CTCF ne se fixe pas et les enhancers peuvent interagir avec les promoteurs du gène *Igf2* (Figure 23). De ce fait, **le gène *Igf2* est exprimé de façon monoallélique paternelle**. L'ICR contient également une séquence silencer qui permet la répression de l'allèle paternel du gène *H19* (Drewell *et al.*, 2000). De plus le promoteur de ce gène est également méthylé paternellement après la fécondation (Bartolomei *et al.*, 1993; Ferguson-Smith *et al.*, 1993; Sasaki *et al.*, 1995; Tremblay *et al.*, 1995; Srivastava *et al.*, 2000). Le promoteur maternel est protégé de cette méthylation *de novo* par la fixation de CTCF sur



**Figure 24 : Technique 3C (Chromosome Conformation Capture).**

Cette technique permet d'étudier les interactions entre des séquences éloignées de l'ADN. La conformation de la chromatine est fixée par pontage (crosslinking) et l'ADN est fragmenté. La ligation permet de refermer les séquences physiquement proches. Après réversion du pontage et digestions des protéines, une PCR peut être réalisée. Seules les séquences physiquement proches pourront donner lieu à l'amplification. Une PCRq avec une sonde Taqman permet d'obtenir un résultat fiable et quantitatif. (Adapté de Hagège *et al*, 2006)



**Figure 25 : Architecture de la chromatine au locus *H19/Igf2*.**

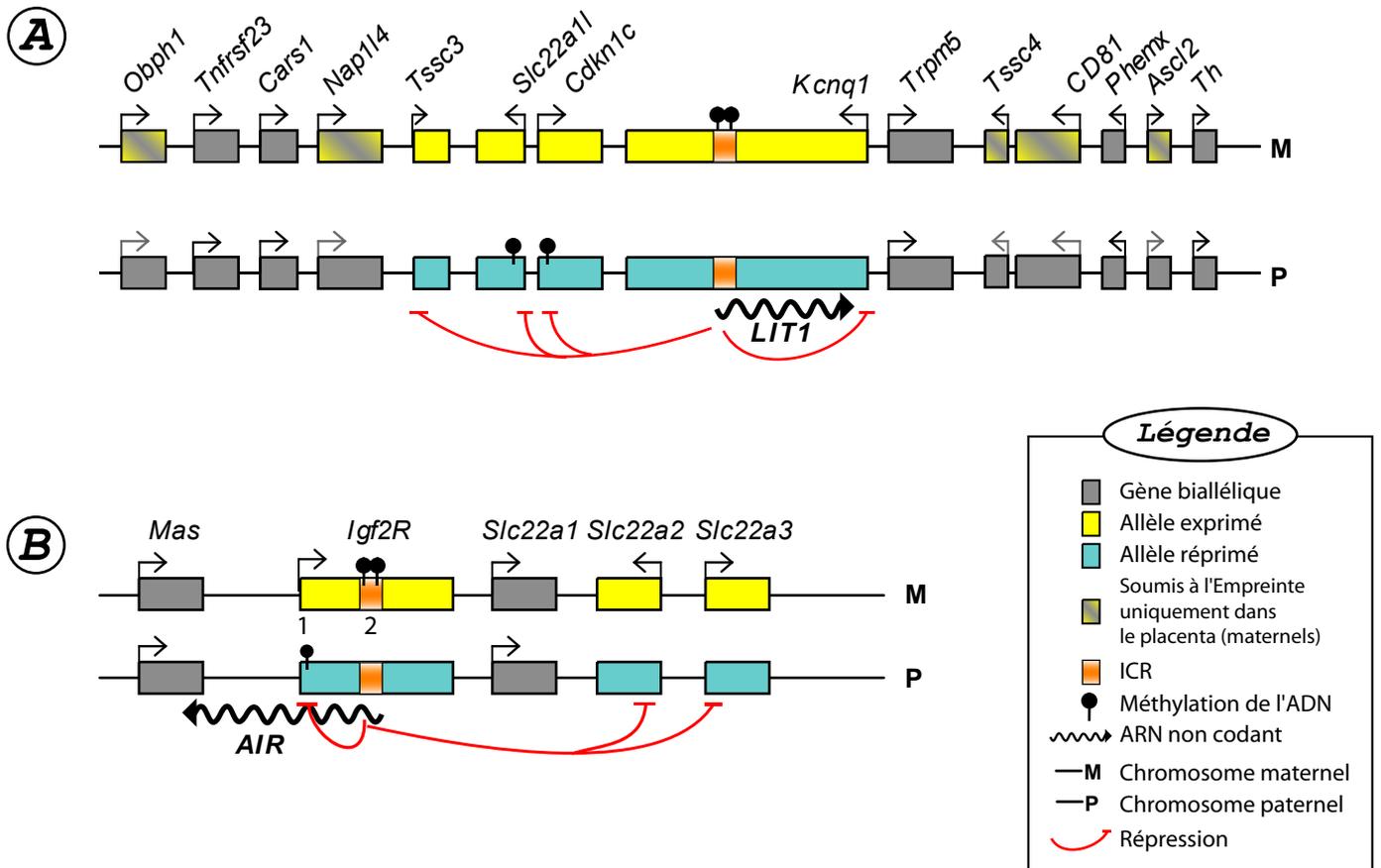
Différentes analyses de 3C ont montré une architecture différente sur les deux allèles. Sur l'allèle maternel, l'ICR interagit avec la DMR1 et MAR3. Le gène *Igf2* est à l'écart des enhancers et l'ADN forme une structure hétérochromatique (tache rouge). Les enhancers contactent le promoteur du gène *H19* et la structure de l'ADN est euchromatique (tache verte). Sur l'allèle paternel, l'ICR interagit avec la DMR2, probablement via des MBP. Les enhancers sont en contact avec les promoteurs du gène *Igf2* au niveau de la DMR1. La chromatine au niveau du gène *H19* est répressive.

l'ICR (Szabo *et al.*, 2004). Ainsi, **le gène *H19* est exprimé uniquement à partir de l'allèle maternel.**

Le gène *Igf2* comporte également trois DMR et quatre séquences d'attachement à la matrice nucléaire (MAR pour Matrix Attachment Region, [Figure 23](#)). La DMR1 permet l'inhibition de l'expression du gène *Igf2* sur l'allèle maternel où elle n'est pas méthylée (Constancia *et al.*, 2000). La DMR2 en revanche agit comme activateur transcriptionnel sur l'allèle paternel où elle est méthylée (Murrell *et al.*, 2001). De façon intéressante, MAR0 et MAR2 présentent un attachement différentiel à la matrice, l'allèle paternel exprimé étant le seul à être attaché dans les organes où les promoteurs adjacents sont actifs (Weber *et al.*, 2003).

De plus, des données récentes montrent que les histones sont différentiellement modifiées au niveau du locus et que CTCF est nécessaire à l'établissement de ces patrons épigénétiques. Ainsi, l'allèle paternel du promoteur du gène *H19* et de l'ICR et l'allèle maternel du gène *Igf2* sont marqués par les marques répressives H3K27me, H3K9me et par la présence du variant macroH2A. L'allèle maternel du gène *H19* et de l'ICR et l'allèle paternel du gène *Igf2* sont marqués par les marques d'activation H3K9ac et H3K4me (Choo *et al.*, 2006; Delaval *et al.*, 2007; Mikkelsen *et al.*, 2007; Han *et al.*, 2008).

L'ensemble de ces données suggère que la vision linéaire de la régulation de l'Empreinte à ce locus est simpliste. Plusieurs équipes ont donc tenté de mettre en évidence **l'architecture de la chromatine** de ce domaine. Pour ce faire, la technique 3C (Chromosome Conformation Capture) a été développée ([Figure 24](#)). Les résultats indiquent que l'ICR interagit avec les DMR du gène *Igf2* de façon différente sur les deux allèles ([Figure 25](#)) (Murrell *et al.*, 2004; Kurukuti *et al.*, 2006; Hagege *et al.*, 2007; Qiu *et al.*, 2008). Sur l'allèle maternel, l'ICR interagit avec la DMR1 et MAR3. L'architecture ainsi formée placerait le gène *Igf2* dans un domaine chromatinien inactif et empêcherait les enhancers d'interagir avec les promoteurs mais favoriserait l'interaction avec le promoteur du gène *H19*. Sur l'allèle paternel, l'ICR méthylé interagit avec la DMR2 méthylée et les enhancers sont proches de la DMR1. Cette boucle permettrait aux enhancers d'activer l'expression du gène *Igf2*. Les interactions entre l'ICR et les DMR pourraient se faire *via* CTCF ainsi que d'autres protéines. Il a notamment été rapporté que plusieurs MBP (methyl-CpG binding proteins) peuvent se lier



**Figure 26 : Régulation de l'Empreinte par les ARN non codants de grande taille.**

**A.** Le locus *Kcnq1* : l'ARNnc *LIT1* est exprimé à partir du centre d'Empreinte (ICR) sur l'allèle paternel et entraîne la répression bidirectionnelle *en cis* des gènes *Tssc3*, *Slc22a1*, *Cdkn1c* et *Kcnq1*. Dans le placenta, la répression *en cis* concerne aussi les gènes *Obph1*, *Nap114*, *Tssc4*, *CD81* et *Ascl2*. Ces gènes ont ainsi une expression monoallélique maternelle.

**B.** Le locus *Igf2R* : l'ARNnc *Air* est exprimé à partir de l'ICR sur l'allèle paternel et entraîne la répression bidirectionnelle *en cis* des gènes *Igf2R*, *Slc22a2* et *Slc22a3*. Ces gènes ont ainsi une expression monoallélique paternelle.

(Adapté de Gabory *et al.*, 2005).

à l'ICR sur l'allèle paternel méthylé. Elles pourraient ainsi contribuer à la régulation de l'Empreinte du locus (Drewell *et al.*, 2002b; Filion *et al.*, 2006; Reese *et al.*, 2007).

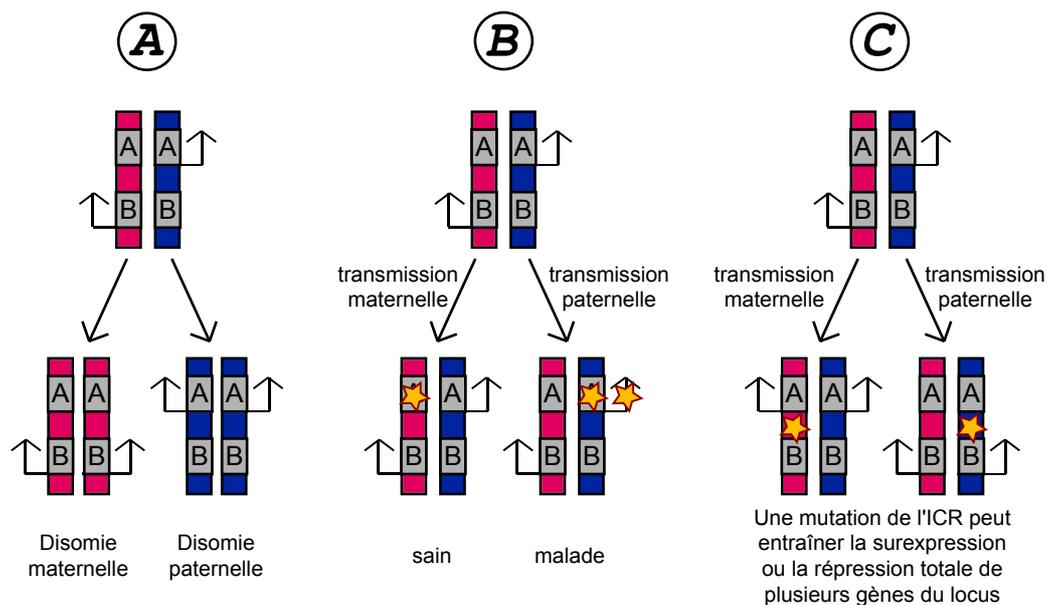
### 3-4-2) Régulation par les ARN non codants

Dans plusieurs régions chromosomiques, l'Empreinte est régulée par l'expression d'un ARN non codant de longue taille (Tableau 2). C'est le cas des loci PWS/AS (Prader-Willi Syndrome/Angelman Syndrome, dont un des gènes principaux est *SNRPN*), *CDKN1C* et *IGF2R*. Le mode de fonctionnement de ces deux derniers loci présente de nombreuses similarités qui seront détaillées dans la suite de cette partie.

Le locus *Kcnq1* est localisé sur la partie distale du chromosome 7 murin, région synténique du locus 11p15.5 humain. Il comporte un dizaine de gènes soumis à l'Empreinte répartis sur 700 kb (Figure 26A), la plupart sont exprimés à partir de l'allèle maternel, dont notamment *Kcnq1* (qui code un canal potassium voltage dépendant impliqué dans les battements cardiaques), *Cdkn1c* (ou *p57<sup>KIP2</sup>* une kinase cycline dépendante impliquée dans le cycle cellulaire) et *Tssc3* (ou *Ipl* qui code une protéine à domaine d'homologie pleckstrin). Un long ARNnc (60kb), *LIT1* (*Long intronic transcript 1* ou *Kcnq1ot1* pour *Kcnq1-overlapping transcript 1*), est transcrit sur l'allèle paternel, en antisens à partir de l'ICR (KvDMR dans l'intron 1 de *Kcnq1*) (Smilnich *et al.*, 1999; Fitzpatrick *et al.*, 2002).

Le locus *Igf2R* est localisé sur le chromosome 17 murin, région synténique du locus 6q26 humain. Il comporte quatre gènes soumis à l'Empreinte : *Igf2R* (Récepteur de IGFII), *Slc22a2* et *Slc22a3* (*solute carrier 22a*) exprimés maternellement et *Air* (*Antisense of Igf2R*) exprimé à partir de l'allèle paternel. *Air* est un ARNnc de grande taille (108 kb) transcrit en sens opposé au gène *Igf2R* (Figure 26B). Le locus contient deux DMR : la région 1 qui est méthylée sur l'allèle paternel après la fécondation et correspond au promoteur d'*Igf2R* et la région 2 qui est méthylée maternellement dans les gamètes et qui est localisée dans l'intron 2 du gène *Igf2R*. Cette séquence est l'ICR du locus et sert de promoteur pour transcrire l'ARN *Air* à partir de l'allèle paternel (Wutz *et al.*, 1997; Sleutels *et al.*, 2002).

Pour ces deux loci, la transcription de l'ARNnc à partir de l'ICR non méthylé est nécessaire à la répression bidirectionnelle *en cis* des gènes soumis à l'Empreinte du locus. La délétion de la DMR ainsi que la production d'ARN tronqués entraînent l'expression biallélique



**Figure 27 : Trois mécanismes associés aux pathologies d'Empreinte.**

Les rectangles roses représentent les chromosomes reçus de la mère et les bleus ceux reçus du père. Les rectangles grisés représentent deux gènes A et B soumis à l'Empreinte d'un même locus.

**A.** Altération de la transmission chromosomique : la disomie uniparentale maternelle entraîne la surexpression de A et l'absence d'expression de B. Un effet en miroir est observé après disomie paternelle.

**B.** Une mutation (étoile) dans un gène soumis à l'Empreinte aura un effet seulement s'il est transmis avec l'allèle exprimé (ici transmission paternelle puisque l'allèle paternel de A est exprimé et l'allèle maternel réprimé).

**C.** Une mutation ou une épimutation (étoile) du centre d'Empreinte peut entraîner une perte d'Empreinte de tous les gènes du locus. Les gènes seront alors ou surexprimés de façon biallélique ou totalement réprimés. (Adapté de Ferguson-Smith *et al.*, 2003)

des gènes soumis à l'Empreinte, suggérant l'implication de l'ARNnc en lui-même dans le processus de répression (Sleutels *et al.*, 2002; Mancini-Dinardo *et al.*, 2006). Trois mécanismes sont proposés (Pauler *et al.*, 2007) :

- l'ARNnc agit lui-même selon un modèle proche de celui de *Xist* ou via la production de petits ARNnc.
- la transcription de l'ARN pourrait entraîner la formation d'une architecture spécifique de la chromatine en induisant des boucles ou en plaçant les gènes dans des compartiments transcriptionnellement inactifs.
- la transcription de l'ARNnc au niveau du locus réprimerait la fixation d'un facteur important pour l'initiation de l'expression des autres gènes.

Dans le cas du locus *Kcnq1*, l'Empreinte de plusieurs gènes est restreinte au placenta. Dans cet organe, les modifications des histones, plus que la méthylation de l'ICR, semblent gouverner l'expression monoallélique via *LIT1* (Lewis *et al.*, 2004; Umlauf *et al.*, 2004). De plus, l'ICR aurait également une fonction insulatrice puisqu'elle est capable de fixer la protéine CTCF (Fitzpatrick *et al.*, 2007). Selon les tissus, l'Empreinte des différents gènes du locus *Kcnq1* pourrait être régulée soit par *LIT1* soit par l'insulateur (Shin *et al.*, 2008).

### 3-5) Empreinte Parentale et Pathologie humaine

La dérégulation de gènes soumis à l'Empreinte Parentale a été impliquée dans diverses pathologies humaines. Les principales causes de maladies de l'Empreinte sont les disomies uniparentales (présence de deux copies d'une région chromosomique provenant d'un parent et absence de la copie de l'autre parent) et les pertes d'Empreinte (LOI pour loss of imprinting) dues à des défauts de marques épigénétiques comme la méthylation de l'ADN ou à des mutations dans les ICR (Figure 27) (Ferguson-Smith *et al.*, 2003). Ainsi des défauts d'Empreinte sont à l'origine de certains syndromes de retards mentaux, tels que les syndromes de Prader-Willi et Angelman ou des dérégulations de la croissance, comme les syndromes de Beckwith-Wiedemann et de Silver-Russell. De plus, le dysfonctionnement de gènes soumis à l'Empreinte a été constaté dans divers cas de cancers.

Les gènes *H19* et *Igf2* ont notamment été impliqués dans les syndromes de Beckwith-Wiedemann et de Silver-Russell. Ces pathologies seront développées dans le chapitre III.

## *Chapitre II*

### *Croissance et Myogenèse*

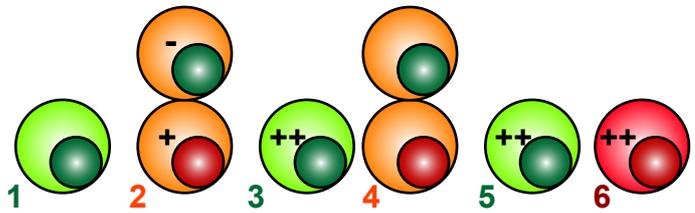
**A**

Inhibition des IGF par séquestration

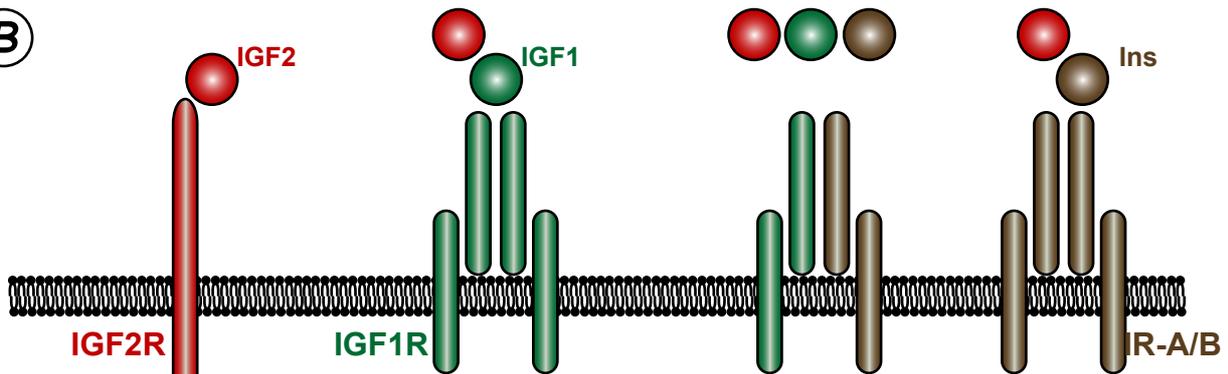
Prolongation de la vie des IGF dans le plasma

Potentialisation de l'action des IGF

IGFBP

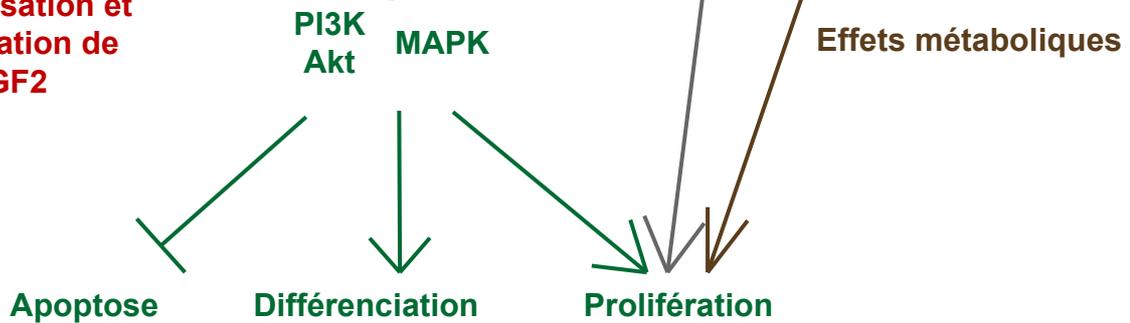


**B**



**C**

Internalisation et dégradation de l'IGF2



**Figure 28 : Les acteurs de la voie Insulin-like Growth-Factor (IGF).**

**A.** Les IGF-binding proteins (IgfBP) peuvent lier avec affinité (forte + ou ++, faible -) les ligands Igf1 (vert clair), Igf2 (rouge) ou les deux (orange) dans le sang et les espaces extra cellulaires.

**B.** Les ligands Igf1, Igf2 et insuline (Ins) peuvent lier quatre types de récepteurs. Igf2 lie avec forte affinité le récepteur de type 2, Igf2R. Le récepteur de type 1, Igf1R, peut fixer Igf1, Igf2 et Ins avec des affinités décroissantes. Le récepteur à insuline, IR, est présent sous deux isoformes IR-A et IR-B. Ces deux isoformes lient l'insuline, alors que seule IR-A foetale peut lier Igf2. Enfin, une forme hétérodimérique Igf1R/IR peut lier les trois ligands.

**C.** Igf2-mannose6P-R est couplé à l'internalisation et à la dégradation d'Igf2. Igf1R est couplé aux voies PI3K et MAPK et entraîne ainsi la prolifération, la protection contre l'apoptose ou la différenciation. IR et Igf1R/IR entraînent les effets métaboliques de l'insuline.

## 1) Croissance embryonnaire et post-natale

### 1-1) Introduction

La croissance est le **processus par lequel un individu atteint sa taille définitive**. C'est un mécanisme continu depuis le stade blastocyste jusqu'à l'âge adulte. Chez l'embryon, la croissance est maximale au cours de la période fœtale et ralentie en fin de gestation. Après la naissance, la croissance est dans un premier temps rapide (au cours des 3 premières semaines chez la souris, ce qui peut correspondre à l'enfance chez l'Homme) puis plus lente. Une dernière poussée a lieu au moment du sevrage (qui pourrait correspondre à l'adolescence chez l'Homme). A l'âge adulte, la croissance est quasiment nulle. Les signaux qui favorisent ou inhibent la croissance sont en balance et leur résultante entraîne ces périodes de développement ou d'arrêt.

Les deux principaux acteurs de la croissance sont les **facteurs de croissance** et les **hormones** (Efstratiadis, 1998). Les facteurs de croissance sont des protéines, sécrétées par différents types de tissus, nécessaires à la prolifération cellulaire. Les hormones sont des substances synthétisées par des cellules spéciales et sécrétées directement dans le sang ou la lymphe ; certaines d'entre elles influencent la croissance. Les principales molécules décrites pour agir sur la croissance de l'individu dans son ensemble sont **les protéines de l'axe IGF** (insulin-like growth factor) et **l'hormone de croissance (GH)** pour Growth Hormone). L'axe IGF a un rôle majeur dans la croissance embryonnaire et la GH, en revanche, joue plutôt un rôle dans la vie post-natale, via la stimulation de l'expression du gène *Igf1*.

### 1-2) L'axe Insulin-like Growth-Factor : une analyse génétique poussée

La voie de signalisation insuline/IGF (Figure 28) se compose de :

- trois peptides ligands : l'insuline, IGF1 et IGF2,
- quatre récepteurs : le récepteur à insuline (IR), le récepteur à IGF de type I (IGF1R), un récepteur hybride IGF1R/IR et le récepteur à IGF de type II (IGF2R)
- six protéines sériques solubles de liaison de l'IGF1 et/ou de l'IGF2 : IGFBP1 à 6 (IGF-binding proteins).

Elle joue un rôle critique dans la régulation de l'homéostasie glucidique et la croissance d'un individu (Randhawa and Cohen, 2005). Ce rôle dans la croissance a été particulièrement mis

Mutant	Poids à la naissance	Poids adulte
sauvage	 100%	100% 
<i>Igf1</i> <sup>-/-</sup>	 60%	30% 
<i>Igf2</i> <sup>-/-</sup>	 60%	60% 
<i>Igf1R</i> <sup>-/-</sup> <i>Igf1</i> <sup>-/-</sup> <i>Igf1R</i> <sup>-/-</sup>	 45%	léthalité à la naissance
<i>Igf2R</i> <sup>-/-</sup>	 140%	léthalité périnatale
<i>Igf2</i> <sup>-/-</sup> <i>Igf2R</i> <sup>-/-</sup>	 70%	70% 
<i>IR</i> <sup>-/-</sup> <i>Ins1</i> <sup>-/-</sup> <i>ins2</i> <sup>-/-</sup>	 90%	léthalité due au diabète

**Tableau 3 : Mutations nulles de la voie Insulin-like Growth-Factor (IGF).**

L'analyse des simples, doubles et triples mutants des acteurs de la voie IGF a permis de mettre en évidence le rôle crucial de cette voie dans la croissance embryonnaire et post-natale. Ce tableau représente les poids de naissance et adultes de certains mutants comparés à celui de souris sauvages.

en lumière par la production de mutants invalidés pour les différents acteurs de cette voie (**Tableau 3**) (Louvi *et al.*, 1997; Efstratiadis, 1998). Pour ces modèles murins, les animaux présentent des défauts de croissance harmonieux, ce qui permet d'utiliser le poids comme caractéristique de leur croissance globale.

Les souris simples mutantes pour *Igf1* ou *Igf2* ont un poids de naissance réduit de 40% par rapport aux souris sauvages, les souris *Igf2<sup>-/-</sup>* présentant également une atrophie du placenta. La différence de taille et de poids avec les souris sauvages reste constante après la naissance pour les *Igf2<sup>-/-</sup>*, en revanche les *Igf1<sup>-/-</sup>* ont un défaut majeur de croissance post-natale et leur taille n'atteint que 30% de la taille normale adulte (DeChiara *et al.*, 1990; Liu *et al.*, 1993). Les doubles mutants *Igf1<sup>-/-</sup> Igf2<sup>-/-</sup>* présentent un retard de croissance considérable puisque leur poids de naissance est équivalent à 30% du poids normal. Cela montre bien le **rôle central des IGF dans la croissance embryonnaire**, sans préjudice sur la survie.

De plus, les mutants des récepteurs ont également des phénotypes de croissance caractéristiques : les animaux *IGF1R<sup>-/-</sup>* font 45% du poids normal et les animaux *Igf2R<sup>-/-</sup>* ont une surcroissance de 40% (Liu *et al.*, 1993; Wang *et al.*, 1994). Ces récepteurs sont donc les régulateurs principaux du signal IGF, IGF1R induisant le signal de prolifération et IGF2R étant un régulateur négatif.

Enfin, la voie insuline a un rôle métabolique très important mais un rôle plus secondaire dans la croissance. Les souris mutantes n'exprimant pas IR ou n'exprimant pas d'insuline ont en effet un retard de croissance de 10% à la naissance et meurent rapidement en raison d'un diabète sévère (Joshi *et al.*, 1996; Duvillie *et al.*, 1997).

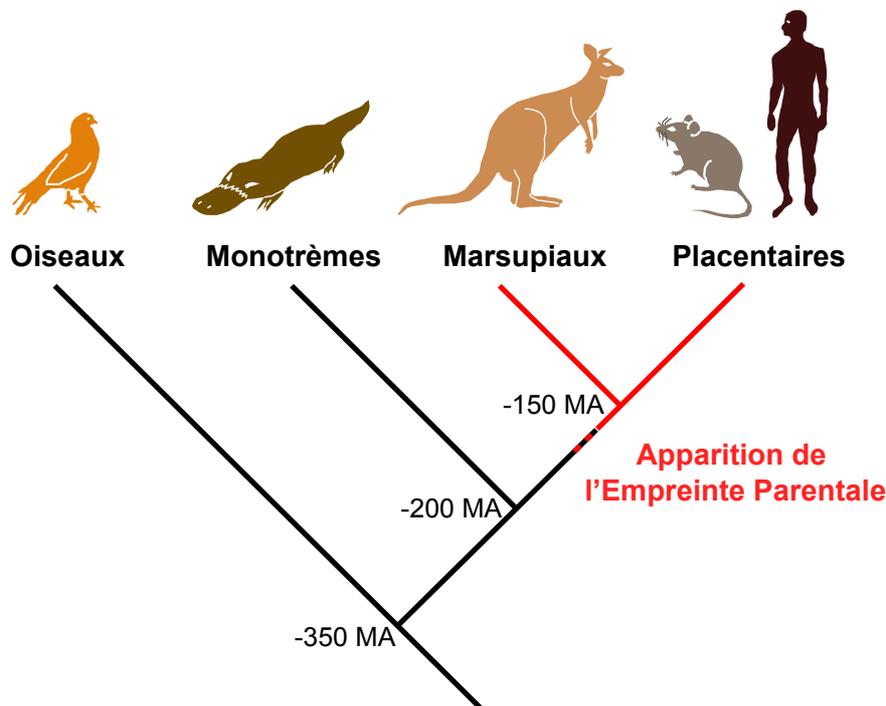
Ces nombreux croisements entre les mutants ainsi que des expériences cellulaires de diverses constructions ont permis de déchiffrer les interactions entre ligands et récepteurs et un modèle a ainsi pu être établi (**Figure 28**) (Randhawa and Cohen, 2005; Riedemann and Macaulay, 2006). Le récepteur à l'insuline existe sous deux isoformes IR-A, forme fœtale qui lie IGF2 avec une grande affinité et permet la prolifération et la protection contre l'apoptose, et IR-B qui est l'effecteur de l'insuline dans le foie, les muscles et les tissus adipeux. IGF1R fixe IGF1 et IGF2 et induit, *via* les voies PI3K, MAPK et RAF, la survie cellulaire, la prolifération ou la différenciation. IGF2R n'a pas de fonction de transmission d'un signal mais ce récepteur, également appelé MPR (pour Mannose-6-phosphate receptor), est couplé à la voie de l'internalisation des protéines lysosomales. Aussi il est postulé qu'IGF2R permet de réguler le taux d'IGF2 circulant par endocytose et dégradation. Enfin, la biodisponibilité des

IGF dans le milieu sérique est régulée par les protéines de liaison aux IGF (IGFBP pour IGF binding-proteins).

### 1-3) Croissance embryonnaire et croissance post-natale.

La croissance d'un individu peut être partagée en deux périodes : la croissance embryonnaire et la croissance post natale. La croissance embryonnaire est régulée par des facteurs environnementaux, extra-embryonnaires et embryonnaires. En effet, chez les mammifères, le régime alimentaire de la mère influe sur le développement de l'embryon et le passage efficace des nutriments et de l'oxygène à travers le filtre placentaire est essentiel. Au niveau embryonnaire, **l'axe IGF est un des facteurs clé de la croissance**. L'étude des mutants murins a montré que ces deux phases ne sont pas gouvernées par les mêmes facteurs. Par exemple, la comparaison des mutants *Igf2<sup>-/-</sup>* et *Igf1<sup>-/-</sup>* montre que le premier agit au cours de l'embryogenèse alors que le second agit pendant les deux périodes (DeChiara *et al.*, 1990; Liu *et al.*, 1993). En effet, la différence de poids de 40% des souris invalidées pour *Igf2* par rapport aux sauvages à la naissance reste stable alors que dans le cas des souris *Igf1<sup>-/-</sup>*, elle s'aggrave pour atteindre 70% à l'âge adulte (Tableau 3). De plus, les souris *Igf1<sup>-/-</sup>* n'ont pas d'anomalie placentaire alors que les souris *Igf2<sup>-/-</sup>* ont un placenta atrophié. L'effet d'IGF2 sur le placenta a été avéré par des expériences de délétion du variant placentaire de l'ARNm *Igf2* (P0). Cette mutation entraîne la perte de l'expression d'IGF2 dans les cellules trophoblastiques du placenta mais le taux fœtal d'IGF2 circulant n'est pas atteint. Le placenta est plus petit que celui d'une souris sauvage dès E14 pour atteindre une différence de 30% et l'embryon présente un défaut de croissance à partir de E16 pour atteindre une différence de 30% à la naissance (Constancia *et al.*, 2002). Ces expériences montrent un rôle particulier pour IGF2 dans le placenta et le transfert de nutriment de la mère à l'embryon, processus important pour la croissance fœtale.

Il existe également des lignées de souris avec des mutations spontanées induisant un phénotype de nanisme, comme les souris *little* ou *dwarf*. Ces souris sont de taille normale à la naissance mais présentent un retard de croissance visible à partir du jour 15 et leur taille adulte ne dépasse pas la moitié de celle de souris sauvages. Ces animaux présentent des mutations au niveau des gènes *Pit1* et *Ghrh* qui codent des facteurs de la voie de l'hormone de croissance (Pfaffle *et al.*, 1992; Radovick *et al.*, 1992; Godfrey *et al.*, 1993; Lin *et al.*, 1993). Chez ces souris, le taux de GH circulante est dramatiquement diminué.



**Figure 29 : Co-évolution de l'Empreinte et de la placentation.**

L'Empreinte existe chez la plupart des mammifères placentaires et chez les marsupiaux, qui ont un placenta rudimentaire. Ce mécanisme n'est présent ni chez les monotrèmes qui, bien qu'ils allaitent leurs petits, pondent des oeufs, ni chez les oiseaux. L'Empreinte serait donc apparue dans la branche commune aux marsupiaux et aux placentaires, ce qui suggère une co-évolution avec la placentation.

MA : millions d'années. (Adapté de Kilian *et al.* 2001)

	Croissance Embryonnaire	Croissance Placentaire	Effet sur le métabolisme		Autre
			Glucides	Lipides	
<i>Igf2</i>	favorise	favorise	Tg : hypoglycémie	KO (cerveau) : adiposité accrue Tg : adiposité réduite	
<i>Igf2R</i>	inhibe	inhibe			KO : léthal
<i>Cdkn1c</i>		inhibe			
<i>Dlk1</i>	favorise		Tg : résistance à l'insuline tolérance au glucose	KO : adiposité accrue Tg : adiposité réduite	Tg : hyperphagie
<i>Grb10</i>	inhibe	inhibe	KO : sensibilité à l'insuline Tg : hyperglycémie résistance à l'insuline	KO : adiposité réduite Tg : adiposité accrue	
<i>Dio3</i>	inhibe				
<i>Mest</i>	favorise	inhibe		Tg : adiposité accrue	
<i>Peg3</i>	favorise	inhibe		KO : adiposité accrue	KO : hypophagie
<i>Zac1</i>	favorise		Tg : hyperglycémie néonatale légère résistance à l'insuline		KO : léthal
<i>Gnas</i>			KO : hyperglycémie résistance à l'insuline	KO : adiposité accrue	KO : léthal
<i>Gnasxl</i>			KO : sensibilité à l'insuline	KO : adiposité réduite	KO : léthal
<i>Rasgrf1</i>			KO : intolérance au glucose	KO : adiposité réduite	
<i>Atp10c</i>			KO : résistance à l'insuline	KO : adiposité accrue	

**Tableau 4 : Gènes soumis à l'Empreinte et leurs effets.**

Un grand nombre de gènes soumis à l'Empreinte ont un effet sur la croissance embryonnaire et le métabolisme des glucides et lipides. Les gènes d'expression maternelle (rouge) ont en général tendance à ralentir la croissance embryonnaire alors que les gènes paternels (bleu) la favorise. Les effets sur le métabolisme sont également décrits dans ce tableau.

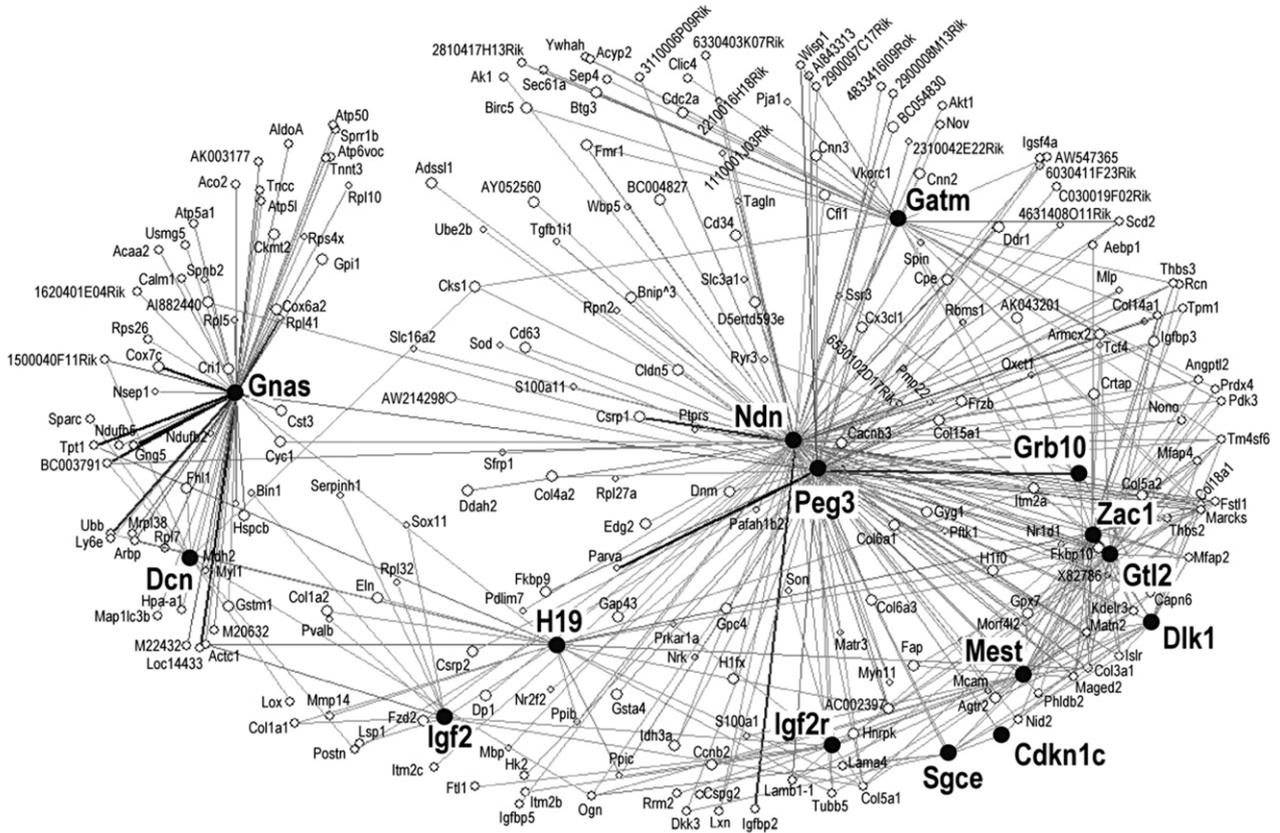
(Adapté de Charalambous *et al.* 2007 et Smith *et al.* 2006)

L'hypothalamus sécrète un facteur de libération de l'hormone de croissance (GH pour growth hormone) : la GHRH (GH releasing hormone) qui agit sur son récepteur au niveau de l'hypophyse et permet la sécrétion de la GH dans le sang. Cette sécrétion est pulsatile et régulée en fonction du cycle circadien. La GH circule vers le foie où la fixation à son récepteur (GHR) permet la sécrétion d'IGF1, qui va pouvoir induire ses effets de prolifération, etc... Par ailleurs, IGF1 exerce un rétrocontrôle négatif en inhibant la production de GH. Ainsi, **la croissance post natale est finement régulée par la voie GH-IGF1.**

#### **1-4) Croissance embryonnaire et Empreinte Parentale.**

Peu de temps après la découverte des premiers gènes soumis à l'Empreinte, une **théorie dite du "conflit parental"** a vu le jour pour tenter d'analyser l'intérêt, sur le plan évolutif, d'un tel mécanisme (Moore and Haig, 1991). Cette théorie s'appuie sur le rôle des gènes soumis à l'Empreinte dans la croissance embryonnaire, les fonctions placentaires et les relations de la mère avec ses petits. Selon cette théorie, les intérêts du père et de la mère dans la dispersion de leur patrimoine génétique sont antagonistes. La mère doit avoir de nombreuses portées et pour cela doit limiter la croissance et les apports aux embryons, afin de préserver son propre organisme et ses futures portées. Le père doit avoir des petits forts et en bonne santé et pour cela permettre la croissance de l'embryon, afin de contribuer à la dissémination de son propre patrimoine génétique. Les portées suivantes sont assurées par des femelles différentes. Ainsi, l'intérêt du développement de l'Empreinte parentale serait lié au développement *in utero* des mammifères, qui fait intervenir une connexion foeto-maternelle *via* le placenta tout au long de la gestation. Cette théorie est appuyée par le fait que l'Empreinte est spécifique des mammifères placentaires et semble avoir co-évolué avec la placentation (Figure 29). L'ornithorynque, qui allaite ses petits mais pond des œufs et n'a donc pas de placenta, est dépourvu d'Empreinte : les gènes *Igf2* et *Igf2R* s'expriment de façon biallélique. Chez les marsupiaux, qui ont un placenta rudimentaire, quelques gènes comme *Igf2* sont soumis à l'Empreinte (Killian *et al.*, 2001). La relation intra-utérine materno-fœtale doit ainsi être régulée selon les intérêts du père et de la mère.

De façon concordante avec la théorie du conflit parental, les gènes soumis à l'Empreinte jouent **un rôle important dans la croissance embryonnaire et le métabolisme.** Les souris invalidées ou surexprimant un gène soumis à l'Empreinte ont presque toutes un phénotype de croissance embryonnaire ou de défaut métabolique (Smith *et al.*, 2006;

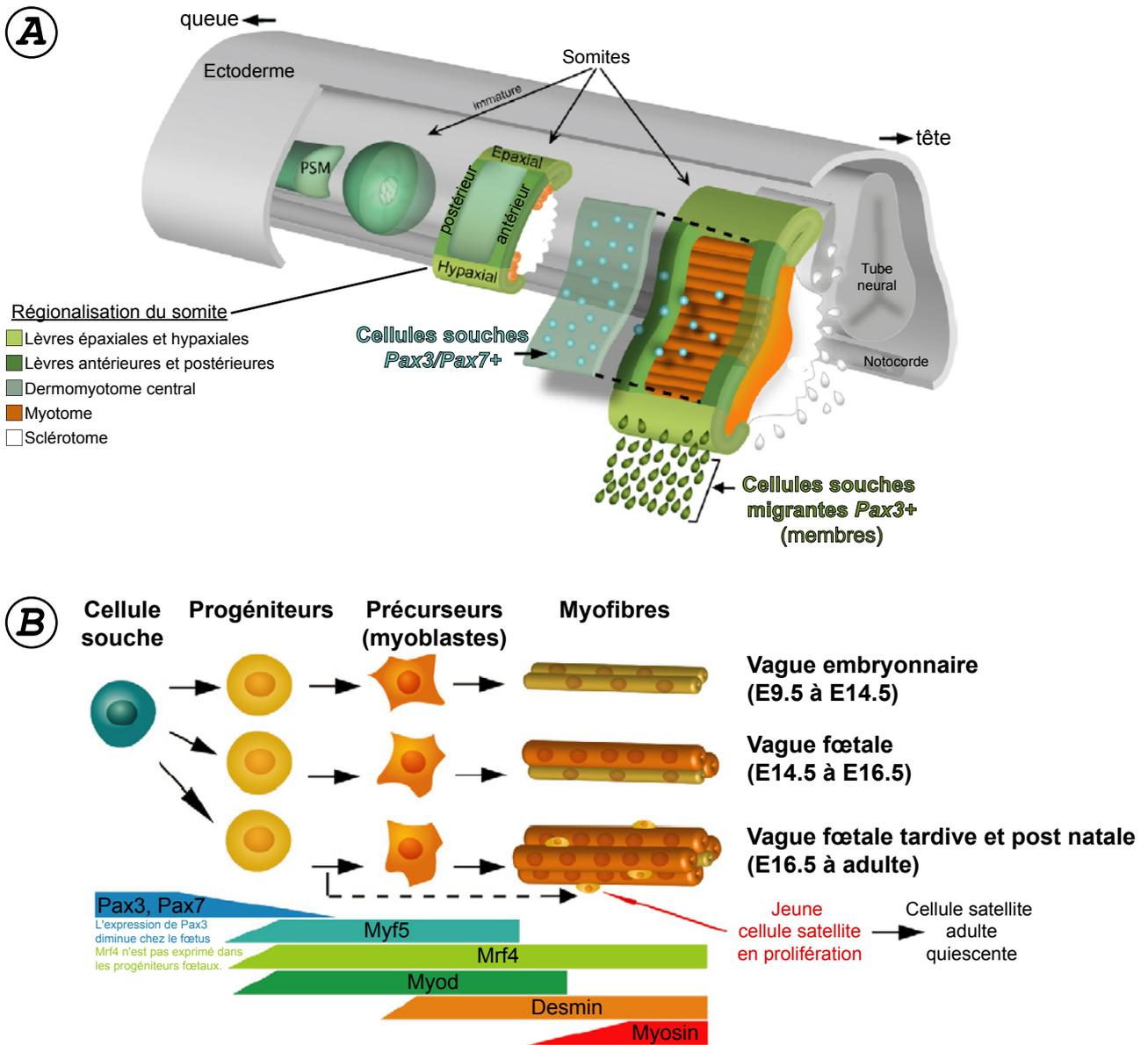


**Figure 30 : Réseau de gènes soumis à l'Empreinte.**

Une méta-analyse des gènes co-exprimés avec les gènes soumis à l'Empreinte a permis de mettre en évidence un réseau (IGN pour Imprinted genes network). Sur ce réseau sont représentés 246 gènes liés à au moins 3 gènes soumis à l'Empreinte. L'intensité des lignes représente la force du lien, les lignes noires correspondant aux gènes les plus fréquemment co-exprimés. (Adapté de Varrault *et al.* 2006)

Charalambous *et al.*, 2007). Les gènes à expression paternelle sont en général des facteurs de croissance et de métabolisme du glucose et des lipides alors que les gènes maternels régulent négativement la croissance embryonnaire et le métabolisme. Ainsi au niveau de la croissance embryonnaire, le gène *Igf2*, d'expression paternelle, est crucial pour la croissance alors que le gène *Igf2R*, d'expression maternelle, la réprime (DeChiara *et al.*, 1990; Ludwig *et al.*, 1996). De même, le gène maternel *Grb10* inhibe la croissance et le gène paternel *Dlk1* la favorise, de façon indépendante de l'axe IGF (Moon *et al.*, 2002; Charalambous *et al.*, 2003). Au niveau du métabolisme du glucose et des lipides, les gènes paternels *Igf2*, *Gnas*, *Dlk1*, *Rasgrf1* ou encore *Zac1* ont un effet de régulation négative, les souris invalidées pour ces gènes présentent une hypoglycémie et/ou une adiposité réduites et les souris qui surexpriment un de ces gènes ont au contraire une hyperglycémie et/ou une adiposité augmentées (Tableau 4).

Il est récemment apparu que les gènes soumis à l'Empreinte sont fréquemment co-exprimés dans les mêmes tissus, suggérant un **réseau fonctionnel**. En analysant les données d'une centaine d'analyses de puces à ADN, Varrault *et al* ont observé que, parmi les gènes co-exprimés avec le gène *Zac1*, les gènes soumis à l'Empreinte sont surreprésentés. La même analyse systématiquement appliquée aux gènes co-exprimés avec les gènes soumis à l'Empreinte a permis d'établir un réseau (Figure 30) de gènes soumis à l'Empreinte ou non (IGN pour imprinted genes network). L'IGN comporte les gènes *Zac1*, *Gtl2*, *H19*, *Mest*, *Dlk1*, *Peg3*, *Grb10*, *Igf2*, *Igf2r*, *Dcn*, *Gnas*, *Gatm*, *Ndn*, *Slc38a4*, *Cdkn1c* et *Sgce*. Lorsque l'expression du gène *Zac1* est perturbée, dans le cas de la mutation nulle, l'expression des autres gènes de l'IGN est modifiée, ce qui suggère un lien fonctionnel entre les acteurs de l'IGN (Varrault *et al.*, 2006). Ce réseau pourrait ainsi jouer un rôle dans la régulation de la croissance et du métabolisme.



**Figure 31 : La myogenèse embryonnaire.**

**A.** Le mésoderme paraxial présomitique (PSM) est à l'origine des somites. Les somites se différencient en sclérotome, qui donnera le squelette axial, et en dermomyotome, qui donnera le derme cutané et les muscles squelettiques. Le myotome, qui donnera les muscles du tronc, est enrichi de cellules Pax3/Pax7+ qui migrent depuis le dermomyotome central. A partir des lèvres hypaxiales, migrent des cellules Pax3+, qui seront à l'origine des muscles des membres, du diaphragme et de la langue.

**B.** La myogenèse se fait selon trois vagues : embryonnaire, foetale et post-natale. Pour chaque vague, les cellules souches se différencient en progéniteurs puis en myoblastes et fusionnent en myofibres, grâce à un programme d'expression génique spécifique. La dernière vague voit le positionnement des cellules satellites, les cellules souches musculaires chez l'adulte, en position basale.

(Adapté de Sambasivan et Tajbakhsh, 2007)

## 2) Myogenèse

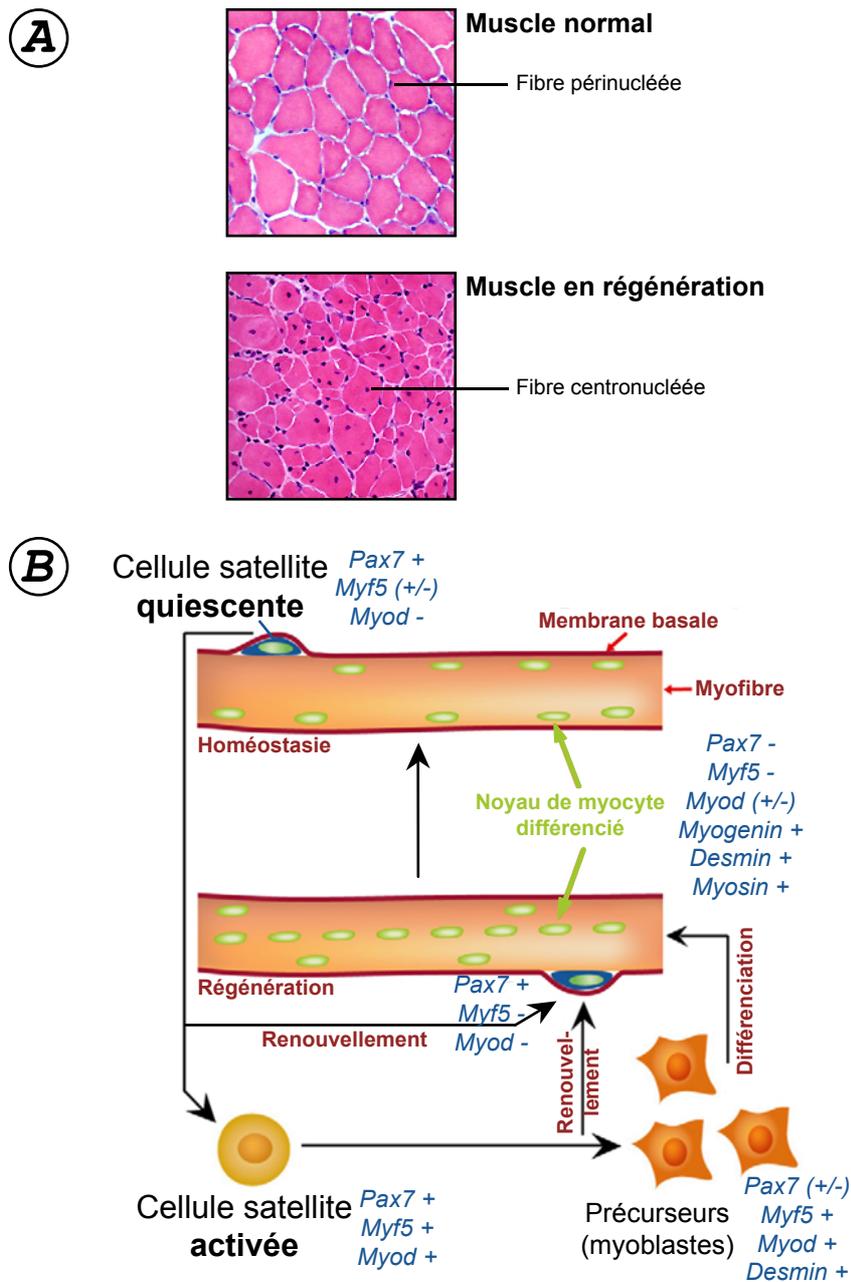
Chez les animaux, le mouvement, volontaire ou non, nécessite le fonctionnement de muscles. Il existe trois types de muscles chez les mammifères : le muscle lisse, le muscle strié cardiaque et le muscle strié squelettique. Les muscles lisses sont contrôlés par le système nerveux parasympathique et leur contraction est involontaire. Ils assurent par exemple le mouvement du système digestif et des vaisseaux sanguins. Les muscles striés squelettiques sont en général associés au squelette via les tendons, leur contraction est volontaire régie par le système nerveux sympathique et ils assurent le fonctionnement de l'appareil locomoteur. Le muscle cardiaque est strié mais sa contraction rythmique est sous le contrôle du système nerveux autonome. Dans ce chapitre, le développement et la régénération du muscle strié squelettique seront développés.

### 2-1) Myogenèse embryonnaire

Les muscles squelettiques sont des tissus dérivés du mésoderme. Dans l'embryon, le mésoderme présomitique paraxial est segmenté en somites. Ces structures se différencient ensuite en sclérotome, qui donnera le squelette axial, et en dermomyotome (DM), régionalisé en un dermomyotome central et quatre lèvres épaxiales, hypaxiales, antérieures et postérieures (Figure 31). A partir des bordures du DM, les cellules prolifèrent et migrent entre le DM et le sclérotome pour former le myotome. Les fibres des myotomes adjacents issus des somites vont fusionner pour donner la musculature du tronc. Aux niveaux cervical, occipital et des bourgeons de membres, des cellules de la lèvre hypaxiale migrent pour former respectivement les muscles de la langue, du diaphragme et des membres. Le processus de différenciation musculaire fait intervenir les cellules souches dont la division génère des progéniteurs déterminés dans un premier temps en précurseurs myoblastes et se différenciant ensuite en myofibres. Les myofibres sont produites par la fusion de cellules musculaires pour former un syncytium (Sambasivan and Tajbakhsh, 2007). Les myofibres matures sont donc de grandes et longues cellules plurinucléées à noyaux périphériques.

La différenciation musculaire se fait sous l'action des facteurs de transcription paired-box, Pax3 et Pax7, et de facteurs de transcription bHLH (basic helix-loop-helix) MRF (myogenic regulatory factors), MyoD, Myf5, MRF4 (ou Myf6) et Myogenin (Sambasivan and Tajbakhsh, 2007) (Figure 31). Les protéines Pax3 et Pax7 sont toutes deux nécessaires, et ont des fonctions redondantes, pour l'apparition des progéniteurs musculaires. Tout au long de la

myogenèse, des cellules souches *Pax3* et *Pax7* positives migrent depuis le DM central pour assurer un apport constant de progéniteurs au niveau du myotome (Relaix *et al.*, 2005). En revanche les progéniteurs migrants à partir de la lèvre hypaxiale sont sous la seule dépendance de *Pax3* et les souris mutantes *Pax3*<sup>-/-</sup> ne développent pas de muscle au niveau des membres (Tremblay *et al.*, 1998). *Pax7* serait plus important pour la survie cellulaire et le renouvellement du stock de progéniteurs, les souris *Pax7*<sup>-/-</sup> ne présentent pas de phénotype musculaire à la naissance mais sont incapables de régénérer leurs fibres lésées (Relaix *et al.*, 2006). Les gènes *MyoD* et *Myf5* sont des cibles des facteurs de transcription *Pax3* et *Pax7* (Bajard *et al.*, 2006; Relaix *et al.*, 2006). *MyoD* et *Myf5* sont des facteurs de transcription majeurs pour déterminer les progéniteurs vers la lignée musculaire (Tajbakhsh *et al.*, 1997). Les simples mutants *MyoD*<sup>-/-</sup> ou *Myf5*<sup>-/-</sup> présentent une musculature appropriée à la naissance, en revanche le phénotype des doubles mutants *MyoD*<sup>-/-</sup>*Myf5*<sup>-/-</sup> est très sévère avec une absence totale de musculature squelettique au niveau du tronc et des membres (Rudnicki *et al.*, 1993). Ce phénotype rend compte de la redondance fonctionnelle et de l'importance cruciale de ces facteurs myogéniques. *Myogenin* est requis pour la différenciation des précurseurs en myofibres (Hasty *et al.*, 1993; Nabeshima *et al.*, 1993). La fonction de MRF4 au cours de la différenciation musculaire de l'embryon est moins évidente. Il joue un rôle au cours de la vague embryonnaire dans la détermination des myoblastes et peut aussi activer l'expression de *Myogenin* pour induire la différenciation (Kassar-Duchossoy *et al.*, 2004).



**Figure 32 : Cellules satellites et régénération musculaire.**

**A.** L'histologie d'un tissu musculaire normal présente des fibres à noyaux périphériques. Un muscle en régénération présente des fibres à noyaux centraux.

**B.** Une myofibre normale, en haut, présente des noyaux périphériques et des cellules satellites situées entre le plasmalemme (rose pâle) et la membrane basale (rouge). Lorsque le muscle est lésé, les cellules satellites sont activées et se différencient en myoblastes qui fusionnent pour donner de nouvelles myofibres. Un autorenouvellement des cellules satellites est également assuré. La myofibre du muscle en régénération présente des noyaux centraux et des cellules satellites. Au cours de ce processus, différents gènes (bleus) sont activés ou réprimés à des moments précis pour assurer l'activation, la détermination, la différenciation et la fusion des cellules musculaires ainsi que l'autorenouvellement des cellules satellites.

(Adapté de Sambasivan et Tajbakhsh, 2007)

## 2-2) Régénération musculaire

A l'âge adulte, les muscles striés squelettiques sont composés de cellules syncytiales, les myofibres ([figure 32](#)). Au cours de la maturation des cellules musculaires, les noyaux se décentrent en périphérie. Autour de ces myofibres, entre le plasmalemme et la membrane basale, se trouvent de petites cellules quiescentes : les cellules satellites. Ces cellules sont considérées comme les cellules souches musculaires, progéniteurs capables de régénérer le tissu musculaire en cas de lésions (Le Grand and Rudnicki, 2007; Sambasivan and Tajbakhsh, 2007). Lorsque les cellules satellites quittent l'état de quiescence, elles sont alors activées : elles prolifèrent, migrent au niveau de la lésion, se différencient et fusionnent pour donner de nouvelles fibres musculaires. Au cours de ce processus, une petite partie des cellules garde les caractéristiques de cellules satellites, assurant ainsi l'auto-renouvellement du stock de cellules souches/progéniteurs.

L'activation des cellules satellites se fait sous le contrôle de plusieurs facteurs micro-environnementaux tels que le HGF (hepatocyte growth factor) et les FGF (fibroblast growth factors). En effet, les fibres musculaires lésées peuvent libérer du HGF ou des oxydes d'azote capables d'induire la sécrétion du HGF. Ce facteur de croissance peut entraîner les cellules satellites dans la voie de la régénération (Tatsumi *et al.*, 1998). Les FGF permettent de promouvoir la voie myogénique MAPK/p38 et ainsi d'activer les cellules satellites (Jones *et al.*, 2005). Par ailleurs, les IGF ont également un rôle au cours de la régénération dans la prolifération et la différenciation/fusion des myoblastes. Les facteurs sécrétés par les cellules immunitaires, migrant vers un muscle lésé, ont également un rôle important dans le processus de régénération (Charge and Rudnicki, 2004).

La population des cellules satellites semble être hétérogène. Les cellules satellites expriment les marqueurs Pax3 et Pax7, Pax 7 étant essentiel à la survie et au renouvellement du pool de cellules progénitrices (Relaix *et al.*, 2006). Certaines expriment également Myf5. Il semble que la cellule satellite se divise de façon asymétrique, certaines cellules filles resteront Myf5- et assureront l'auto renouvellement, d'autres vont exprimer Myf5 et être plutôt déterminées vers un destin musculaire (Kuang *et al.*, 2007). Ces cellules activées vont subir une expression séquentielle des facteurs de transcription pour se différencier, fusionner et former les nouvelles fibres ([figure 32](#)). MyoD permet le passage de cellules satellites activées en myoblastes, Pax3 est alors réprimé par ubiquitination et dégradation. Ensuite, l'activation de *Myogenin* est nécessaire à la formation de myotubes et myofibres. Myogenin inhibe

l'expression de *Pax7*. *Myf5* est également réprimé suivi de *MyoD* (revues par (Le Grand and Rudnicki, 2007; Sambasivan and Tajbakhsh, 2007)).

Plusieurs études apportent des données intéressantes concernant la division des cellules satellites pour assurer les fonctions de régénération et d'auto-renouvellement. Dans ce but, les cellules satellites se divisent de façon asymétrique et au moins trois caractéristiques ont été rapportées.

La première concerne l'ADN : en phase S, l'ADN de la cellule mère sert de matrice à la synthèse du nouveau brin d'ADN. Dans les cellules différenciées, au cours de la mitose qui suit, les brins matrices et néosynthétisés sont partagés au hasard entre les cellules filles, ce qui est visible par dilution d'un marquage BrdU réalisé en phase S. En revanche, les cellules souches ont tendance à "retenir" les brins matrices dans une cellule et séparer les brins néosynthétisés dans l'autre cellule fille, le marquage BrdU étant alors visible dans une seule des deux cellules après division. Cette particularité a été montrée pour les cellules souches musculaires, 50% d'entre elles retenant les brins matrices (Shinin *et al.*, 2006; Conboy *et al.*, 2007). Ce phénomène pourrait permettre aux cellules souches de conserver l'ADN le plus proche possible des conditions originales sans les "erreurs de copies" qui peuvent avoir lieu au cours de la réplication. Cette théorie du "brin immortel" est toutefois controversée. Une autre théorie de la "sœur silencieuse" (silent sister) décrit cette asymétrie comme due à une différence des marques épigénétiques entre les deux chromatides du chromosome mitotique (Lansdorp, 2007). En fonction des marques portées, les brins d'ADN seraient "tractés" par des fuseaux mitotiques différents et ségrégeraient donc de manière différentielle entre les deux cellules filles. La cellule fille qui a hérité des marques actives au niveau des gènes importants pour le maintien de l'état "souche" resterait dans un état plus proche de la cellule mère et permettrait ainsi l'auto-renouvellement.

La deuxième originalité est la ségrégation du facteur Numb. Cette protéine est capable d'inhiber la voie de signalisation Notch. Au cours de la division, Numb est retenu par la cellule fille qui garde le brin d'ADN matrice et conserve les caractéristiques de cellule satellite. Aussi, l'inhibition de la voie Notch pourrait être importante pour l'autorenouvellement de ces cellules souches. A l'inverse, les cellules en autorenouvellement n'exprimeront pas le facteur *Myf5* alors que les cellules déterminées pour le lignage musculaire l'exprimeront (Kuang *et al.*, 2007).

### 2-3) Myogenèse et Epigénétique

Ces dernières années, avec la croissance des connaissances en épigénétique, de nouvelles données ont montré l'implication notamment des micro-ARN et de l'Empreinte dans le processus de myogenèse.

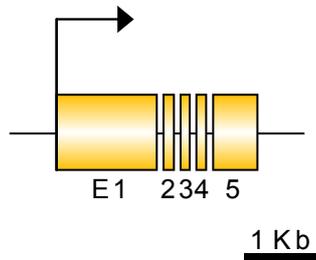
Plusieurs micro-ARN ont été décrits comme enrichis dans le muscle, notamment miR-1, miR-133, miR-181, miR-191 et miR-206 (Callis *et al.*, 2007). Dans un modèle de myoblastes murins en culture, les cellules C2C12 et dans des embryons de Xénope, il a été montré que miR-1 favorise la myogenèse alors que miR-133 a plutôt tendance à réprimer la différenciation des cellules C2C12 (Chen *et al.*, 2006). MiR-1 réprime HDAC4, une histone déacétylase qui peut inhiber la myogenèse en réprimant l'expression du facteur de transcription myogénique MEF2C (McKinsey *et al.*, 2000). MiR-133 a pour cible le facteur de transcription SRF (serum response factor), important pour la croissance et la maturation musculaires grâce à l'activation de gènes spécifiques musculaires (Li *et al.*, 2005; Charvet *et al.*, 2006). Par ailleurs SRF et MyoD possèdent des sites de fixation au niveau du locus des miR-1 et 133. Cela suggère une boucle de régulation et une fonction importante pour ces miRNA dans le développement du muscle squelettique. Enfin, il a récemment été décrit une mutation ponctuelle dans le 3'UTR de l'ARNm au niveau du gène de la *Myostatin*, chez le mouton Texel (Clou *et al.*, 2006). Cette mutation entraîne la création d'un site aberrant pour miR-1 et miR-206, qui partagent la même séquence "seed", inhibant ainsi la production de protéine Myostatin dans le muscle squelettique. La Myostatin étant un régulateur négatif de la myogenèse, cela explique le phénotype d'hypertrophie musculaire des moutons. Ces exemples illustrent bien le rôle régulateur important des miRNA musculaires dans la myogenèse.

Par ailleurs, quelques études suggèrent un rôle des gènes soumis à l'Empreinte dans la myogenèse. Tout d'abord, il a été montré que le phénotype du mouton *callipyge*, caractérisé par une hypertrophie musculaire généralisée postnatale, est du à une mutation au niveau du locus soumis à l'Empreinte *Dlk1/Gtl2* (Charlier *et al.*, 2001). La surexpression de *Dlk1* dans le muscle squelettique post-natal chez la souris reproduit le phénotype *callipyge*, ce qui suggère que la persistance de l'expression du gène paternel *Dlk1*, normalement réprimé en postnatal, est responsable de l'hypertrophie musculaire du mouton (Davis *et al.*, 2004). La régulation de l'expression de ce gène soumis à l'Empreinte est donc importante pour une myogenèse correcte.

D'autre part, une analyse par puce à ADN de gènes dont l'expression est induite ou réprimée au cours de la régénération musculaire murine a été réalisée. Parmi les 3500 gènes altérés, sept gènes soumis à l'Empreinte sont induits au cours du processus de régénération et coréglés avec les facteurs de régulations myogéniques (MRF) tels que *Myogenin* : *H19*, *Igf2*, *Cdkn1c*, *Zac1*, *Mest*, *Peg3* et *Rex3/Bex1* (Yan *et al.*, 2003). Leur induction est également abolie après irradiation, éliminant les cellules satellites et bloquant la régénération, contrairement à d'autres gènes comme *CyclinD1* ou *Csfl* qui restent induits. Cette donnée suggère un rôle important de ces gènes soumis à l'Empreinte dans l'activité des cellules satellites musculaires.

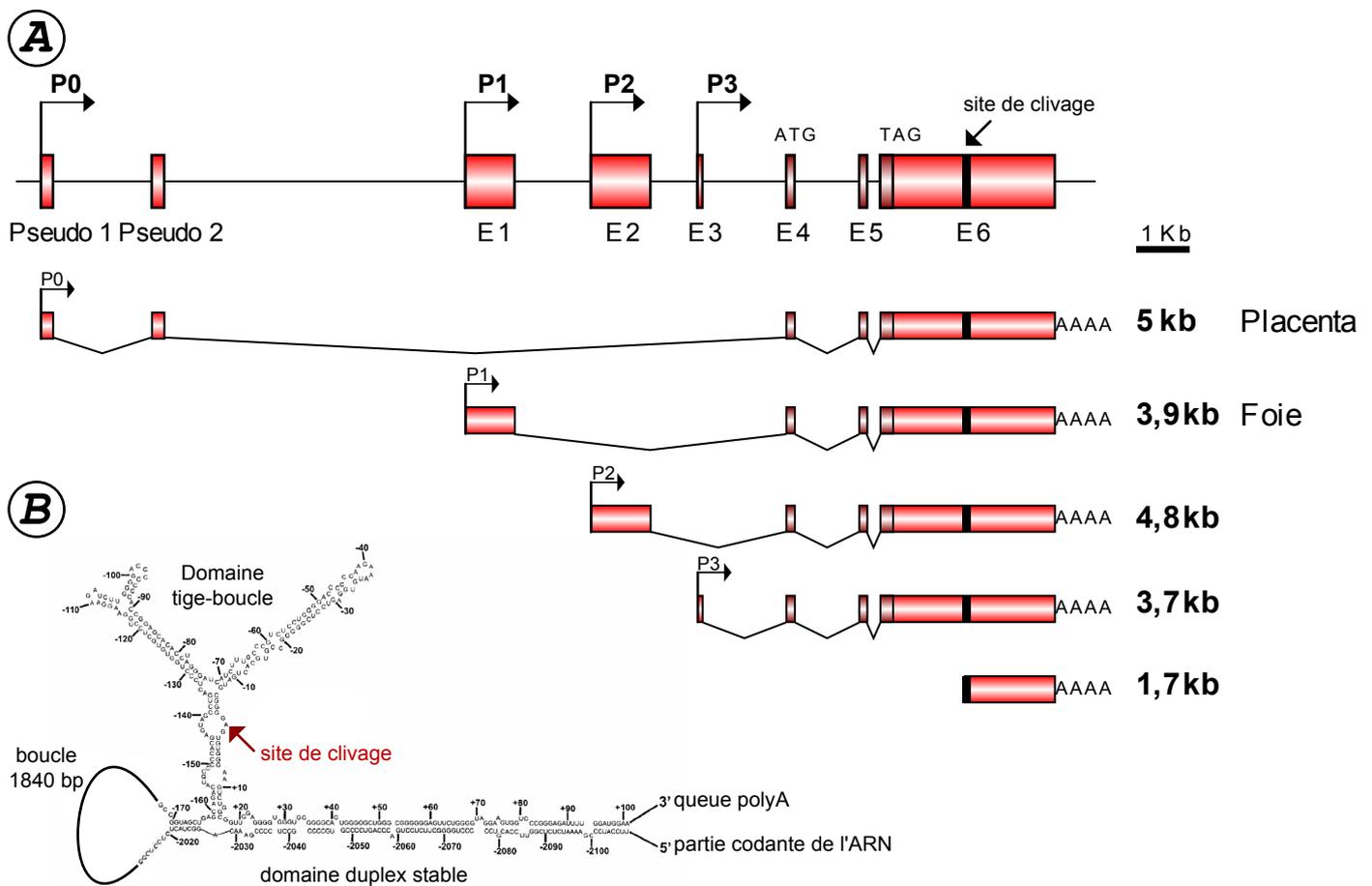
## Chapitre III

*H19 et Igf2 sont deux gènes  
étroitement liés*



**Figure 33 : Structure du gène *H19* murin.**

Le gène *H19* murin est composé de 5 exons et 4 courts introns. Il s'étend sur 3 kb et permet la production d'un ARN de 2,3 kb.



**Figure 34 : Structure du gène *Igf2* murin.**

**A.** Le gène *Igf2* murin est composé de 8 exons (nommés Pseudo 1 et 2, qui sont en fait les exons placentaires, et exons 1 à 6). La séquence codant la pro-IGF2 est située sur les exons 4 à 6 (rouge foncé). La transcription se fait à partir de quatre promoteurs alternatifs à spécificité tissulaire et les transcrits subissent un épissage alternatif. Quatre ARNm sont ainsi obtenus : P0, placentaire, P1, exprimé principalement dans le foie, P2 et P3, qui sont exprimés dans la plupart des tissus embryonnaires et placentaires.

**B.** Dans l'exon 6, au niveau du 3'UTR, deux éléments forment un duplex stable qui permet de former une structure tige-boucle. Cela permet de former une boucle qui va servir de site pour un clivage endonucléotidique spécifique. Ce clivage donne lieu à un ARN 5' instable contenant la séquence codante de l'ARN et un ARN polyA stable de 1,7 kb. (Adapté de van Dijk *et al.*, 2000)

## 1) Le produit du gène *H19* est un ARN non codant

### 1-1) La découverte du gène *H19*

Le gène *H19* a été cloné pour la première fois en 1984 dans une librairie d'ADNc de foie embryonnaire ayant un profil d'expression similaire à celui de l' $\alpha$ -fœtoprotéine (Pachnis *et al.*, 1984). Le nom de ce gène, *H19*, vient du fait que sur le filtre utilisé pour le clonage, l'ADNc correspondant était contenu dans la 19<sup>ème</sup> colonie bactérienne de la ligne H (Tilghman, 2006). Il a ensuite été décrit comme *MyoH* dans la banque d'ADNc spécifiques des myoblastes où a été isolé le gène *MyoD* (Davis *et al.*, 1987). L'ADNc *H19* fait également partie d'une librairie de gènes différentiellement exprimés entre les cellules souches embryonnaires (Embryonic Stem Cells ou ES) et les corps embryonnaires (EB pour Embryoid Bodies) (Poirier *et al.*, 1991).

### 1-2) Le gène *H19* est lié au gène *Igf2*

#### 1-2-1) Structure des gènes *H19* et *Igf2*

Le gène *H19* est situé sur la partie distale du chromosome 7 murin et composé de 5 exons séparés par de courts introns et s'étend sur 2,5 kilobases (Pachnis *et al.*, 1988) (Figure 33). Le gène est transcrit par l'ARN polymérase II, l'ARN est coiffé, polyadénylé et épissé et l'ARN mature a une taille de 2,3 kb (Brannan *et al.*, 1990). Il est exprimé uniquement à partir de l'allèle maternel chez la souris et chez l'humain (Bartolomei *et al.*, 1991; Zhang and Tycko, 1992).

Sur le chromosome 7 murin, *H19* est physiquement lié au gène *Igf2* (Zemel *et al.*, 1992). L'expression du facteur de croissance embryonnaire *Igf2* est très finement régulée sur le plan transcriptionnel : il possède quatre promoteurs alternatifs et subit un épissage alternatif (Figure 34). Les quatre transcrits produisent toutefois le même polypeptide puisque la séquence codante est située dans les exons communs 4 à 6 (Moore *et al.*, 1997). Les ARNm subissent également une régulation post-transcriptionnelle, un site de clivage endonucléotidique spécifique étant localisé dans le 3'UTR commun à tous les transcrits. Ce clivage fait intervenir des séquences conservées sur l'ARNm qui permettent la formation d'une structure tige-boucle complexe avec un site de clivage faisant intervenir des appariement non Watson-Crick (van Dijk *et al.*, 2000). Ce mécanisme produit un ARN non

Stade	Expression du gène <i>H19</i>	Expression du gène <i>Igf2</i>
Oeuf fécondé (E0.5)	Non détecté	Non détecté
Morula (E2.5)	Non détecté	Non détecté
Blastula précoce (E3.5)	Non détecté	Traces
Blastula agée (E4.5)	Trophectoderme	Traces
Post-implantation (E5.5)	Cône ectoplacentaire Endoderme et ectoderme extra-embryonnaires <i>Biallélique</i>	Cône ectoplacentaire Endoderme et ectoderme extra-embryonnaires <i>Biallélique</i>
Gastrulation (E6.5)	Endoderme et ectoderme extra-embryonnaires <i>Biallélique</i>	Endoderme et ectoderme extra-embryonnaires <i>Biallélique</i>
A partir de E8.5	Annexes extra embryonnaires <i>Préférentiellement maternelle</i> Endoderme et mésoderme embryonnaires <i>Monoallélique maternelle</i>	Annexes extra embryonnaires <i>Monoallélique paternelle</i> Endoderme et mésoderme embryonnaires <i>Monoallélique paternelle</i>
Profil à E13.5	 Annexes extra embryonnaires <i>Monoallélique maternelle</i> Endoderme et mésoderme embryonnaires Plexus choroïde et leptoméninges <i>Monoallélique maternelle</i>	 Annexes extra embryonnaires <i>Monoallélique paternelle</i> Endoderme et mésoderme embryonnaires <i>Monoallélique paternelle</i> Plexus choroïde et leptoméninges <i>Biallélique</i>
Post natal	Répression rapide dans la plupart des tissus	Répression rapide dans la plupart des tissus
Adulte (à partir de 3 semaines)	Expression forte dans le muscle squelettique Traces dans thymus, coeur et poumons <i>Monoallélique</i>	Expression <i>biallélique</i> et restreinte au plexus choroïde

**Tableau 5 : Expression embryonnaire coordonnée des gènes *H19* et *Igf2*.**

Le gène *H19* s'exprime à partir du stade péri-implantatoire dans les annexes extra-embryonnaires (indiqué en rouge) puis dans l'embryon proprement dit (indiqué en bleu) après la gastrulation dans les tissus endodermiques et mésodermiques. Le gène *Igf2* a une expression quasiment identique. Les deux gènes sont réprimés après la naissance sauf dans le muscle squelettique pour le gène *H19* et dans le cerveau pour le gène *Igf2*. (Photos tirées de Leighton *et al*, 1995)

codant stable de 1,7 kb correspondant à l'extrémité polyadénylée du 3'UTR dont la fonction est inconnue (Nielsen and Christiansen, 1992).

### **1-2-2) L'expression des gènes *H19* et *Igf2* est étroitement coordonnée**

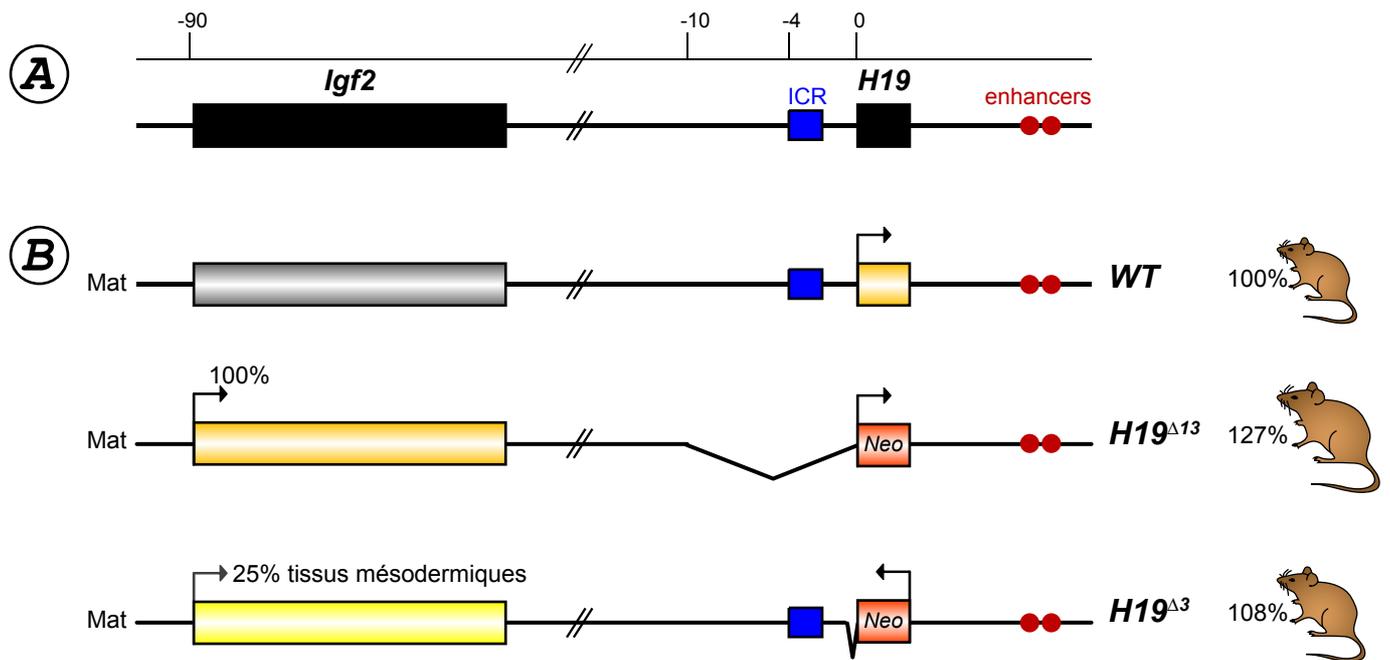
L'expression du gène *H19* est détectable à partir du stade péri-implantatoire dans les annexes extra-embryonnaires. Son expression est alors biallélique avec un biais vers l'expression maternelle qui va s'accroître tout au long de la période utérine (Szabo and Mann, 1995b). L'ARN est détecté dans l'embryon proprement dit après la gastrulation et se focalise dans les tissus dérivés de l'endoderme et du mésoderme. La transcription est réalisée à partir de l'allèle maternel. Au cours de la vie embryonnaire, l'expression du gène *H19* est très forte et comparable au niveau de la  $\beta$ -actin. Elle est ensuite fortement réprimée après la naissance dans la plupart des tissus, mais reste faiblement présente dans le thymus, le cœur et les poumons et forte dans le muscle squelettique (Poirier *et al.*, 1991; Leighton *et al.*, 1995b; Weber *et al.*, 2001). Ce profil est comparable à celui du gène *Igf2* adjacent (Tableau 5) qui s'exprime également assez fortement dans les tissus endodermiques et mésodermiques mais à partir de l'allèle paternel et s'éteint en trois semaines après la naissance dans tous les tissus à l'exception du cerveau (Lee *et al.*, 1990).

Cette expression coordonnée et réciproque d'allèle s'explique par la compétition des enhancers pour les deux gènes et le système d'insulation de CTCF (Figure 23). La seule exception connue est l'expression dans le cerveau au niveau du plexus choroïde et des leptoméninges. Dans ces tissus, la transcription du gène *H19* est monoallélique maternelle mais celle du gène *Igf2* est biallélique (DeChiara *et al.*, 1991; Svensson *et al.*, 1995). Cela peut être dû à des enhancers spécifiques de ce tissu et du gène *Igf2*, décrits dans une région conservée située entre 30 et 40 kb en amont du gène *H19* (Charalambous *et al.*, 2004).

### **1-3) Fonctions proposées pour le gène et l'ARN *H19***

#### **1-3-1) Le produit du gène *H19* est un ARN non traduit**

Rapidement après le clonage du gène, il est apparu que le produit de *H19* était un ARN non codant. Ce gène, particulièrement sa structure exons-introns, est conservé chez le poulet, le kangourou, l'opossum et les mammifères placentaires (Smits *et al.*, 2008). La comparaison des séquences murines et humaines a montré en revanche qu'aucun cadre ouvert



**Figure 35 : Modèles murins d'inactivation du gène *H19*.**

**A.** Structure du locus *H19/Igf2*. Les distances sont indiquées en kilobases.

**B.** Sur un allèle maternel normal (WT pour Wild type), le gène *Igf2* ne s'exprime pas. Les souris *H19*<sup>Δ13</sup> sont invalidées pour le gène *H19* et les 10 kb en amont, qui comprennent le centre d'Empreinte (ICR). Le gène *Igf2* s'exprime alors de façon totalement biallélique dans les tissus endodermiques et mésodermiques. Les souris *H19*<sup>Δ13</sup> ont une surcroissance de 27% par rapport aux souris sauvages. Les souris *H19*<sup>Δ3</sup> sont invalidées pour les 3 kb du gène *H19*, l'ICR est préservé. Le gène *Igf2* est réactivé sur l'allèle maternel dans les tissus mésodermiques à un taux de 25% par rapport à l'allèle paternel d'une souris sauvage. Les souris *H19*<sup>Δ3</sup> ont une surcroissance de 8% par rapport aux souris sauvages.

de lecture (ORF pour open reading frames) n'est conservé (Brannan *et al.*, 1990).

L'ARN *H19* est majoritairement cytoplasmique et pourrait être associé aux polysomes (Li *et al.*, 1998). Sa structure prédite par bioinformatique semble complexe avec de nombreuses tiges-boucles et conservée entre humain, rat et souris (Juan *et al.*, 2000).

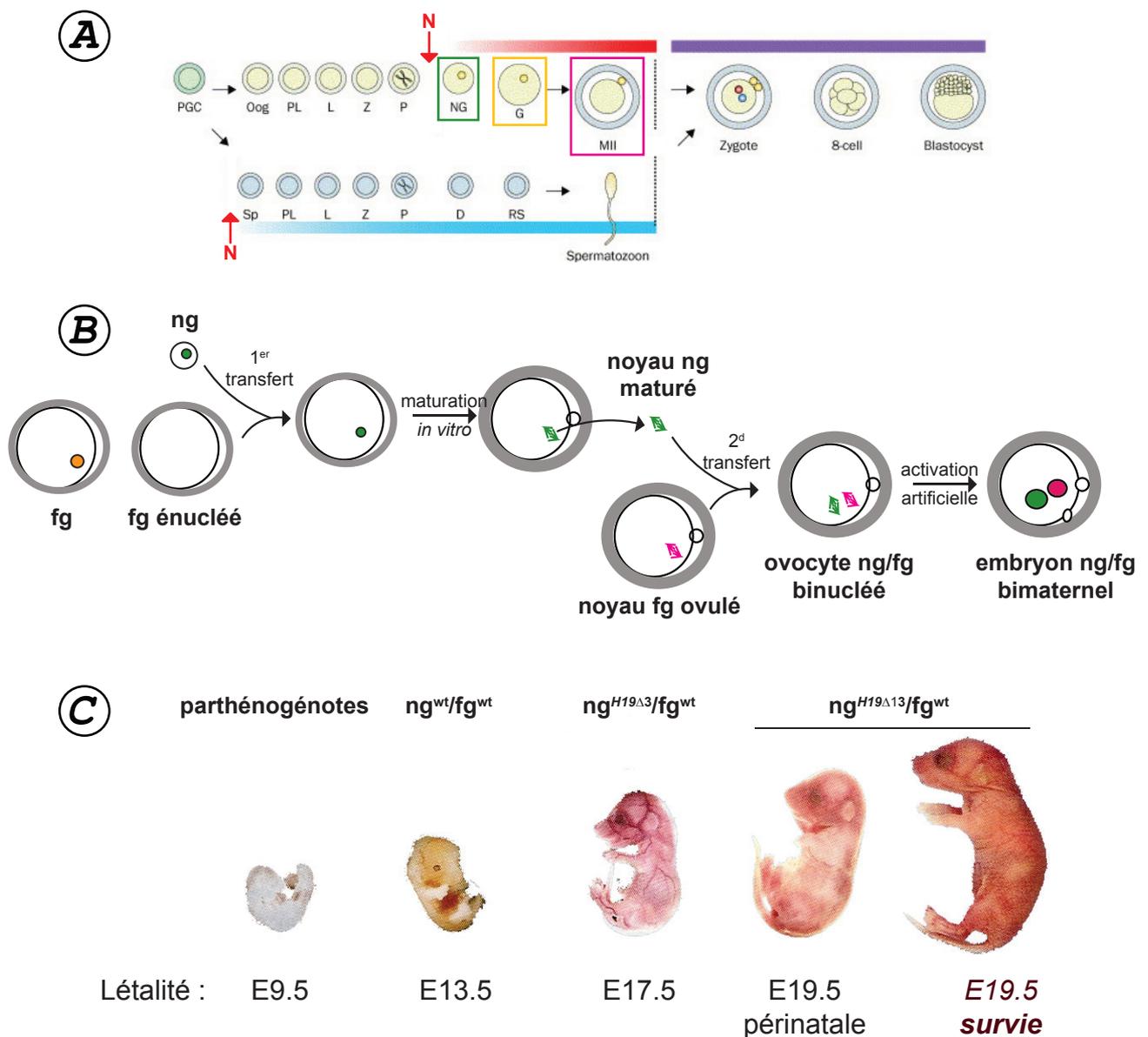
Le gène *H19* dans sa totalité partage 77% d'identité entre les séquences humaines et murines. Cependant certaines séquences semblent particulières, les exons 3, 4 et 5 ne sont conservés qu'à 72% alors que la région 3' de l'exon 1 possède 85% d'identité (Brannan *et al.*, 1990). Cette région très conservée permet la formation d'une structure tige-boucle sur l'ARN (Pfeifer and Tilghman, 1994). Récemment, un miRNA, miR-675, a été décrit dans l'exon 1 du gène *H19* chez la souris et l'humain (Mineno *et al.*, 2006; Cai and Cullen, 2007) et correspond à la tige-boucle précédemment décrite. Ce micro ARN est conservé chez le kangourou et l'opossum, l'apparition de miR-675 remonte donc à plus de 148 millions d'années, ce qui suggère qu'il pourrait avoir un rôle important et conservé (Smits *et al.*, 2008).

**L'ARN non codant pourrait ainsi avoir une fonction sous sa forme ARN entier de 2,3 kb épissé et une autre sous la forme de miR-675.** Le rôle de l'ARN non codant reste encore inconnu et aucune cible du miRNA n'est encore décrite.

### **1-3-2) Modèles murins d'inactivation du gène *H19***

Afin d'élucider le rôle du gène *H19*, des modèles murins d'inactivation du gène par recombinaison homologue ont été produits. Dans le premier modèle, le gène *H19* a été remplacé par le gène de résistance à la néomycine (*Neo*) et porte une délétion des 10 kb en amont du gène (Figure 35). Ce modèle présente donc 13 kb modifiés par rapport à l'allèle sauvage (wt) et porte la référence *H19*<sup>Δ13</sup> (Leighton *et al.*, 1995a). Le second modèle est modifié sur les 3 kb du gène *H19* et porte la référence *H19*<sup>Δ3</sup>. Chez ces souris, la cassette *Neo* remplace le gène *H19* et une petite partie du promoteur (50 bp) est déléetée (Ripoche *et al.*, 1997).

Bien que le gène *H19* soit fortement exprimé au cours de l'embryogenèse, son rôle au cours du développement n'est pas fondamental car les souris knock-out sont viables et fertiles. Elles présentent en revanche une surcroissance somatique différente selon le modèle : les



**Figure 36 : Production expérimentale d'embryons parthénogénétiques.**

**A.** Etablissement de la méthylation des gènes soumis à l'Empreinte au cours de l'ovogenèse (rouge) et de la spermatogenèse (bleu). Les flèches rouges indiquent la naissance. Un ovocyte naïf (ng pour non-growing) est encadré en vert, il est dépourvu de méthylation d'Empreinte et va ensuite subir une phase de croissance. L'ovocyte mature en fin de croissance (fg pour fully-grown) est encadré en jaune. L'ovocyte ovulatoire au stade métaphase II est encadré en rose.

**B.** Production d'embryon bimaternel à partir de noyau d'ovocyte naïf non méthylé (ng) et d'ovocyte ovulé (fg). L'embryon bimaternel parthénogénétique se développe après activation de l'ovocyte binucléé.

**C.** Phénotype des embryons parthénogénétiques. La létalité est retardée lorsque les gènes soumis à l'Empreinte sont exprimés de façon plus appropriée grâce à l'utilisation de noyaux naïfs ng obtenus à partir des mutants du gène *H19*.

(Adapté de Gosden *et al.*, 2003 et Kono *et al.*, 2006)

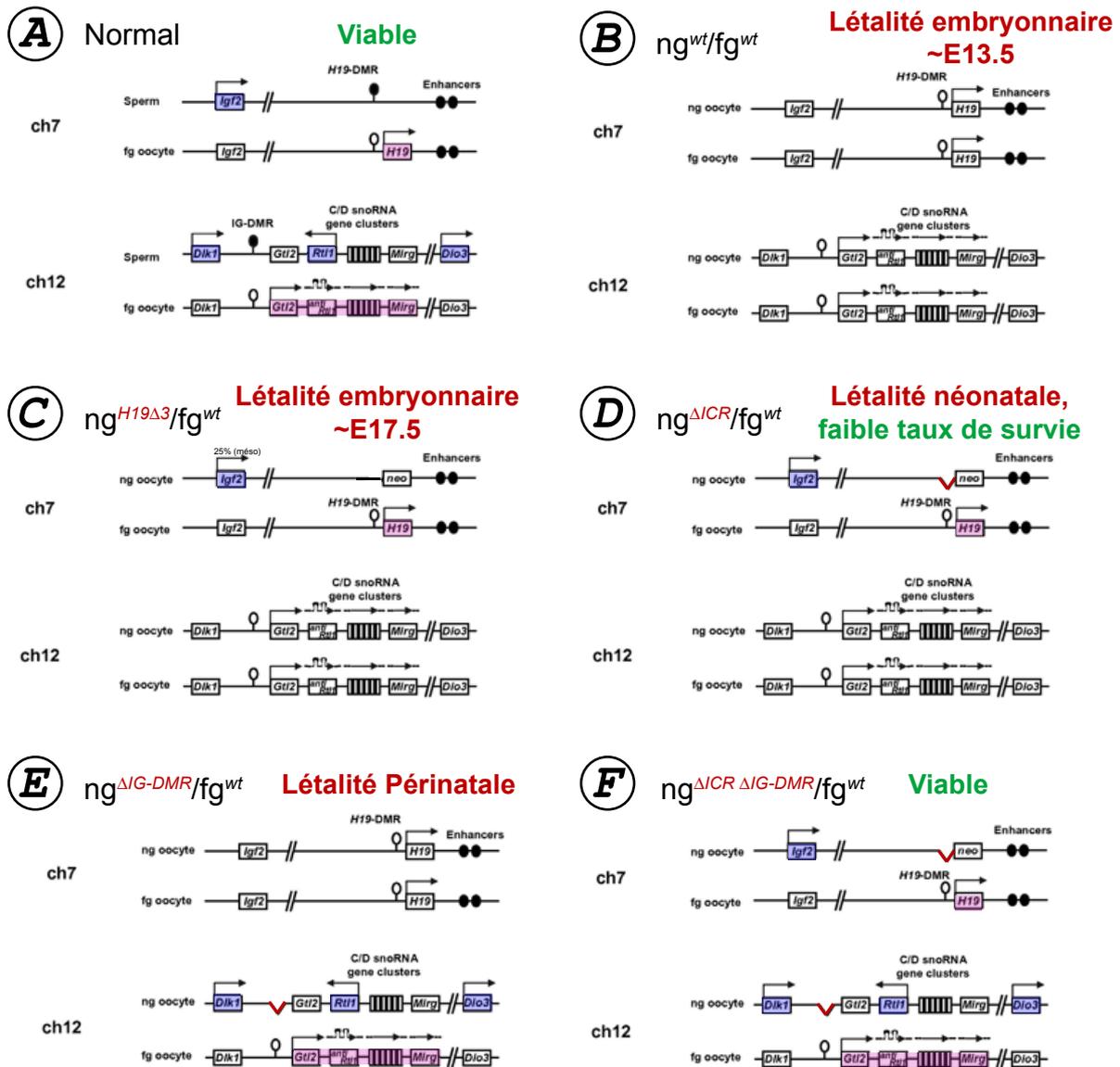
*H19*<sup>Δ13</sup> sont près de 30% plus grandes que les sauvages, les souris *H19*<sup>Δ3</sup> ont une croissance accrue de 8%. Ce phénotype est visible dès la naissance et se conserve tout au long de la vie.

La surcroissance des animaux délétés pour le gène *H19* s'explique par la réactivation de l'allèle maternel du gène du facteur de croissance *Igf2* qui est normalement l'allèle silencieux. Cette réactivation est de 100%, par rapport à l'allèle paternel d'une souris sauvage, et dans tous les tissus exprimant le gène *Igf2*, pour les souris *H19*<sup>Δ13</sup> (Leighton *et al.*, 1995a). Chez les animaux *H19*<sup>Δ3</sup>, la réactivation est partielle, de 25%, et restreinte aux tissus musculaires striés (Ripoche *et al.*, 1997; Yoshimizu *et al.*, 2008). De ces données est née l'idée d'une séquence de contrôle de l'Empreinte du gène *Igf2* dans les 10 kb en amont du gène *H19*. En fait, les souris *H19*<sup>Δ13</sup> sont délétées pour le centre d'Empreinte (ICR), situé entre 2 et 4 kb en amont du gène *H19*, et ont donc une perte d'Empreinte du gène *Igf2* dans tout l'organisme murin (Tremblay *et al.*, 1995; Hark *et al.*, 2000).

La réactivation de 25% dans les tissus musculaires des souris *H19*<sup>Δ3</sup> n'est pas encore expliquée mais il pourrait s'agir d'une modification de l'architecture de la chromatine qui permettrait le dialogue entre les promoteurs du gène *Igf2* et des enhancers mésodermiques.

### **1-3-3) Souris parthénogénétiques et locus *H19/Igf2***

La notion d'Empreinte est apparue avec les expériences montrant que la parthénogenèse est impossible chez les mammifères. L'équipe de Tomohiro Kono a pourtant réussi à obtenir des souris parthénogénètes grâce à des techniques d'embryologie expérimentale complexes (Figure 36). En partant du principe que la méthylation est une marque importante d'Empreinte établie au cours de la gamétogenèse et que la plupart des IC sont méthylés au cours de l'ovogenèse, ils ont produit des ovocytes binucléés avec un noyau méthylé et l'autre non. Cet ovocyte mimerait ainsi un noyau marqué de l'Empreinte maternelle (fg) et un noyau sans Empreinte (ng). Pour cela, un ovocyte naïf (ng pour non-growing) est prélevé chez une souris nouveau-née. Il est introduit dans un ovocyte mature (fg pour fully-grown) au stade vésicule germinale et énucléé. Les ovocytes ng et fg sont fusionnés et l'ovocyte obtenu est maturé *in vitro*. Le noyau de cet ovocyte est prélevé et transplanté dans un ovocyte fg fraîchement ovulé pour obtenir un ovocyte binucléé avec un noyau naïf et un noyau mature. L'ovocyte est activé avec du chlorure de strontium et cultivé



**Figure 37 : Résumé des travaux sur les embryons bi-maternels.**

Les rectangles roses indiquent une expression appropriée des gènes normalement exprimés à partir de l'allèle maternel, les rectangles bleus indiquent une expression appropriée des gènes paternels. Les délétions des centres d'Empreinte sont indiquées par des traits rouges.

**A. Embryon normal.** Expression monoallélique des gènes *Igf2* et *H19* et des gènes *Dlk1*, *Rtl1*, *Dio3* et des ARN non codants maternels. Génotype viable.

**B. Embryon bimaternel à partir d'un ovocyte ng wt.** Expression biallélique des ARN non codants, répression des gènes codant des protéines aux loci *H19/Igf2* et *Dlk1/Gtl2/Dio3*. Génotype léthal à E13.5.

**C. Embryon bimaternel à partir d'un ovocyte ng  $H19\Delta 3$ .** Expression appropriée au locus *H19/Igf2*, mais dérégulation au locus *Dlk1/Gtl2/Dio3*. Génotype léthal à E17.5.

**D. Embryon bimaternel à partir d'un ovocyte ng  $\Delta ICR$  *H19/Igf2*.** Expression appropriée au locus *H19/Igf2*, mais dérégulation au locus *Dlk1/Gtl2/Dio3*. Génotype à létalité néonatale avec faible taux de survie.

**E. Embryon bimaternel à partir d'un ovocyte ng  $\Delta IG-DMR$  *Dlk1/Gtl2/Dio3*.** Expression appropriée au locus *Dlk1/Gtl2/Dio3*, mais dérégulation au locus *H19/Igf2*. Génotype à létalité périnatale.

**F. Embryon bimaternel à partir d'un ovocyte ng  $\Delta ICR$   $\Delta IG-DMR$ .** Expression appropriée aux loci *H19/Igf2* et *Dlk1/Gtl2/Dio3*. Génotype viable.

(Adapté de Kawahara *et al.* 2007).

*in vitro* pendant 3.5 jours avant d'être transplanté au stade morula/blastocyste dans des femelles pseudogestantes (Kono, 2006).

Les embryons parthénogénotes et gynogénotes meurent autour de E9.5 (McGrath and Solter, 1984; Surani *et al.*, 1984). Les embryons ng/fg ont un développement prolongé et se résorbent à E13.5 (Kono *et al.*, 1996). Ces embryons par rapport aux parthénotes ont une expression appropriée des gènes soumis à l'Empreinte *Mest*, *Peg3*, *Snrpn*, *Igf2R* et *Cdkn1c*. En revanche, le gène *H19* est exprimé de façon biallélique et le gène *Igf2* est réprimé (Obata *et al.*, 1998) (Figure 37). Ces expériences ont suggéré que **le locus *H19/Igf2* pouvait jouer un rôle important dans l'arrêt de développement** des embryons parthénogénotes.

Aussi, l'équipe de Kono a utilisé les mutants *H19*<sup>Δ3</sup> et *H19*<sup>Δ13</sup> comme source de noyau ng afin de rétablir une expression monoallélique du gène *H19* et une expression du gène *Igf2*. Les souris ng<sup>*H19*<sup>Δ3</sup>/fg<sup>wt</sup></sup> ont un développement prolongé jusqu'à E17.5 (Kono *et al.*, 2002), ce qui confirme le rôle de ces gènes dans l'arrêt de développement. A partir des ovocytes ng<sup>*H19*<sup>Δ13</sup>/fg<sup>wt</sup></sup>, deux nouveaux-nés vivants ont pu être obtenus (Kono *et al.*, 2004). L'analyse de l'expression des gènes chez un nouveau-né vivant comparé à des souris sauvages montre une expression appropriée des gènes soumis à l'Empreinte. Ces expériences montrent **l'importance du locus *H19/Igf2* dans le contrôle du développement parthénogénétique**.

Par ailleurs les animaux ng<sup>*H19*<sup>Δ13</sup>/fg<sup>wt</sup></sup> qui n'ont pas survécu présentent une répression du gène *Dlk1* et la surexpression du gène *Gil2*. Aussi, les mêmes expériences ont été reconduites avec des embryons ng<sup>ΔIG-DMR/fg<sup>wt</sup></sup> pour lesquels le noyau ng porte une délétion du centre d'Empreinte du locus *Dlk1* (cet ICR est appelé IG-DMR) (Figure 37). Ces embryons survivent jusqu'à la naissance mais meurent par incapacité respiratoire (Kawahara *et al.*, 2007a). En revanche, les embryons ng<sup>ΔICR ΔIG-DMR/fg<sup>wt</sup></sup>, dont le noyau ng est délété à la fois pour le centre d'Empreinte des loci *H19/Igf2* (ICR) et *Dlk1* (IG-DMR), sont viables et le taux d'obtention est équivalent à celui d'embryons obtenus par fécondation *in vitro* (Kawahara *et al.*, 2007b).

L'ensemble de ces résultats suggèrent un rôle à la fois des loci *H19/Igf2* et *Dlk1* dans le contrôle de la parthénogenèse. La non-viabilité des souris ng<sup>ΔIG-DMR/fg<sup>wt</sup></sup> peut s'expliquer par l'absence d'expression du gène *Igf2* ou l'expression biallélique du gène *H19*. Or les souris

*Igf2*<sup>+/-</sup> sont viables et fertiles, ce qui suggère que la surexpression de l'ARN *H19* pourrait être un évènement important de la létalité (Kawahara *et al.*, 2007a).

### **1-3-4) Le gène *H19* : un gène suppresseur de tumeur ?**

Le locus 11p15.5 humain est impliqué dans les pathologies humaines de Beckwith-Wiedemann (voir point 2 de ce chapitre) et dans les tumeurs de Wilms, qui sont des tumeurs rénales. Ces observations ont permis d'émettre l'hypothèse de la présence d'un gène suppresseur de tumeur à ce locus. Parmi les candidats, l'ARN *H19* de fonction inconnue a été extensivement étudié. L'activation du gène dans plusieurs types de tumeurs du sein, de la vessie ou des poumons a été rapportée, ouvrant une controverse sur le rôle oncogénique ou suppresseur de tumeur de l'ARN *H19*.

En faveur du **rôle suppresseur de tumeur**, les premières expériences menées sur des cultures de cellules issues de tumeurs embryonnaires transfectées par un vecteur d'expression d'*H19* ont montré son rôle dans le ralentissement de la croissance des cellules tumorales. De plus, ces cellules transfectées perdent leur capacité de clonage en agar et de tumorigénicité dans les souris *nude* (Hao *et al.*, 1993). Par ailleurs, l'expression du gène *H19* est souvent réduite ou abrogée dans les tumeurs de Wilms (Moulton *et al.*, 1994). Ce phénomène est un évènement précoce puisque les gènes *H19* et *Igf2* sont exprimés dans les cellules interstitielles des reins de patients atteints de tumeurs de Wilms mais l'expression de l'ARN *H19* est perdue dès le stade néphrogénique (Cui *et al.*, 1997).

D'autres équipes ont montré la réactivation du gène *H19* dans les cancers du sein, des poumons ou de la vessie (Berteaux *et al.*, 2005). Par ailleurs, l'oncogène *c-Myc* est capable de se fixer au promoteur du gène pour induire l'expression de l'ARN *H19* dans des cellules de lignées tumorales de poumons et de sein. Un knock-down de l'ARN *H19* dans ces lignées inhibe la capacité de clonage en agar, suggérant un rôle de l'ARN dans la croissance cellulaire sans ancrage (Barsyte-Lovejoy *et al.*, 2006). Ces résultats suggèrent un **rôle oncogénique** pour le gène *H19*.

Dans notre laboratoire, nous avons très récemment démontré le rôle suppresseur de tumeur du gène *H19* dans trois modèles murins (Yoshimizu *et al.*, 2008). Cet article vous sera présenté dans la partie résultats de ma thèse, qui va suivre.

## 2) Pathologies humaines associées aux gènes *H19/Igf2*

### 2-1) Le syndrome de Beckwith-Wiedemann

Le syndrome de Beckwith-Wiedemann (**BWS**, OMIM 130650) est une maladie de surcroissance fœtale associée à une prédisposition au développement de tumeurs embryonnaires. Les signes cliniques de cette pathologie sont hétérogènes et les plus fréquents sont une macroglossie (95%), une macrosomie (80%), des anomalies de fermeture de la paroi abdominale (65%) et des viscéromégalias (50%). On trouve également une hypoglycémie post-natale, des anomalies des oreilles et une hémihypertrophie dans environ 30 à 40% des cas. Par ailleurs, 7 à 10% des enfants atteints développent des tumeurs au cours de leurs premières années de vie, majoritairement des tumeurs rénales de Wilms mais aussi des hépatoblastomes, des tumeurs corticosurrénales, des neuroblastomes ou des rhabdomyosarcomes.

Le BWS est une maladie liée à l'Empreinte parentale, qui résulte d'une dérégulation d'un ou plusieurs gènes du locus 11p15.5. Cette région est composée de deux domaines soumis à l'Empreinte<sup>1</sup> : le premier comporte les gènes *H19* et *Igf2* (Figure 23) et le second est régulé par l'expression de l'ARNnc *LIT1* (Figure 26). Dans 60% des cas, une perte de méthylation de l'ICR2 est observée, conduisant à une expression biallélique de l'ARN régulateur *LIT1* et à une réduction de l'expression du gène de la kinase inhibitrice du cycle cellulaire *Cdkn1c*. L'ICR1 et le promoteur du gène *H19* présentent quant à eux un gain de méthylation dans 10% des cas, conduisant à la perte de l'ARN *H19* et à l'expression biallélique du gène *Igf2*. Parfois, une LOI du gène *Igf2* est trouvée sans modification de l'expression du gène *H19*. Dans 20% des cas, une unidisomie paternelle (patUPD pour paternal uniparental disomy) est constatée et conduit à la surexpression du gène *Igf2* et à la répression du gène *Cdkn1c*. Dans 5% des BWS, la pathologie est associée à une mutation causant une perte de fonction de la protéine CDKN1C. Enfin dans moins de 2% des cas, des anomalies cytogénétiques (inversion, trisomie partielle ...) sont présentes (revues par Enklaar *et al.*, 2006).

Pour la plupart des malades, le syndrome de Beckwith-Wiedemann semble associé à la réduction de l'expression du gène antiprolifératif *Cdkn1c* et à la surexpression du facteur de

---

<sup>1</sup> Pour ce locus 11p15.5, les ICR sont nommés ainsi : l'ICR1 est celui qui contrôle les gènes *H19/Igf2* et l'ICR2 est celui du domaine *Kcnq1/Lit1*.

croissance IGF2. Le modèle murin du BWS associe d'ailleurs la perte de fonction de *Cdkn1c* (également appelé *p57<sup>KIP2</sup>*) avec la LOI du gène *Igf2*, par croisements entre les souris *p57<sup>KIP2</sup>-/+* et *H19<sup>Δ13/+</sup>* (Casparly *et al.*, 1999). Les souris doubles hétérozygotes maternelles n'expriment pas les gènes *Cdkn1c* et *H19* et surexpriment le gène *Igf2*. Elles présentent une macrosomie, une macroglossie et des malformations de la paroi abdominale, qui sont les caractéristiques principales du BWS.

Le BWS est associé à une prédisposition au cancer. De nombreuses études ont été développées pour établir un lien entre les défauts moléculaires des patients et le risque de tumeur, afin d'assurer un meilleur suivi médical. Le risque de tumeur est maximal en cas de défaut de méthylation de l'ICR1 (~30%) et de patUPD (~25%). Le développement de cancer est beaucoup moins fréquent chez les patients présentant une mutation du gène *Cdkn1c* (~10%) ou un défaut de méthylation de l'ICR2 (<5%), ces groupes de patients ne développant par ailleurs jamais de tumeurs de Wilms (Cooper *et al.*, 2005). Aussi, ces résultats indiquent que le locus *H19/Igf2* est plus impliqué dans la tumorigenèse que le domaine *Cdkn1c*. Cela peut être dû au dérèglement du facteur de croissance IGF2 ou à celui d'un gène suppresseur de tumeur, comme *H19*.

## 2-2) Le syndrome de Silver-Russell

Le syndrome de Silver-Russell (**SRS**, OMIM 180860) est une maladie caractérisée par un retard de croissance intra-utérin (RCIU) avec périmètre crânien préservé, une petite stature conservée en post-natal et à l'âge adulte, une dysmorphie faciale (front bombant et bas du visage fin) et une hémihypertrophie du corps (Abu-Amero *et al.*, 2008). Les enfants atteints peuvent également présenter des difficultés alimentaires ainsi que des troubles de l'apprentissage variables.

Le SRS est en général sporadique, mais des transmissions autosomiques dominantes ou autosomiques récessives ont été rapportées. La présence d'unidisomie parentale dans ce syndrome a conduit à soupçonner que cette maladie est liée à l'Empreinte Parentale. Jusqu'en 2005, le chromosome 7 humain était très étudié puisque dans 10% des cas de SRS environ, une unidisomie maternelle de ce chromosome est présente (Eggermann *et al.*, 1997; Preece *et al.*, 1997). Ce chromosome porte deux loci soumis à l'Empreinte comprenant les gènes *Grb10*, *Sgce*, *Peg10*, *Mest* et *Copγ2*.

Le SRS étant une maladie de croissance tout comme le BWS et compte tenu du rôle crucial de la région 11p15.5 associée dans la croissance, Gicquel *et coll* ont étudié cette région dans plusieurs patients SRS. Ils ont observé une déméthylation de l'ICR du locus *H19/Igf2* (Gicquel *et al.*, 2005). Parallèlement des patients présentant une unidisomie maternelle de la région ont été génotypés (Eggermann *et al.*, 2005). Une analyse d'une grande cohorte de patient SRS a pu impliquer la déméthylation de l'ICR et du promoteur du gène *H19* dans 60% des cas de SRS (Netchine *et al.*, 2007). De plus cette analyse lie l'hypométhylation avec la sévérité phénotypique des patients. Il a été proposé que la déméthylation de l'allèle paternel de l'ICR permette la fixation de CTCF de façon biallélique. Les enhancers ne peuvent donc pas activer le gène *Igf2* qui est totalement réprimé alors que le gène *H19* s'exprime de façon biallélique. La répression du facteur de croissance IGF2 expliquerait ainsi le RCIU et la petite stature conservée en post-natal.

# Article 1

## The H19 gene: regulation and function of a non-coding RNA

Gabory, A., Ripoche, M. A., Yoshimizu, T. and Dandolo, L.  
(2006). *Cytogenet Genome Res* 113, 188-93.

## **Introduction**

Lors du congrès "Genomic Imprinting Workshop" de Montpellier, organisé à l'occasion des 20 ans de la découverte de l'Empreinte Parentale par Davor Solter et Azim Surani en 1984, nous avons été sollicitées pour participer au numéro spécial "Genomic Imprinting" de la revue *Cytogenetic and Genome Research* (vol 113, No. 1-4, 2006).

Nous avons décidé de consacrer notre article à une revue de synthèse portant sur le gène *H19*. En effet, ce gène présente un intérêt particulier puisqu'il est très abondamment transcrit au cours de l'embryogenèse, qu'il fut parmi les premiers gènes soumis à l'Empreinte découverts et que son produit fut le premier ARN non codant transcrit par l'ARN polymérase II décrit. Depuis son clonage en 1984, de nombreuses études ont été réalisées sur la régulation de l'expression et les séquences impliquées ont été étudiées grâce à de nombreux modèles murins transgéniques. Par ailleurs, la question de la fonction de l'ARN *H19* a été adressée par des modèles d'inactivations et des études *in vitro*. Il a également été très étudié dans les tumeurs humaines où il est souvent abondamment exprimé.

Le but de cette revue était de rassembler les données de la littérature sur le gène atypique qu'est *H19*, afin de proposer une vision d'ensemble de sa régulation et des pistes apportées quant à sa fonction. Cette revue a été soumise à peer-review.

# The *H19* gene: regulation and function of a non-coding RNA

A. Gabory M.-A. Ripoché T. Yoshimizu L. Dandolo

Department of Genetics and Development, Institut Cochin, INSERM U567, CNRS UMR 8104, University Paris V Descartes, Paris (France)

Manuscript received 9 July 2005; accepted in revised form for publication by F. Ishino, 14 November 2005.

**Abstract.** The *H19* gene encodes a 2.3-kb non-coding mRNA which is strongly expressed during embryogenesis. This gene belongs to an imprinted cluster, conserved on mouse chromosome 7 and human chromosome 11p15. *H19* is maternally expressed and the neighbouring *Igf2* gene is transcribed from the paternal allele. These two genes are co-expressed in endoderm- and mesoderm-derived tissues during embryonic development, which suggests a common mechanism of regulation. The regulatory elements (imprinted control region, CTCF insulation, different enhancer sequences, promoters of the two genes, matrix attachment regions) con-

fer a differential chromatin architecture to the two parental alleles leading to reciprocal expression. The role of the *H19* gene is unclear but different aspects have been proposed. *H19* influences growth by way of a *cis* control on *Igf2* expression. Although *H19*<sup>-/-</sup> mice are viable, a role for this gene during development has been suggested by viable *H19*<sup>-/-</sup> parthenogenetic mice. Finally it has been described as a putative tumour suppressor gene. *H19* has been studied by numerous laboratories over the last fifteen years, nevertheless the function of this non-coding RNA remains to be elucidated.

Copyright © 2006 S. Karger AG, Basel

The *H19-Igf2* locus has been a paradigm for the study of genomic imprinting for over fifteen years. This locus contains the *H19* gene, which is maternally expressed and the paternally expressed *Igf2* gene (insulin-like growth factor 2). These genes are 90 kb apart and define a region in which many regulatory elements have been identified, mainly using targeted germ line deletions or transgenes in the mouse. This locus belongs to a large 1 Mb cluster of imprinted genes on the distal part of mouse chromosome 7.

## *H19* RNA

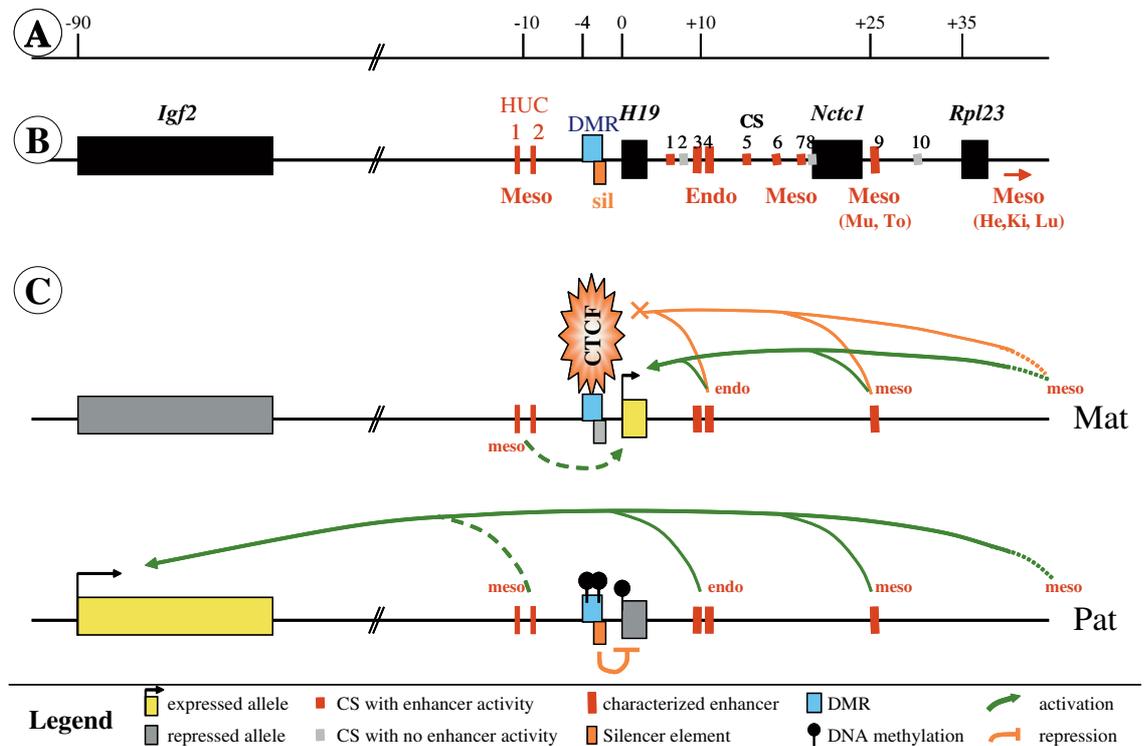
The murine *H19* gene was originally isolated in four different laboratories: i) as a member of the myogenic factors, at the same time as *MyoD* and called at that time *MyoH* (Davis et al., 1987), ii) during a search for genes coordinately regulated with the  $\alpha$ -foetoprotein gene by the *Raf* gene (Pachnis et al., 1988), iii) as an upregulated gene upon embryonal carcinoma differentiation (Wiles, 1988), and iv) as a differentially expressed gene upon screening of non-differentiated ES cells and differentiated embryoid bodies (Poirier et al., 1991).

The 2.5-kb gene is composed of five exons and four small introns and the *H19* RNA is fully capped, spliced and polyadenylated. The *H19* gene was found to be monoallelically expressed from the chromosome of maternal origin in mouse and in human and, with *Igf2*, was defined as one of the first imprinted genes (Bartolomei et al., 1991; Zhang and Tycko, 1992).

Upon comparison of the murine and human sequences, it did not appear to have any conserved long open reading frame. It was concluded that the product of this gene was an untranslated RNA (Brannan et al., 1990).

Supported by the INSERM, by an Action Concertée Incitative (ACI Biologie du Développement), by the Association de Recherche contre le Cancer (ARC) and by the AFM. A.G. is supported by a fellowship from the Ministère de la Recherche et de l'Éducation Nationale. T.Y. was supported by a post-doctoral fellowship from the Institut National pour la Santé et la Recherche Médicale (INSERM) and from the Association Française contre les Myopathies (AFM).

Request reprints from Luisa Dandolo  
Department of Genetics and Development, Institut Cochin  
INSERM U567, CNRS UMR 8104, University Paris V Descartes  
24 rue Fbg. St-Jacques, Paris 75014 (France)  
telephone: +33 1 44 41 24 55; fax: +33 1 44 41 24 21  
e-mail: [dandolo@cochin.inserm.fr](mailto:dandolo@cochin.inserm.fr)



**Fig. 1.** Common enhancers of the *H19* and *Igf2* genes. **(A)** Distances in kb from the *H19* gene transcription start. **(B)** *Igf2/H19* locus and enhancer positions. CS: human/mouse conserved sequences. CS3 and 4 are endodermal enhancers (Endo), CS9 is a mesodermal enhancer (Meso, Mu: skeletal muscle, To: tongue, He: cardiac muscle, Ki: kidney, Lu: lung). HUC: *H19* Upstream Conserved regions, HUC 1 and 2 are mesodermal enhancers. The Imprinting Control Region is composed of a differentially methylated region (DMR) and a silencer element (sil).

The DMR can bind the CTCF insulator protein and the silencer is able to repress *H19* paternal expression. **(C)** Model for *H19* and *Igf2* expression. On the paternal allele (Pat), *Igf2* is activated by the different mesodermal and endodermal enhancers but the enhancers downstream of *H19* are blocked by the CTCF insulation on the maternal allele (Mat). *H19* is activated by HUC and the downstream enhancer on the maternal allele but repressed by the silencer on the paternal allele.

### Expression profiles

The expression profile of *H19* was examined in the mouse during development; it was found by in situ hybridisation to be strongly expressed in extraembryonic tissues as early as E5.5 but not in the embryo itself before E8.5 (Poirier et al., 1991). Later, it was shown, using the more sensitive and allele-specific RT-PCR SNUPE assay, that *H19* expression could be detected soon after implantation (E6.5). It is biallelically expressed at that time and strict monoallelic *H19* expression is not detected until E10.5 (Szabó and Mann, 1995a).

*H19* is then highly expressed all through fetal development, its level being comparable to the level of  $\beta$ -actin gene expression. The adjacent *Igf2* gene is expressed in the same tissues as *H19* (Lee et al., 1990). Thus these two genes have a very tightly coordinated expression, with tissue and stage specific patterns. Mostly expressed in endoderm and mesoderm derived tissues during fetal life, they are not found in the brain or central nervous system, with the exception of leptomeninges and choroid plexus. In all organs, these two genes are reciprocally expressed, i.e. *H19* is maternally and *Igf2* paternally expressed.

After birth, both genes are strongly down-regulated in all tissues with the exception of skeletal muscle, in which *H19* continues to be expressed (Weber et al., 2001).

### Regulatory elements

In vivo systems are required to study genomic imprinting since imprints are set during gametogenesis, to be then further expressed during development or in adult life. During the past ten years, more than 35 gene targeting events and transgenes have been produced in the mouse around this locus (Arney, 2003). These mouse models have permitted identification of a large number of regulatory elements, which are involved in determining the differential expression of imprinted genes (Fig. 1).

#### *The H19 ICR, DMD or DMR*

In the 10-kb region upstream from *H19*, several elements of regulation have been identified, some having an important role and others not.

An ICR (imprinting control region) or DMD (differentially methylated domain) or DMR (differentially methylated

region) was identified between -4 and -2 kb upstream of the *H19* transcription start site (Tremblay et al., 1995, 1997). This sequence contains CpGs which are methylated on the paternal allele and fully unmethylated on the maternal allele. Differential methylation of the DMR occurs during spermatogenesis, is present in the early embryo, is maintained during the genome-wide demethylation that occurs before implantation, and is found all through development up to adulthood.

Deletion of this DMR leads to disruption of both *H19* and *Igf2* expression thus suggesting that this region is an imprinting center, capable of controlling *in cis* the expression of several genes in this chromosomal region (Thorvaldsen et al., 1998, 2002; Srivastava et al., 2000).

*In vivo* footprinting has shown that a zinc finger protein called CTCF can bind to four sites on the unmethylated maternal ICR (Bell and Felsenfeld, 2000; Hark et al., 2000; Kanduri et al., 2000; Szabó et al., 2000). The binding of CTCF creates a boundary element on the maternal allele and inhibits interaction of downstream enhancers with the *Igf2* promoters.

The *H19* gene is controlled not only by the ICR but also by a promoter, which is also differentially methylated. This promoter region is methylated *de novo* after fertilisation on the paternal chromosome and is a tissue-specific regulatory element (Bartolomei et al., 1993; Ferguson-Smith et al., 1993; Sasaki et al., 1995; Tremblay et al., 1995; Srivastava et al., 2000). It is unmethylated on the maternal chromosome. However, when the CTCF sites are mutated on the maternal allele, the ICR can no longer bind CTCF. It becomes fully methylated post-fertilisation and this leads to methylation of the *H19* promoter (perhaps through a spreading mechanism). Reduction of maternal *H19* expression is observed as well as increase of maternal *Igf2* expression (Szabó et al., 2004).

It is now established that the ICR possesses three functions: i) on the maternal chromosome, the ICR acts as an insulator and mediates silencing of the *Igf2* gene *in cis*; ii) on the paternal chromosome, it is a transcriptional silencer which represses the paternal *H19* promoter (Drewell et al., 2000); iii) the ICR is responsible for imprinting establishment with germ line differential methylation acquired during male gametogenesis. This constitutes the primary epigenetic mark responsible for imprinting of this locus.

#### Enhancers

*Igf2* and *H19* share other regulatory elements, including several spatially and temporally regulated enhancers downstream of the *H19* gene (Fig. 1).

Early on, two endodermal enhancers were identified downstream from *H19*. Their targeted deletion affects both *Igf2* and *H19* expression, therefore implying that these enhancers are common to the two genes (Yoo-Warren et al., 1988; Leighton et al., 1995b).

In a search for mesodermal enhancers, a study of sequence conservation between human and mouse identified ten conserved sequences (CS) in the 35-kb region downstream of *H19* (Ishihara et al., 2000). CS3 and 4 are actually the endodermal enhancers previously determined. A 130-kb YAC transgene extending as far as +35 kb can recapitulate *H19* imprinting

in endoderm and mesoderm derived tissues, with the exception of heart, kidney and lung (Ainscough et al., 1997, 2000). A BAC transgene extending further downstream (+120 kb) was shown to contain not only skeletal muscle but also cardiac muscle-specific mesodermal enhancers (Kaffer et al., 2000). Finally, the minute (*Mnt*) mouse mutant has lost expression of *Igf2* in mesoderm tissues and this phenotype is due to a large inversion starting at CS9 (Davies et al., 2002). Altogether these studies demonstrate that the skeletal muscle and tongue specific mesodermal enhancers must be CS9.

In summary, three sets of enhancers have been determined downstream of the *H19* gene: endoderm-specific enhancers located at +10 kb, skeletal muscle mesoderm-specific enhancers at +25 kb and other mesoderm-specific enhancers between +35 kb and +120 kb for cardiac muscle, kidney and lung.

Another couple of putative mesoderm enhancers have been identified upstream of *H19*: HUC 1 and 2 (*H19* Upstream Conserved; Drewell et al., 2002).

The *Igf2* gene seems to also possess its own mesodermal enhancers in a region described as CCD (Centrally Conserved Domain) at -32 to -44 kb from the *H19* transcription start. CCD may have a silencer function in skeletal muscle and positively regulate *Igf2* expression in brain (Ainscough et al., 2000; Jones et al., 2001; Charalambous et al., 2004).

#### Chromatin architecture

A primary imprint must be established in the germ line, then stabilised by differential methylation on the ICR and/or by histone modifications, which can be called a secondary imprint. Tissue-specific factors are finally superimposed on the basic imprinting signals to be read by the cellular transcription machinery.

Aside from the *H19* DMR, several DMRs have been described for the *Igf2* gene as well. A hierarchy of these different DMRs in the *H19-Igf2* locus has been established by bringing together numerous targeted deletions all along the locus (Lopes et al., 2003).

MAR sequences (matrix attachment regions) have also been identified in this region and contribute to the complex chromatin architecture of the locus which is starting to be unravelled (Weber et al., 2003; Purbowasito et al., 2004).

A model of chromatin loops has also been proposed, leading to distribution of the genes in active and inactive nuclear compartments (Banerjee and Smallwood, 1995; Murrell et al., 2004; Kato and Sasaki, 2005).

#### Imprinting in the germ line

Imprinting is set during gametogenesis, is maintained throughout development and in the adult somatic divisions. It is then erased in the primordial germ cells (PGCs) at the time when they enter and colonize the genital ridges (Szabó and Mann, 1995b; Hajkova et al., 2002; Lee et al., 2002; Li et al., 2004). In PGCs at E10.5, *H19* and *Igf2* are monoallelically expressed (Szabó et al., 2002). At E11.5, the postmigratory PGCs enter the gonad and the imprint is erased: the two genes become biallelically expressed. For the *H19* gene, the

imprint is then reset soon after in the male germ line: establishment of paternal-specific methylation of the DMR occurs in spermatogonia as well as in spermatocytes (Davis et al., 1999, 2000; Ueda et al., 2000). *H19* is biallelically expressed in spermatogonia, and *H19* is not detected in spermatocytes or spermatids (Szabó and Mann, 1995b). In the female germ line imprints are acquired in the postnatal growth phase at the diplotene stage of prophase I. Thus imprinting is erased in non-growing (ng) oocytes and established again in fully-grown (fg) oocytes (Obata et al., 1998, 2002; Lucifero et al., 2002).

### Function of the *H19* gene

In order to identify the function of *H19*, several targeted deletions of this gene were produced and all the mice were viable and fertile (Leighton et al., 1995a; Ripoche et al., 1997; Jones et al., 1998; Schmidt et al., 1999).

The *H19Δ13* mutant consists of the deletion of a 13-kb region covering 10 kb of the upstream region and the *H19* transcription unit; the *H19Δ3* mutation deletes only the 3-kb transcription unit. Both mutations lead to an overgrowth phenotype, 27% in the case of *H19Δ13* and 8% in the case of *H19Δ3*.

The molecular mechanism underlying this overgrowth is linked with loss of imprinting of the adjacent *Igf2* gene on the maternal chromosome. In the *H19Δ13* deletion, *Igf2* is expressed at a level equivalent to the paternal allele in most organs such as muscle and liver. In the *H19Δ3* deletion, loss of imprinting is found in muscle but not in liver and the level of expression of the maternal *Igf2* is approximately 25%. This suggested that the *H19* locus was able to control *Igf2* expression *in cis*.

Since the *H19Δ13* mutant deletes the DMR, expression of the maternal *Igf2* allele can be explained by removal of the boundary element, which allows access of the downstream enhancers to the *Igf2* promoter. The reason for maternal *Igf2* expression in the *H19Δ3* mutant is still unclear.

### Role in development

Since *H19* was highly expressed during embryogenesis, it had been expected that the deletion of this gene would lead to embryonic lethality. It was therefore a surprise to find viable *H19<sup>-/-</sup>* mice and the possible conclusion from these experiments was that *H19* did not play a role during development. However, the recent experiments of T. Kono, investigating possibilities to prolong the development of parthenogenetic embryos, strongly suggest that *H19* plays an important role in development. T. Kono performed nuclear transfer experiments using nuclei from non-growing oocytes (ng) and from fully-grown oocytes (fg) and showed that the embryos can survive to E13.5 (Kono et al., 1996). When *H19<sup>-/-</sup>* females were used as a source of ng oocytes, embryos survived to E17.5 (*H19Δ3* mutation) and a live female was born and shown to be fertile (*H19Δ13* mutation) (Kono et al., 2002, 2004). This could be due either to overexpression of

*Igf2* or to reduced expression of *H19*. The transcriptome analysis of this viable mouse shows that the situation is far from being simple.

If *H19* plays an important role in development, it is difficult to explain why no effect was seen in the targeted deletions of *H19*. This is usually explained by redundancy. However no family of *H19*-like genes has been found by sequence homology analysis. One could postulate that the *H19* RNA is associated with proteins in order to fulfill its role and that this RNA can adopt a spatial structure specific to its protein partners. Perhaps other RNAs could bind to the protein partners and this association would lead to a functional redundancy.

### Trans effect of *H19* RNA

The overgrowth phenotype of the *H19* deletion mutants could be due to several causes: it could be the deletion of the locus itself which would modify the chromatin structure of the region and change the accessibility of the *Igf2* promoter, allowing its expression from the normally silent maternal allele. Or it could be the *H19* RNA itself which could interfere in the loss of imprinting of *Igf2*. This would imply that *H19* could play a role not only *in cis* but also *in trans*. This has been suggested by several different approaches:

(i) In the *H19Δ3* and *Δ13* mutants, methylation profiles of the *Igf2* DMRs have shown a modification of the methylation levels on the opposite allele, thus suggesting a *trans* effect upon deletion of the *H19* gene (Forne et al., 1997).

(ii) The *H19* RNA was found to be associated with polyosomes and thus potentially act on the translation of *Igf2* (Li et al., 1998).

(iii) In vitro experiments showed that *H19* RNA can have a *trans* effect on the levels of *Igf2* expression (Wilkin et al., 2000).

(iv) *H19* RNA can bind to the IMP1 protein (Igf2mRNA binding-protein1) which also binds to the 5'-UTR of the *Igf2* mRNA. The bound IMP1 could interfere with post-transcriptional regulation of *Igf2* or with its translation (Runge et al., 2000; Hansen et al., 2004).

The *trans* effect of the *H19* RNA could be mediated through several different mechanisms:

(i) As it has been described in plants, this RNA could induce DNA methylation and serve as a carrier to bring DNA-methyl transferase or methyl binding proteins to specific sequences such as the DMRs of the locus.

(ii) Both *H19* and *Igf2* could compete for the same IMP1 molecules. It was shown that lack of IMP1 leads to down-regulation of *Igf2* translation (Hansen et al., 2004). Since growth control appears to be very tightly regulated at the *Igf2* locus (Sun et al., 1997), the binding of *H19* to IMP1 (reducing the number of free IMP1 molecules capable of binding to *Igf2* mRNA) could be an additional element involved in the control of the IGF2 protein level.

(iii) Finally, it cannot be excluded that the *H19* RNA may function through an miRNA or siRNA mechanism. The small RNA molecules could interfere with post-transcriptional regulation or translation of the *Igf2* mRNA.

## H19 in human disease

The genomic organisation and imprinting of the cluster in the distal chromosome 7 of the mouse is conserved in humans and corresponds to chromosome 11p15.5. This region is linked to the Beckwith-Wiedemann syndrome (BWS), an imprinting disorder with specific features of postnatal overgrowth, macroglossia, anterior abdominal wall defects and an increased susceptibility to tumourigenesis. This 1-Mb region contains two independent domains, one containing the *H19-IGF2* locus and the other a large cluster of genes, controlled by the DMR of *KCNQ1OT1 (LIT1)*. It has been suggested that because of the higher incidence of tumours, such as Wilms' tumour in particular, a tumour suppressor gene must be located in this region (Maher and Reik, 2000). Several data point to *H19* as being the tumour suppressor gene (Hao et al., 1993; Juan et al., 2000). An extensive study of Wilms' cases has shown that *IGF2* is biallelically expressed in these tu-

mours and *H19* fully repressed, with the *H19* DMR being hypermethylated on both alleles (Dao et al., 1999; Frevel et al., 1999).

It was also recently shown that microdeletions in the *H19* DMR of BWS patients lead to loss of imprinting with biallelic expression of *Igf2* and silencing of *H19* (Sparago et al., 2004; Prawitt et al., 2005).

## Conclusion

The *H19-Igf2* locus has brought into the light many aspects of regulatory mechanisms linked to monoallelic expression of these two genes, as well as knowledge of the establishment and maintenance of primary imprinting marks and new insights into the chromatin architecture in the nucleus. However the question of the role of the *H19* gene and its abundant non-coding RNA still remains open.

## References

- Ainscough JF, Koide T, Tada M, Barton S, Surani MA: Imprinting of *Igf2* and *H19* from a 130-kb YAC transgene. *Development* 124:3621–3632 (1997).
- Ainscough JF, Dandolo L, Surani MA: Appropriate expression of the mouse *H19* gene utilises three or more distinct enhancer regions spread over more than 130 kb. *Mech Dev* 91:365–368 (2000).
- Arney KL: *H19* and *Igf2* – enhancing the confusion? *Trends Genet* 19:17–23 (2003).
- Banerjee S, Smallwood A: A chromatin model of *IGF2/H19* imprinting. *Nat Genet* 11:237–238 (1995).
- Bartolomei MS, Zemel S, Tilghman SM: Parental imprinting of the mouse *H19* gene. *Nature* 351:153–155 (1991).
- Bartolomei MS, Webber AL, Brunkow ME, Tilghman SM: Epigenetic mechanisms underlying the imprinting of the mouse *H19* gene. *Genes Dev* 7:1663–1673 (1993).
- Bell AC, Felsenfeld G: Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* 405:482–485 (2000).
- Brannan CI, Dees EC, Ingram RS, Tilghman SM: The product of the *H19* gene may function as an RNA. *Mol Cell Biol* 10:28–36 (1990).
- Charalambous M, Menheniott TR, Bennett WR, Kelly SM, Dell G, Dandolo L, Ward A: An enhancer element at the *Igf2/H19* locus drives gene expression in both imprinted and non-imprinted tissues. *Dev Biol* 271:488–497 (2004).
- Dao D, Walsh CP, Yuan L, Gorelov D, Feng L, Hensle T, Nisen PY, Yamashiro DJ, Bestor TH, Tycko B: Multipoint analysis of human chromosome 11p15/mouse distal chromosome 7: inclusion of *H19/IGF2* in the minimal WT2 region, gene specificity of H19 silencing in Wilms' tumorigenesis and methylation hyper-dependence of H19 imprinting. *Hum Mol Genet* 8:1337–1352 (1999).
- Davies K, Bowden L, Smith P, Dean W, Hill D, Furuumi H, Sasaki H, Cattanaach B, Reik W: Disruption of mesodermal enhancers for *Igf2* in the *minute* mutant. *Development* 129:1657–1668 (2002).
- Davis RL, Weintraub H, Lassar AB: Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* 51:987–1000 (1987).
- Davis TL, Trasler JM, Moss SB, Yang GJ, Bartolomei MS: Acquisition of the *H19* methylation imprint occurs differentially on the parental alleles during spermatogenesis. *Genomics* 58:18–28 (1999).
- Davis TL, Yang GJ, McCarrey JR, Bartolomei MS: The H19 methylation imprint is erased and re-established differentially on the parental alleles during male germ cell development. *Hum Mol Genet* 9:2885–2894 (2000).
- Drewell RA, Brenton JD, Ainscough JF, Barton SC, Hilton KJ, Arney KL, Dandolo L, Surani MA: Deletion of a silencer element disrupts *H19* imprinting independently of a DNA methylation epigenetic switch. *Development* 127:3419–3428 (2000).
- Drewell RA, Arney KL, Arima T, Barton SC, Brenton JD, Surani MA: Novel conserved elements upstream of the *H19* gene are transcribed and act as mesodermal enhancers. *Development* 129:1205–1213 (2002).
- Ferguson-Smith AC, Sasaki H, Cattanaach BM, Surani MA: Parental-origin-specific epigenetic modifications of the mouse H19 gene. *Nature* 362:751–755 (1993).
- Forne T, Oswald J, Dean W, Saam JR, Bailleul B, Dandolo L, Tilghman SM, Walter J, Reik W: Loss of the maternal *H19* gene induces changes in *Igf2* methylation in both *cis* and *trans*. *Proc Natl Acad Sci USA* 94:10243–10248 (1997).
- Frevel MA, Sowerby SJ, Petersen GB, Reeve AE: Methylation sequencing analysis refines the region of *H19* epimutation in Wilms' tumor. *J Biol Chem* 274:29331–29340 (1999).
- Hajkova P, Erhardt S, Lane N, Haaf T, El-Maarri O, Reik W, Walter J, Surani MA: Epigenetic reprogramming in mouse primordial germ cells. *Mech Dev* 117:15–23 (2002).
- Hansen TV, Hammer NA, Nielsen J, Madsen M, Dalbaeck C, Wewer UM, Christiansen J, Nielsen FC: Dwarfism and impaired gut development in insulin-like growth factor II mRNA-binding protein 1-deficient mice. *Mol Cell Biol* 24:4448–4464 (2004).
- Hao Y, Crenshaw T, Moulton T, Newcomb E, Tycko B: Tumour-suppressor activity of H19 RNA. *Nature* 365:764–767 (1993).
- Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM: CTCF mediates methylation-sensitive enhancer-blocking activity at the *H19/Igf2* locus. *Nature* 405:486–489 (2000).
- Ishihara K, Hatano N, Furuumi H, Kato R, Iwaki T, Miura K, Jinno Y, Sasaki H: Comparative genomic sequencing identifies novel tissue-specific enhancers and sequence elements for methylation-sensitive factors implicated in *Igf2/H19* imprinting. *Genome Res* 10:664–671 (2000).
- Jones BK, Levorse JM, Tilghman SM: *Igf2* imprinting does not require its own DNA methylation or *H19* RNA. *Genes Dev* 12:2200–2207 (1998).
- Jones BK, Levorse J, Tilghman SM: Deletion of a nuclease-sensitive region between the *Igf2* and *H19* genes leads to *Igf2* misregulation and increased adiposity. *Hum Mol Genet* 10:807–814 (2001).
- Juan V, Crain C, Wilson C: Evidence for evolutionarily conserved secondary structure in the H19 tumor suppressor RNA. *Nucleic Acids Res* 28:1221–1227 (2000).
- Kaffer CR, Srivastava M, Park KY, Ives E, Hsieh S, Battle J, Grinberg A, Huang SP, Pfeifer K: A transcriptional insulator at the imprinted *H19/Igf2* locus. *Genes Dev* 14:1908–1919 (2000).
- Kanduri C, Pant V, Loukinov D, Pugacheva E, Qi CF, Wolfe A, Ohlsson R, Lobanenkova VV: Functional association of CTCF with the insulator upstream of the *H19* gene is parent of origin-specific and methylation-sensitive. *Curr Biol* 10:853–856 (2000).
- Kato Y, Sasaki H: Imprinting and looping: epigenetic marks control interactions between regulatory elements. *Bioessays* 27:1–4 (2005).
- Kono T, Obata Y, Yoshimizu T, Nakahara T, Carroll J: Epigenetic modifications during oocyte growth correlates with extended parthenogenetic development in the mouse. *Nat Genet* 13:91–94 (1996).
- Kono T, Sotomaru Y, Katsuzawa Y, Dandolo L: Mouse parthenogenetic embryos with monoallelic *H19* expression can develop to day 17.5 of gestation. *Dev Biol* 243:294–300 (2002).

- Kono T, Obata Y, Wu Q, Niwa K, Ono Y, Yamamoto Y, Park ES, Seo JS, Ogawa H: Birth of parthenogenetic mice that can develop to adulthood. *Nature* 428:860–864 (2004).
- Lee JE, Pintar J, Efstratiadis A: Pattern of the insulin-like growth factor II gene expression during early mouse embryogenesis. *Development* 110:151–159 (1990).
- Lee J, Inoue K, Ono R, Ogonuki N, Kohda T, Kaneko-Ishino T, Ogura A, Ishino F: Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells. *Development* 129:1807–1817 (2002).
- Leighton PA, Ingram RS, Eggenschwiler J, Efstratiadis A, Tilghman SM: Disruption of imprinting caused by deletion of the H19 gene region in mice. *Nature* 375:34–39 (1995a).
- Leighton PA, Saam JR, Ingram RS, Stewart CL, Tilghman SM: An enhancer deletion affects both *H19* and *Igf2* expression. *Genes Dev* 9:2079–2089 (1995b).
- Li JY, Lees-Murdock DJ, Xu GL, Walsh CP: Timing of establishment of paternal methylation imprints in the mouse. *Genomics* 84:952–960 (2004).
- Li YM, Franklin G, Cui HM, Svensson K, He XB, Adam G, Ohlsson R, Pfeifer S: The *H19* transcript is associated with polysomes and may regulate *IGF2* expression in *trans*. *J Biol Chem* 273:28247–28252 (1998).
- Lopes S, Lewis A, Hajkova P, Dean W, Oswald J, Forne T, Murrell A, Constância M, Bartolomei M, Walter J, et al: Epigenetic modifications in an imprinting cluster are controlled by a hierarchy of DMRs suggesting long-range chromatin interactions. *Hum Mol Genet* 12:295–305 (2003).
- Lucifero D, Mertineit C, Clarke HJ, Bestor TH, Trasler JM: Methylation dynamics of imprinted genes in mouse germ cells. *Genomics* 79:530–538 (2002).
- Maher ER, Reik W: Beckwith-Wiedemann syndrome: imprinting in clusters revisited. *J Clin Invest* 105:247–252 (2000).
- Murrell A, Heeson S, Reik W: Interaction between differentially methylated regions partitions the imprinted genes *Igf2* and *H19* into parent-specific chromatin loops. *Nat Genet* 36:889–893 (2004).
- Obata Y, Kaneko-Ishino T, Koide T, Takai Y, Ueda T, Domeki I, Shiroishi T, Ishino F, Kono T: Disruption of primary imprinting during oocyte growth leads to the modified expression of imprinted genes during embryogenesis. *Development* 125:1553–1560 (1998).
- Obata Y, Kono T, Hatada I: Gene silencing: maturation of mouse fetal germ cells in vitro. *Nature* 418:497 (2002).
- Pachnis V, Brannan CL, Tilghman SM: The structure and expression of a novel gene activated in early mouse embryogenesis. *EMBO J* 7:673–681 (1988).
- Poirier F, Chan C-TJ, Timmons PM, Robertson EJ, Evans MJ, Rigby PWJ: The murine *H19* gene is activated during embryonic stem cell differentiation in vitro and at the time of implantation in the developing embryo. *Development* 113:1105–1114 (1991).
- Prawitt D, Enklaar T, Gartner-Rupprecht B, Spangenberg C, Oswald M, Lausch E, Schmidtke P, Reutzel D, Fees S, Lucito R, et al: Microdeletion of target sites for insulator protein CTCF in a chromosome 11p15 imprinting center in Beckwith-Wiedemann syndrome and Wilms' tumor. *Proc Natl Acad Sci USA* 102:4085–4090 (2005).
- Purbowasito W, Suda C, Yokomine T, Zubair M, Sado T, Tsutsui K, Sasaki H: Large-scale identification and mapping of nuclear matrix-attachment regions in the distal imprinted domain of mouse chromosome 7. *DNA Res* 11:391–407 (2004).
- Ripoche MA, Kress C, Poirier F, Dandolo L: Deletion of the H19 transcription unit reveals the existence of a putative imprinting control element. *Genes Dev* 11:1596–1604 (1997).
- Runge S, Nielsen FC, Nielsen J, Lykke-Andersen J, Wewer UM, Christiansen J: H19 RNA binds four molecules of insulin-like growth factor II mRNA-binding protein. *J Biol Chem* 275:29562–29569 (2000).
- Sasaki H, Ferguson-Smith AC, Shum ASW, Barton SC, Surani MA: Temporal and spatial regulation of H19 imprinting in normal and uniparental mouse embryos. *Development* 121:4195–4202 (1995).
- Schmidt JV, Levorse JM, Tilghman SM: Enhancer competition between *H19* and *Igf2* does not mediate their imprinting. *Proc Natl Acad Sci USA* 96:9733–9738 (1999).
- Sparago A, Cerrato F, Vernucci M, Ferrero GB, Silengo MC, Riccio A: Microdeletions in the human *H19* DMR result in loss of *IGF2* imprinting and Beckwith-Wiedemann syndrome. *Nat Genet* 36:958–960 (2004).
- Srivastava M, Hsieh S, Grinberg A, Williams-Simons L, Huang SP, Pfeifer K: *H19* and *Igf2* monoallelic expression is regulated in two distinct ways by a shared *cis* acting regulatory region upstream of *H19*. *Genes Dev* 14:1186–1195 (2000).
- Sun FL, Dean WL, Kelsey G, Allen ND, Reik W: Transactivation of *Igf2* in a mouse model of Beckwith-Wiedemann syndrome. *Nature* 389:809–815 (1997).
- Szabó PE, Mann JR: Allele-specific expression and total expression levels of imprinted genes during early mouse development: Implications for imprinting mechanisms. *Genes Dev* 9:3097–3108 (1995a).
- Szabó PE, Mann JR: Biallelic expression of imprinted genes in the mouse germ line: Implications for erasure establishment and mechanisms of genomic imprinting. *Genes Dev* 9:1857–1868 (1995b).
- Szabó P, Tang SH, Rentsendorj A, Pfeifer GP, Mann JR: Maternal-specific footprints at putative CTCF sites in the H19 imprinting control region give evidence for insulator function. *Curr Biol* 10:607–610 (2000).
- Szabó PE, Hubner K, Scholer H, Mann JR: Allele-specific expression of imprinted genes in mouse migratory primordial germ cells. *Mech Dev* 115:157–160 (2002).
- Szabó PE, Tang SH, Silva FJ, Tsark WM, Mann JR: Role of CTCF binding sites in the *Igf2/H19* imprinting control region. *Mol Cell Biol* 24:4791–4800 (2004).
- Thorvaldsen JL, Duran KL, Bartolomei MS: Deletion of the *H19* differentially methylated domain results in loss of imprinted expression of *H19* and *Igf2*. *Genes Dev* 12:3693–3702 (1998).
- Thorvaldsen JL, Mann MR, Nwoko O, Duran KL, Bartolomei MS: Analysis of sequence upstream of the endogenous *H19* gene reveals elements both essential and dispensable for imprinting. *Mol Cell Biol* 22:2450–2462 (2002).
- Tremblay KD, Saam JR, Ingram RS, Tilghman SM, Bartolomei MS: A paternal-specific methylation imprint marks the alleles of the mouse *H19* gene. *Nat Genet* 9:407–413 (1995).
- Tremblay KD, Duran K, Bartolomei M: A 5' 2-kilobase-pair region of the imprinted mouse *H19* gene exhibits exclusive paternal methylation throughout development. *Mol Cell Biol* 17:4322–4329 (1997).
- Ueda T, Abe K, Miura A, Yuzuriha M, Zubair M, Noguchi M, Niwa K, Kawase Y, Kono T, Matsuda Y, et al: The paternal methylation imprint of the mouse *H19* locus is acquired in the gonocyte stage during foetal testis development. *Genes Cells* 5:649–659 (2000).
- Weber M, Milligan L, Delabre A, Antoine E, Brunel C, Cathala G, Forne T: Extensive tissue-specific variation of allelic methylation in the *Igf2* gene during mouse fetal development: relation to expression and imprinting. *Mech Dev* 101:133–141 (2001).
- Weber M, Hagege H, Murrell A, Brunel C, Reik W, Cathala G, Forne T: Genomic imprinting controls matrix attachment regions in the *Igf2* gene. *Mol Cell Biol* 23:8953–8959 (2003).
- Wiles MV: Isolation of differentially expressed human cDNA clones: similarities between mouse and human embryonal carcinoma cell differentiation. *Development* 104:403–413 (1988).
- Wilkin F, Paquette J, Ledru E, Hamelin C, Pollak M, Deal CL, Hamelin C: *H19* sense and antisense transgenes modify insulin-like growth factor-II mRNA levels. *Eur J Biochem* 267:4020–4027 (2000).
- Yoo-Warren H, Pachnis V, Ingram RS, Tilghman SM: Two regulatory domains flank the mouse *H19* gene. *Mol Cell Biol* 8:4707–4715 (1988).
- Zhang Y, Tycko B: Monoallelic expression of the human *H19* gene. *Nat Genet* 1:40–44 (1992).

# *Résultats*

## Article 2

### The H19 locus acts in vivo as a tumour suppressor

Yoshimizu, T., Miroglio, A., Ripoche, M. A., Gabory, A., Vernucci, M., Riccio, A., Colnot, S., Godard, C., Terris, B., Jammes, H. and Dandolo, L. (2008). *Proc Natl Acad Sci U S A* 105, 12417-22.

## **Introduction**

Le gène *H19* transcrit un ARN non traduit de fonction inconnue (Brannan *et al.*, 1990). Ce gène est fortement exprimé au cours de l'embryogenèse et est ensuite réprimé rapidement après la naissance dans tous les tissus à l'exception des muscles squelettiques. Le locus humain 11p15.5, auquel appartient le gène *H19*, a été impliqué dans diverses pathologies telles que le syndrome de Beckwith-Wiedemann (BWS) et les tumeurs de Wilms. Le BWS est caractérisé par une surcroissance somatique et des malformations congénitales. Par ailleurs les enfants atteints par ce syndrome ont une susceptibilité augmentée de développer des tumeurs comme les tumeurs de Wilms (Cooper *et al.*, 2005). Ces observations ont conduit à émettre l'hypothèse de la perte d'expression d'un gène suppresseur de tumeur au locus 11p15.5. L'analyse des tumeurs de Wilms a montré l'association fréquente de ce type tumoral à une épimutation du centre d'Empreinte des gènes *H19* et *Igf2* (Cui *et al.*, 1997). L'hyperméthylation de la région de méthylation différentielle conduit, selon le modèle de compétition des enhancers et d'insulation par CTCF, à la répression totale du gène *H19* et à la surexpression du gène *Igf2*. L'ensemble de ces données a conduit à suspecter un rôle de suppresseur de tumeur pour le gène *H19*.

Des expériences de transfection d'un vecteur d'expression du gène *H19* ont montré que les cellules tumorales transfectées ont une réduction de la clonogénicité en agar et de l'invasivité dans des souris *nude* (Hao *et al.*, 1993). Cependant d'autres observations ont ouvert une controverse. En effet, l'ARN *H19* est fortement surexprimé dans un certain nombre de tissus tumoraux humains et d'autres équipes ont montré qu'un knock-down de l'ARN *H19* diminue l'inhibition de contact de cellules en culture (Berteaux *et al.*, 2005; Barsyte-Lovejoy *et al.*, 2006; Matouk *et al.*, 2007). Ces résultats penchent plutôt en faveur d'un effet oncogénique de l'ARN *H19*.

Afin d'apporter de nouveaux arguments dans ce débat, nous avons voulu développer des systèmes d'études *in vivo* de la fonction anti ou pro-oncogénique de l'ARN *H19*. Nous avons utilisé deux modèles murins n'exprimant pas le gène *H19*. Le premier modèle est notre lignée de souris invalidées pour ce gène, appelée *H19* $\Delta$ 3 (Ripoche *et al.*, 1997). La seconde lignée, *H19* <sup>$\Delta$ enh</sup>, possède une délétion au niveau des enhancers endodermiques communs des gènes *H19/Igf2*. La transmission maternelle de la délétion conduit à l'absence d'expression du

gène *H19* (Leighton *et al.*, 1995). Nous avons développé trois systèmes d'induction de tumeurs *in vivo* :

- le développement de tératocarcinomes par greffe d'embryons sauvages ou *H19*<sup>-/-</sup> sous la capsule rénale.
- le croisement des souris *H19*<sup>-</sup> avec le modèle murin de cancer colorectal *Apc*<sup>Δ14</sup>, qui développe spontanément des tumeurs intestinales et coliques.
- l'induction d'hépatocarcinomes chez des souris sauvages ou n'exprimant pas le gène *H19* par l'antigène T de l'oncogène SV40 sous le contrôle d'un promoteur hépatique, en collaboration avec l'équipe d'A. Riccio.

Ces trois modèles nous ont permis de montrer que le gène *H19* possède bien un effet suppresseur de tumeur chez la souris.

# The *H19* locus acts *in vivo* as a tumor suppressor

Tomomi Yoshimizu\*<sup>†</sup>, Audrey Miroglio\*<sup>†</sup>, Marie-Anne Ripoche\*, Anne Gabory\*, Maria Vernucci<sup>‡</sup>, Andrea Riccio<sup>‡</sup>, Sabine Colnot<sup>§</sup>, Cécile Godard<sup>§</sup>, Benoit Terris<sup>§</sup>, Hélène Jammes\*, and Luisa Dandolo\*<sup>¶</sup>

\*Department of Genetics and Development and <sup>§</sup>Department of Endocrinology, Metabolism and Cancer, Institut Cochin, University Paris Descartes, Centre National de la Recherche Scientifique Unité Mixte de Recherche 8104, Institut National de la Santé et de la Recherche Médicale U567, Paris 75014, France; and <sup>‡</sup>Department of Environmental Science, Second University of Naples and Institute of Genetics and Biophysics A. Buzzati Traverso, CNR, Naples 81100, Italy

Edited by Benjamin Tycko, Department of Pathology, Institute of Cancer Genetics, New York, NY, and accepted by the Editorial Board June 18, 2008 (received for review February 15, 2008)

The *H19* locus belongs to a cluster of imprinted genes that is linked to the human Beckwith-Wiedemann syndrome. The expression of *H19* and its closely associated *IGF2* gene is frequently deregulated in some human tumors, such as Wilms' tumors. In these cases, biallelic *IGF2* expression and lack of expression of *H19* are associated with hypermethylation of the imprinting center of this locus. These observations and others have suggested a potential tumor suppressor effect of the *H19* locus. Some studies have also suggested that *H19* is an oncogene, based on tissue culture systems. We show, using *in vivo* murine models of tumorigenesis, that the *H19* locus controls the size of experimental teratocarcinomas, the number of polyps in the *Apc* murine model of colorectal cancer and the timing of appearance of SV40-induced hepatocarcinomas. The *H19* locus thus clearly displays a tumor suppressor effect in mice.

genomic imprinting | *Igf2* | murine models

The *H19-Igf2* locus is subject to genomic imprinting and has often been used as a paradigm for the study of this particular epigenetic regulation. The *H19* locus produces a 2.5-kb noncoding, spliced, and polyadenylated RNA of yet-unknown function (1, 2). The *Igf2* gene encodes a fetal growth factor, insulin-like growth factor 2. These two genes are located 90 kb apart and are oppositely imprinted: *H19* is maternally expressed and *Igf2* paternally expressed (1, 3). They belong to a large imprinted domain localized on chromosome 7 in mice and chromosome 11p15.5 in humans. The imprinting of *Igf2* and *H19* is controlled by a region located 4 kb upstream from the *H19* transcription unit, defined as the *H19* differentially methylated region (DMR) or imprinting control region (ICR) (4).

The 11p15.5-imprinted domain is linked to the Beckwith-Wiedemann syndrome (BWS), which is characterized by overgrowth phenotypes of affected children as well as a predisposition to develop embryonal tumors such as Wilms' tumor or rhabdomyosarcomas (5). Among the molecular alterations involved in BWS, certain cases (20%) show paternal uniparental disomy (UPD); other cases (5–10%) have hypermethylation of the *H19* DMR; and both lead to lack of expression of *H19* as well as activation of *IGF2*. These patients have higher risk of developing tumors than patients with other molecular defects (6). Genetic and epigenetic alterations at 11p15.5 similar to those found in the BWS have also been demonstrated in nonsyndromic Wilms' tumors. A great number of these cases have either loss of the maternal allele (LOH) or LOI (7, 8). It has thus been suggested that the *H19* gene could have a possible tumor suppressor function (9). The first direct evidence for this tumor suppressor function was provided by *in vitro* experiments in which transfection of *H19* cDNA into G401-transformed kidney cells resulted in loss of tumorigenicity of these cells (10). Subsequent experiments performed with *in vitro* culture systems suggested that *H19* played a role as an oncogene rather than a tumor suppressor gene (11, 12). This controversy has not yet been resolved, as numerous human tumors have been shown to display either overexpression or lack of *H19* expression (13–15).

We decided to investigate the potential role of the *H19* locus *in vivo* by producing murine models of tumorigenesis. We used *H19Δ3*

(16) and *H19ΔEnh* (17) mice (Fig. 1A) and 3 distinct models of tumorigenesis to investigate the potential tumor suppressor activity of the *H19* locus. In the first model, experimental teratocarcinomas induced by grafting embryos under the kidney capsule were compared for size, weight, and histopathology (18). In the second model, the *H19Δ3* mice were bred with mutants of the *Apc* gene, *ApcΔ14/+*, which represent a murine model for colorectal cancer (19). The double mutants lacking *H19* and *Apc* show an increase in number of polyps compared with their *Apc* littermates. Finally, using a transgenic SV40 hepatocarcinoma model (20, 21), we show that the delay of appearance of these tumors is greatly reduced in the absence of *H19*. Interestingly, these models derive from the 3 germ layers (endoderm, mesoderm, and ectoderm) and result in similar phenotypes, showing a tumor suppressor function for the *H19* locus.

## Results

**Teratocarcinoma Model. *H19Δ3* phenotype and *Igf2* expression.** We originally described that in the *H19Δ3* mutants the maternal *Igf2* allele was slightly reexpressed in skeletal muscle but not in liver (16). To identify the precise levels of *Igf2* expression, we extended our analysis to other organs using a cross between *H19Δ3* and *SD7* mice. *Igf2* is biallelically expressed only in mesoderm-derived tissues (limb muscle, tongue, diaphragm, and heart) and not in endoderm-derived tissues (kidney, lung, and liver) (Fig. 1B). Importantly, maternal *Igf2* reexpression reaches at most 20–30% of the paternal allele in 5-day neonates, showing that there is only a slight increase in *Igf2* mRNA levels in the *H19Δ3* mice.

**Production of tumors on wt background.** Experimental teratocarcinomas were produced by grafting E 6.5 embryos under the kidney capsule of syngenic mice. We first compared the weight of tumors obtained after grafting *wt* or *H19Δ3* embryos into *wt* recipient mice. The results showed a clear difference for the two genotypes (Fig. 2A). Although there was some heterogeneity, the overall weight of *H19Δ3* derived tumors was ≈1.6-fold higher than that of *wt*-derived tumors ( $P = 0.015$ ).

In all cases, the tumor and the kidney were clearly individualized, with no invasion of the tumor into the kidney (Fig. 2B). Histologic analysis revealed the presence of different tissues derived from all 3 embryonic germ layers (ectoderm, endoderm, and mesoderm). There were no striking differences in the type of tissue, suggesting that the absence of *H19* does not affect the development of any 1 particular tissue and that all 3 germ layers are involved.

Author contributions: T.Y., A.M., M.-A.R., A.G., M.V., A.R., C.G., B.T., H.J., and L.D. performed research; T.Y., A.M., M.-A.R., A.G., M.V., A.R., B.T., H.J., and L.D. analyzed data; A.R. and L.D. designed research; S.C. and C.G. contributed new reagents/analytic tools; and A.R. and L.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. B.T. is a guest editor invited by the Editorial Board. †T.Y. and A.M. contributed equally to this work.

¶To whom correspondence should be addressed at: Department of Genetics and Development, Institut Cochin, 24 rue Fbg St Jacques, Paris 75014, France. E-mail: luisa.dandolo@cochin.inserm.fr.

© 2008 by The National Academy of Sciences of the USA







mRNA levels ( $4.72 \pm 2.6$  vs.  $4.73 \pm 1.6$  arbitrary units, data not shown).

**Latency of tumor development.** The male mice which developed liver tumors larger than 3 mm by 120 days of age were 11/31 (35%) for the *MatΔEnh* and 5/24 (20%) for the *wt* genotypes, respectively. By 127 and 134 days, the mice with tumors were 6/8 and 6/9 for the *MatΔEnh* genotype and 7/19 and 3/7 for the *wt* genotype, respectively (Fig. 4C). Overall, 23/48 (48%) *MatΔEnh* and 15/50 (30%) *wt* male mice ( $P = 0.037$ ) had liver tumors when analyzed between 120 and 134 days of age. In contrast, *PatΔEnh* mice, expressing high levels of *H19* but low levels of *Igf2*, developed tumors with a delayed time course (20). These results suggest that the lack of *H19* expression causes acceleration in the development of liver tumors, consistent with its proposed role of tumor suppressor.

## Discussion

The presence of a potential tumor suppressor gene in the 11p15.5 chromosomal region has been hypothesized for many years. This stemmed from the observation that in patients with BWS, there was a predisposition toward development of embryonal tumors. Two domains of imprinted genes have been identified in this region, the IC1 domain with the *H19-IGF2* locus and the IC2 domain with the *KCNQ1* locus associated with several other genes such as *CDKN1C* (or *p57KIP2*). The *H19* locus, with no known function was one of the candidates for a tumor suppressor (9).

The initial evidence in support of this hypothesis mainly involved *in vitro* experiments (10). Our aim was to use animal models to investigate the potential tumor suppressor activity of *H19 in vivo*. The *H19Δ3* mice we had produced never developed tumors, whatever background they were bred on (outbred C57BL/6/129 or inbred 129/SvPas). For this reason, we challenged these mutant mice in different tumor models.

The teratocarcinoma model has been an experimental system for producing tumors in the mouse which was described many years ago (27). Teratocarcinomas are composed of highly undifferentiated embryonal carcinoma (EC) cells and of differentiated cells derived from all 3 embryonic layers (mesoderm, endoderm, and ectoderm) (18). Teratocarcinomas can be produced either by grafting embryos or by injecting ES cells under the skin. We chose to graft E 6.5 embryos under the kidney capsule because this approach provided a better control of the number of grafted cells. We were aiming not only to investigate whether tumors were produced in presence or absence of *H19*, but also whether there was a size difference in these different cases.

The comparison of tumors induced by either *wt* or *H19Δ3* embryos clearly shows a difference in the weight of the tumors that are produced after 35 days. There is, of course, some heterogeneity, but overall *H19Δ3* tumors are larger than their *wt* counterparts. Importantly, the size difference between *H19Δ3* and *wt* induced teratocarcinomas was maintained on the *Igf2*<sup>-/-</sup> background (2-fold) compared with the *wt* background (1.6-fold).

These tumors displayed cells derived from endoderm, mesoderm, and ectoderm. This is interesting with regard to observations made on teratocarcinomas produced from androgenetic ES cells, which display lack of *H19* (maternally expressed) as well as disruption of many other imprinted genes (28). These tumors consisted predominantly of striated muscle. Since we find no difference in the type of tissues of *H19Δ3* or *wt* derived teratocarcinomas, the *H19* locus is probably not involved in the overproduction of muscle cells found in the androgenetic tumors. This could be of interest with respect to the occurrence of rhabdomyosarcomas in BWS patients and suggests that perhaps other genes of the 11p15.5 region are responsible for this type of tumor.

The choice of our second colorectal cancer model was prompted by data showing increase in size and number of polyps in mice carrying the *H19Δ13* mutation compared with *wt* mice (24). These mutant mice lack *H19* but also overexpress maternally derived *Igf2*

because of the ICR deletion. To discriminate between the effects of each one of these genes, we performed the cross between *H19Δ3* females (in which the 3-kb transcription unit only is deleted) and *Apc* mutant males.

The number of polyps is >2-fold higher in absence of *H19*, with an increase in the number of small polyps, suggesting that lack of *H19* may play a role in the initiation step of tumorigenesis. Only low levels of *Igf2* mRNA were found in normal tissue (intestine and colon) and in the polyps, with no significant difference between mutant and *wt* polyps. No difference in the crypt size was detected, whereas previously published data suggested that relative levels of *Igf2* are responsible for crypt depth (24). Taken together, the results obtained from both our *H19Δ3* mutants and the *H19Δ13* mutants allow to postulate that *H19* is playing a definite role in the production and size of polyps, whereas *Igf2* may be contributing to the growth of these polyps by affecting the intestinal crypt size.

The SV40 induced hepatocarcinomas interestingly revealed an acceleration in the latency of appearance of the tumors in the absence of *H19* expression. Expression of the *Igf2* and *H19* genes is completely shut off in the liver of adult *CRP-Tag* mice, but is reactivated in a coordinate manner during liver carcinogenesis with conservation of their imprinted expression (26). In addition, loss of the maternal and duplication of the paternal copy of the chromosomal region bearing the *Igf2* and *H19* genes occur at high frequency in the hepatocellular carcinomas. These genetic events resemble the LOH occurring at chromosome 11p15.5 loci in human cancers and result in activation of *IGF2* and lack of *H19* expression.

It has been recently reported that the *H19* gene may act as an oncogene in studies using human cells maintained in culture and injected into mice to produce tumors (15). The discrepancy with our results could be explained by the difference in the systems. The main interest of our study resides in the use of mouse genetics. Our models reproduce a situation in which the potential oncogenic teratocarcinomas, *Apc*<sup>-/+</sup> polyps or SV40 induced hepatocarcinomas are challenged with mice in which the *H19* locus is present (or absent) throughout embryogenesis and the whole life of the mouse. Its effect is therefore constant and this may represent a more biologic situation than a cell culture system. It must also be acknowledged that the *H19* locus may play a more complex role in humans than in mice.

Because *H19* KO mice never spontaneously develop tumors *in vivo*, as other murine models of tumor suppressor genes, *H19* may play the role of a “modifier gene” suppressing tumorigenesis. It could act either through its long noncoding RNA or through the microRNA (miR-675) that has been recently described in exon 1 (29, 30). Targets of *H19* remain to be identified and linked to a biologic function possibly related to pathways involved in tumorigenesis.

## Materials and Methods

**Mouse Strains and Genotyping.** The *H19Δ3* strain carries a 3kb deletion of the *H19* transcription unit and was initially established on an outbred C57BL/6/CBA background (16). Because an isogenic 129 background was required for the teratocarcinoma experiment, we also produced a 129 *H19Δ3* strain by reinjection of the original *H19Δ3*<sup>+</sup> ES cell line into blastocysts. The 129/SvPas *wt* strain and the 129 *Igf2*<sup>-/-</sup> strain (31) were used as recipients for the production of teratocarcinomas. The *H19ΔEnh* mice (17) were maintained on a C57BL/6 background. These mice can be bred as maternal heterozygotes (*MatH19ΔEnh*) which lack *H19* expression since the deletion is carried on the maternal allele (Fig. 1A). The *SD7* strain is a C57BL/6/CBA strain carrying the distal part of *Mus spretus* chromosome 7. The *ApcΔ14*<sup>+</sup> mice were bred on a C57BL/6 background (19) and were crossed with either the outbred *H19Δ3* or the inbred 129 *H19Δ3*. The *CRP-Tag* 60-3 strain (25) was maintained on a BALB/c background. The protocol of animal handling and treatment was performed in accordance with the guidelines of the animal ethics committee of the Ministère de l'Agriculture of France.

Genotyping was done by PCR on tail DNA. Primers used for detecting the *H19Δ3* allele were *neo* primers: 5'-GTCCTGATAGCGGTCGCCA-3' and 5'-GTGTTCCGGCTGTCAGCGCA-3' (500 bp). The *ApcΔ14* allele was detected using primers that distinguish the *wt* allele (180 bp) from the exon14-excised allele (160

bp): Primer 1: 5'-CTGTTCTGCAATGTTATCA-3'; Primer 2: 5'-CTATGAGTCAACA-CAGGATTA-3'; Primer 3: 5'-TATAAGGGCTAACAGTCAATA-3'.

**Teratocarcinoma Production.** Wt or *H19Δ3* embryos were dissected at 6.5dpc and the ectoplacental cone was taken off. These embryos were introduced under the kidney capsule of isogenic males (8–10 weeks old), as previously described (18). The recipient mice were either 129/SvPas wt or 129 *Igf2<sup>-/-</sup>* strain. 35 days after grafting, tumors were surgically dissected from the mice and weighed. A fraction of the tumor was used to prepare DNA and RNA. The remaining part was fixed in Bouin's fixative, embedded in paraffin, and 5 μm sections were stained with hematoxylin and eosin (H & E).

**Polyp Analysis and Tumor Scoring.** The progeny from the cross between *H19Δ3/+* and *ApcΔ14/+* mutant mice were killed at 180 days and genotyped. The entire intestinal tract was removed, flushed with PBS, and stained with Indigo carmine (0.08%). The small intestine and colon were opened longitudinally, flattened on filter paper, and fixed in 4% PFA. The number and size of the polyps were determined by double counting on mice blinded for genotype. The whole intestine was then rolled and embedded in paraffin for histologic analysis ("Swiss rolls"). Crypt length and tumor grading were performed on 5-μm H & E sections.

**Analysis of Liver Tumorigenesis.** Homozygous *CRP-Tag 60-3* males were mated with *Math19ΔEnh* females, and liver carcinogenesis was analyzed in their progeny. The mice were genotyped for the presence of the *H19ΔEnh*, as described in ref. 21. The males were killed between 120 and 134 days of age. Livers were dissected from the mice and carefully examined for the presence of tumors. Only nodules larger than 3 mm were considered.

**Statistical Analyses.** Data are shown as averages and s.e.m. We used ANOVA analysis and Student *t* test with Excel X and Statview.

**RNA Preparation and Analysis.** Total RNA was extracted from 5-day neonate organs or tumors with TRIzol reagent (Invitrogen). For RT-PCR analysis, DNase I treated RNA (0.5 μg) was reverse-transcribed with SuperScript II and random hexamer primers (Invitrogen). For semiquantitative RT-PCR, 50 ng of cDNA were amplified using gene-specific primers and TaqDNA polymerase (Invitrogen) during 20 cycles. Detection of *Igf2* transcripts derived from the *SD7* cross was performed using primers: forward 5'-GACGTGTCTACTCTCAGCCGCTACT-3' and reverse 5'-GGGTGTCAATTGGGTTGTTAGACCA-3'. The 517-bp product was digested with *BsaA1*, yielding a 241-bp paternal fragment detected by the

internal primer 5'-TCAAATTTGGTTTTTAAAG-3'. RT-PCR products were separated on 1% agarose gels and transferred onto Hybond N+ membranes in 0.4 M NaOH. Blots were probed with  $\gamma^{32}$ P-ATP kinase-labeled primers at 42°C in Church buffer. Membranes were washed in 0.4×SSC, 0.5% SDS at 42°C and results were quantified using PhosphorImager analysis and ImageQuant software.

Quantitative q-PCR was performed on a Light Cycler system using Sybr Green PCR kits (Roche). 1 to 5 ng of cDNA were amplified in duplicate using primers for *H19*, *Igf2*, *GAPDH* and *TBP*. H19 F 5'-GGAGACTAGGCCAGGTCTC-3'; H19 R 5'-GCCCATGGTGTCAAGAAGGC-3'; *Igf2* F 5'-GGCCCCGAGAGACTCTGTGC-3'; *Igf2* R 5'-TGGGGTGGGTAAGGAGAAAC-3'; *GAPDH* F 5'-ACAGTCCATGCCATCACTGCC-3'; *GAPDH* R 5'-GCCTGCTTACCACCTTCTTG-3'; *TBP* F 5'-GCAATCAACATCTCAGCAACC-3' and *TBP* R 5'-CGAAGTGCAATGGTCTTTAGG-3'. Genom calculations were used for normalization.

Northern analysis and RNase protection assays were carried out as previously described (20).

**Methylation Assay.** DNA from tumors, polyps or control tissue was incubated at 55°C in lysis buffer (Tris 100 mM pH 8, EDTA 5 mM, SDS 0.2%, NaCl 20 mM, and 0.4 mg/ml Proteinase K (Sigma), followed by phenol-chloroform extraction and ethanol precipitation. DNA was digested with *SacI* and *HhaI*, separated on 1% agarose gels and transferred onto Hybond N+ membranes. Southern blots were probed with a 200-bp PCR product corresponding to the region covering the CTCF site No. 3: 5' CTGTTATGTGCAACAAGGGAA and 3' GGTCTTACCAGCCACTGA. Blots were washed at 65°C and quantified as described above.

**In situ Hybridization.** *In situ* hybridization on liver sections was performed as previously described (26).

**ACKNOWLEDGMENTS.** We thank Arg Efstratiadis, Wolf Reik, and Shirley Tilghman for their kind gifts of *Igf2<sup>-/-</sup>* and *SD7* mice to L.D and *H19Δenh* mice to A.R. We thank Béatrice Romagnolo and Christine Perret for advice on the *ApcΔ14* colorectal cancer model. We are very grateful to Davor Solter, Jean Gaillard, and Michel Huerre for helpful comments and a critical eye on the teratocarcinoma sections. We thank Debra Wolgemuth and Edith Heard for constructive discussions on the manuscript. We thank Charles Babinet, Wolf Reik, and Azim Surani for constant encouragement. This work was supported by funding from the Ministère de la Recherche (ACI), Association de la Recherche contre le Cancer (ARC), Ligue contre le Cancer, Association Française contre les Myopathies (AFM), MIUR PRIN 2005 (AR), Istituto Superiore di Sanità (AR), Associazione Italiana Ricerca sul Cancro (AR), Telethon-Italia Grant GGP07086 (AR), and fellowships from ARC and AFM to A.M. and to T.Y.

- Bartolomei MS, Zemel S, Tilghman SM (1991) Parental imprinting of the mouse *H19* gene. *Nature* 351:153–155.
- Gabory A, Ripoche MA, Yoshimizu T, Dandolo L (2006) The *H19* gene: Regulation and function of a non-coding RNA. *Cytogenet Genome Res* 113:188–193.
- DeChiara TM, Robertson EJ, Efstratiadis A (1991) Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* 64:849–859.
- Thorvaldsen JL, Duran KL, Bartolomei MS (1998) Deletion of the *H19* differentially methylated domain results in loss of imprinted expression of *H19* and *Igf2*. *Genes Dev* 12:3693–3702.
- Weksberg R, Smith AC, Squire J, Sadowski P (2003) Beckwith-Wiedemann syndrome demonstrates a role for epigenetic control of normal development. *Hum Mol Genet* 12:R61–R68.
- Cooper WN, et al. (2005) Molecular subtypes and phenotypic expression of Beckwith-Wiedemann syndrome. *Eur J Hum Genet* 13:1025–1032.
- Moulton T, et al. (1994) Epigenetic lesions at the *H19* locus in Wilms' tumor patients. *Nat Genet* 7:440–447.
- Feinberg AP (1999) Imprinting of a genomic domain of 11p15 and loss of imprinting in cancer: An introduction. *Cancer Res* 59:1743s–1746s.
- Rump P, Zeegers MP, van Essen AJ (2005) Tumor risk in Beckwith-Wiedemann syndrome: A review and meta-analysis. *Am J Med Genet A* 136:95–104.
- Hao Y, Crenshaw T, Moulton T, Newcomb E, Tycko B (1993) Tumor-suppressor activity of *H19* RNA. *Nature* 365:764–767.
- Lustig-Yariv O, et al. (1997) The expression of the imprinted genes *H19* and *IGF-2* in choriocarcinoma cell lines. Is *H19* a tumor suppressor gene? *Oncogene* 15:169–177.
- Verkerk A, et al. (1997) Unique expression patterns of *H19* in human testicular cancers of different etiology. *Oncogene* 14:95–107.
- Dao D, et al. (1999) Multipoint analysis of human chromosome 11p15/mouse distal chromosome 7: Inclusion of *H19/IGF2* in the minimal WT2 region, gene specificity of *H19* silencing in Wilms' tumorigenesis and methylation hyper-dependence of *H19* imprinting. *Hum Mol Genet* 8:1337–1352.
- Frevel MA, Sowerby SJ, Petersen GB, Reeve AE (1999) Methylation sequencing analysis refines the region of *H19* epimutation in Wilms tumor. *J Biol Chem* 274:29331–29340.
- Matouk IJ, et al. (2007) The *H19* non-coding RNA is essential for human tumor growth. *PLoS ONE* 2:e845.
- Ripoche MA, Kress C, Poirier F, Dandolo L (1997) Deletion of the *H19* transcription unit reveals the existence of a putative imprinting control element. *Genes Dev* 11:1596–1604.
- Leighton PA, Saam JR, Ingram RS, Stewart CL, Tilghman SM (1995) An enhancer deletion affects both *H19* and *Igf2* expression. *Genes Dev* 9:2079–2089.
- Damjanov I, Solter D, Skreb N (1971) Teratocarcinogenesis as related to the age of embryos grafted under the kidney capsule. *Wilhelm Roux' Archiv* 107:288–290.
- Colnot S, et al. (2004) Colorectal cancers in a new murine model of familial adenomatous polyposis: Influence of genetic and environmental modifiers. *Lab Invest* 84:1619–1630.
- Vernucci M, et al. (2000) The *H19* endodermal enhancer is required for *Igf2* activation and tumor formation in experimental liver carcinogenesis. *Oncogene* 19:6376–6385.
- Vernucci M, et al. (2004) Developmentally regulated functions of the *H19* differentially methylated domain. *Hum Mol Genet* 13:353–361.
- Hassan AB, Howell JA (2000) Insulin-like growth factor II supply modifies growth of intestinal adenoma in *Apc(Min/+)* mice. *Cancer Res* 60:1070–1076.
- Leighton PA, Ingram RS, Eggenschwiler J, Efstratiadis A, Tilghman SM (1995) Disruption of imprinting caused by deletion of the *H19* gene region in mice. *Nature* 375:34–39.
- Sakatani T, et al. (2005) Loss of imprinting of *Igf2* alters intestinal maturation and tumorigenesis in mice. *Science* 307:1976–1978.
- Ruther U, Woodrooffe C, Fattori E, Ciliberto G (1993) Inducible formation of liver tumors in transgenic mice. *Oncogene* 8:87–93.
- Casola S, et al. (1995) Loss of heterozygosity of imprinted genes in SV40 t/T antigen-induced hepatocellular carcinomas. *Oncogene* 11:711–721.
- Stevens LC (1970) The development of transplantable teratocarcinomas from intratesticular grafts of pre- and postimplantation mouse embryos. *Dev Biol* 21:364–382.
- Mann JR, Gadi I, Harbison ML, Abbondanzo SJ, Stewart CL (1990) Androgenetic mouse embryonic stem cells are pluripotent and cause skeletal defects in chimeras: Implications for genetic imprinting. *Cell* 62:251–260.
- Mineno J, et al. (2006) The expression profile of microRNAs in mouse embryos. *Nucleic Acids Res* 34:1765–1771.
- Cai X, Cullen BR (2007) The imprinted *H19* noncoding RNA is a primary microRNA precursor. *RNA* 13:313–316.
- DeChiara TM, Efstratiadis A, Robertson EJ (1990) A growth deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by gene targeting. *Nature* 345:78–80.

## **Conclusions**

Le premier modèle murin que nous avons développé consistait à induire des tératocarcinomes par introduction d'un embryon sous la capsule rénale. Dans ce modèle, les tumeurs induites par les embryons *H19*<sup>-/-</sup> sont 40% plus grosses que les tumeurs induites par les embryons sauvages 35 jours après greffe sur des souris sauvages. Le facteur de croissance *Igf2* produit par la souris porteuse pouvant influencer sur le développement des tumeurs, la même expérience a été réalisée sur des souris *Igf2*<sup>-/-</sup>. La même différence de 40% entre embryons greffés sauvages et *H19*<sup>-/-</sup> a été observée sur ce fond génétique, bien que les tumeurs soient quatre fois plus petites que sur porteur sauvage. Ces données ont permis de suggérer un rôle du gène *H19* sur la croissance tumorale.

Le second modèle était l'induction de tumeur dans le modèle *Apc* <sup>$\Delta$ 14</sup>. Ces souris portent une délétion de l'exon 14 du gène suppresseur de tumeur *Apc* qui conduit à l'absence d'expression de cette protéine. Ces souris développent spontanément des tumeurs dans l'intestin et le colon en six mois. Nous avons croisé les femelles *H19*<sup>+/-</sup> avec des mâles *Apc* <sup>$\Delta$ 14</sup> afin de comparer les descendants *APC* <sup>$\Delta$ 14</sup>*H19*<sup>+/+</sup> et *APC* <sup>$\Delta$ 14</sup>*H19*<sup>+/-</sup>. Le nombre de tumeurs développées en absence du gène *H19* est augmenté de plus de deux fois. Par ailleurs, le nombre de petites tumeurs (<2mm) est plus significativement augmenté que celui des grosses tumeurs (>2mm). Ces données suggèrent un rôle dans le processus d'initiation tumoral colique.

Le dernier modèle était l'induction d'hépatocarcinomes par l'antigène T de l'oncogène SV40 sous le contrôle du promoteur du gène spécifique du foie CRP (C reactive protein). Les souris mâles expriment à un taux faible l'antigène T, ce qui conduit à la formation d'hépatocarcinomes multiples à l'âge de trois à cinq mois. Nous avons croisé ces souris avec les femelles *H19* <sup>$\Delta$ enh</sup> et analysé l'atteinte hépatique à 120, 127 et 134 jours. Le nombre de souris atteintes entre 120 et 134 jours est plus important dans la population *H19* <sup>$\Delta$ enh/+</sup> (48%) que dans la population sauvage (30%). Ces résultats suggèrent que l'absence d'expression du gène *H19* entraîne l'accélération du développement des tumeurs hépatocytaires.

Grâce à ces trois modèles, nous avons pu corrélérer l'absence d'expression du gène *H19* avec l'augmentation de la croissance, du nombre et de la vitesse d'apparition des tumeurs. Ces résultats sont donc vraiment en faveur d'un rôle suppresseur de tumeur pour le gène *H19*. Les différences observées avec les précédentes études peuvent s'expliquer par le système utilisé. Nos expériences *in vivo* reproduisent les événements qui peuvent se produire au cours de l'embryogenèse et tout au long de la vie de la souris, reproduisant une situation physiologique. En revanche, les souris invalidées pour le gène *H19* ne développent jamais de tumeurs spontanées, indiquant un rôle de gène modificateur supprimant la tumorigenèse plutôt qu'un rôle de "gate-keeper" comme le gène *Apc* par exemple. Les cibles de l'ARN *H19* restent à définir afin de mieux comprendre le mécanisme de ce suppresseur de tumeur.

## Article 3

*The H19 gene acts as a trans regulator of the  
imprinted gene network controlling growth in mice*

Article soumis pour publication

## **Introduction**

Afin d'élucider la question du rôle du gène *H19* et de son ARN non codant au cours de l'embryogenèse, des modèles d'inactivation du gène par recombinaison homologue ont été produits. Le modèle *H19*<sup>A13</sup> possède une délétion des 3 kb de la séquence transcrite du gène *H19* et 10 kb en amont (Leighton *et al.*, 1995). La délétion *H19*<sup>A3</sup>, restreinte à la séquence transcrite du gène, a été produite dans notre laboratoire (Ripoche *et al.*, 1997). Malgré l'expression très forte de l'ARN *H19* au cours de l'embryogenèse, les souris *H19*<sup>-/-</sup> sont viables et fertiles. Le phénotype est une surcroissance somatique variable selon le modèle qui s'explique par la perte d'empreinte du gène *Igf2*. Ce phénomène est probablement plus lié à une architecture de la chromatine altérée qu'à l'absence de l'ARN *H19* lui-même.

Nous avons décidé d'explorer la fonction du gène *H19* dans des modèles de souris transgéniques. En sortant le gène de son contexte chromatinien, nous pouvions alors étudier l'impact des modifications de l'expression de l'ARN lui-même. Nous avons utilisé deux modèles de souris transgéniques. Une lignée Tg, dans lequel l'unité de transcription du gène *H19* est placée sous le contrôle de séquences promotrices conservées du gène *Necdin*. Le choix de ce promoteur nous permettait également d'exprimer l'ARN dans le cerveau, organe où il est normalement réprimé, afin d'observer si cette induction entraînait un effet phénotypique. La seconde lignée est la lignée YAC décrite précédemment (Ainscough *et al.*, 2000). Une première observation a été la faible surexpression du gène *H19* lorsque le transgène est placé sur fond *H19*<sup>+/+</sup>. En revanche, placé sur fond *H19*<sup>-/+</sup>, l'expression du gène à partir du transgène est plus importante. Nous avons donc décidé de croiser les souris afin d'obtenir dans la même portée trois génotypes d'intérêt : sauvage, *H19*<sup>-/+</sup> et *H19*<sup>-/+</sup>Tg ou *H19*<sup>-/+</sup>YAC selon la lignée. Ces trois génotypes nous permettaient de voir si l'expression ectopique du gène *H19* pouvait restaurer un phénotype sauvage.

Nous avons effectivement observé une récupération du phénotype de surcroissance, les animaux transgéniques étant comparables à des animaux sauvages. Ce sauvetage phénotypique s'accompagne d'un sauvetage de l'expression du gène *Igf2*, l'expression de ce gène étant diminuée lorsque le transgène s'exprime sur fond *H19*<sup>-/+</sup>.

Les gènes *H19* et *Igf2* ont été décrits comme appartenant à un réseau de gènes soumis à l'empreinte (IGN pour Imprinted Gene Network), impliqué dans la régulation de la croissance embryonnaire (Varrault *et al.*, 2006). Nous avons donc exploré l'expression des gènes de ce réseau, en collaboration avec l'équipe de Laurent Journot. Nous avons observé que onze gènes du réseau sont surexprimés dans les animaux *H19*<sup>-/+</sup> et que, en revanche, leur expression est similaire à celle des animaux sauvages dans les souris *H19*<sup>-/+</sup> transgéniques *H19*. Ces données nous ont permis de postuler que l'ARN *H19* est un *trans* régulateur de l'IGN.

**The *H19* gene acts as a *trans* regulator of the imprinted gene network controlling growth in mice**

Anne Gabory<sup>1</sup>, Marie-Anne Ripoché<sup>1</sup>, Anne Le Digarcher<sup>2</sup>, Françoise Watrin<sup>3</sup>, Ahmed Ziyat<sup>1</sup>, Thierry Forné<sup>4</sup>, Hélène Jammes<sup>5</sup>, Justin Ainscough<sup>6</sup>, Azim Surani<sup>6</sup>, Laurent Journot<sup>2</sup> and Luisa Dandolo<sup>1§</sup>

<sup>1</sup>Genetics and Development Department, Inserm U567, CNRS UMR 8104, University Paris Descartes, Institut Cochin, Paris, France

<sup>2</sup>Institut de Genomique Fonctionnelle, CNRS UMR 5203, INSERM U661, University Montpellier II, Montpellier, France

<sup>4</sup>Institut de Génétique Moléculaire, CNRS UMR 5535, University Montpellier II, Montpellier, France

<sup>5</sup> PHASE Department, INRA, Jouy en Josas, France

<sup>6</sup>Wellcome Trust/Cancer Research UK Gurdon Institute of Cancer and Developmental Biology, University of Cambridge, Cambridge, UK

§Corresponding author :

Luisa Dandolo

Department of Genetics and Development, Institut Cochin

24 rue Fbg St Jacques, Paris 75014, France

Phone 33 1 44 41 24 55      Fax 33 1 44 41 24 48

[dandolo@cochin.inserm.fr](mailto:dandolo@cochin.inserm.fr)

Running title: *H19* gene is a *trans* regulator of IGN

Keywords: *H19*, *Igf2*, transgenic mouse models, imprinted gene network (IGN)

21 pages, 4 figures, 4 tables

## **Abstract**

The *H19* gene encodes a non coding RNA of unknown function. Deletion mouse models show an overgrowth phenotype, due to a *cis* effect of the *H19* locus on the adjacent *Igf2* gene. To explore the function of the RNA itself, we produced transgenic mice overexpressing H19. We observed embryonic and post-natal growth reduction in two independent transgenic lines and detected a decrease in *Igf2* expression. We also analysed the expression of multiple genes from the newly described Imprinted Gene Network (IGN) implicated in embryonic growth. We found that *H19* deletion leads to overexpression of eleven genes of the IGN and that this overexpression is rescued by ectopic expression of *H19* RNA in our transgenic animals. We therefore propose that the *H19* RNA participates as a *trans* regulator in the fine tuning of the IGN. Although the precise mechanism remains unclear, our results open new perspectives in the comprehension of genomic imprinting role in embryonic development and in human imprinting pathologies.

## **Introduction**

The *H19* gene was cloned 20 years ago (Pachnis *et al.*, 1988) and its function remains unknown. The main characteristic of this gene is that it encodes an imprinted non coding RNA (Gabory *et al.*, 2006). Genomic imprinting is an epigenetic mechanism that leads to parent-of-origin specific monoallelic expression. Thus *H19* is expressed from the maternal allele in mouse and human (Bartolomei *et al.*, 1991; Zhang and Tycko, 1992). The *H19* RNA is transcribed by RNA polymerase-II, capped, polyadenylated and spliced but lacks conserved ORFs between human and mouse (Brannan *et al.*, 1990). Its complex conserved predicted structure suggests that the functional product is a non-coding RNA (Juan *et al.*, 2000). This RNA is strongly expressed during embryogenesis in mesoderm and endoderm derived tissues. Its expression is then repressed after birth and found only in skeletal muscle and heart in adults (Poirier *et al.*, 1991). This expression pattern is similar to that of the *Igf2* (*insulin-like growth-factor 2*) gene, which is linked to the *H19* gene on the distal part of mouse chromosome 7 and on human 11p15.5 region (Zemel *et al.*, 1992). *Igf2* is paternally expressed and encodes a major fetal growth-factor (DeChiara *et al.*, 1991).

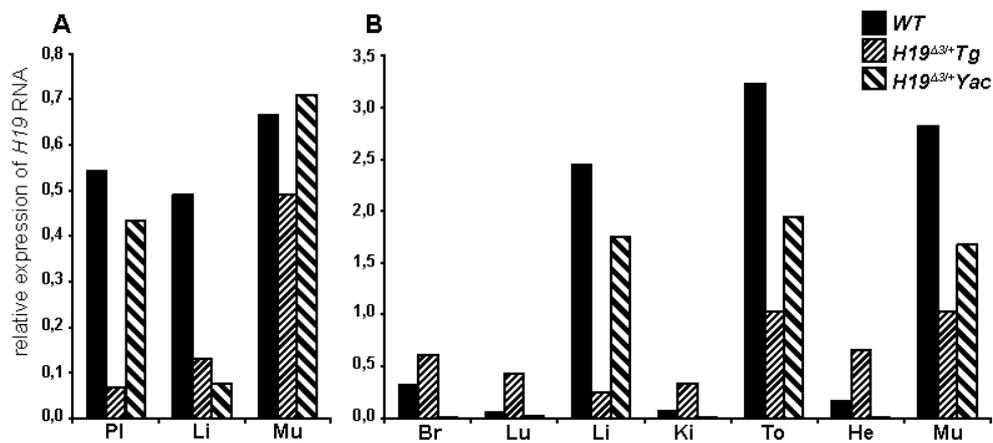
In order to decipher the role of this non coding RNA, two knock-out models were established. In the first model, referred to as *H19*<sup>Δ13</sup>, the 3 kb transcribed region and 10 kb upstream were deleted and replaced by a neomycin resistance gene cassette (neoR) (Leighton *et al.*, 1995). In the second model, referred to as *H19*<sup>Δ3</sup>, the 3 kb transcription unit only was replaced by a neoR cassette (Ripoche *et al.*, 1997). The maternal heterozygous *H19*<sup>-/+</sup> mice, which do not express the RNA because of imprinting, are viable, fertile and present an overgrowth phenotype. In *H19*<sup>Δ13/+</sup> mice, the *Igf2* maternal allele is totally reactivated in all expressing tissues (Leighton *et al.*, 1995). In *H19*<sup>Δ3/+</sup> mice, the *Igf2* maternal allele is expressed at a rate of 25% compared to *Igf2* paternal allele in wild-type mouse, in mesoderm derived tissues only (skeletal muscle, tongue, diaphragm and heart) (Ripoche *et al.*, 1997; Yoshimizu *et al.*, 2008). Imprinted genes are associated with differentially methylated regions (DMRs) which are involved in the regulation of their expression. The DMR located between 2 and 4 kb upstream of the *H19* gene is the imprinting control region (ICR) of the locus, since its deletion affects both *H19* and *Igf2* expression (Thorvaldsen *et al.*, 1998). This sequence binds the insulator protein CTCF on the unmethylated maternal allele, therefore creating a boundary between the enhancers and the *Igf2* maternal allele (Hark *et al.*, 2000). The reactivation of the *Igf2*

maternal allele in  $H19^{+/+}$  mice is probably rather due to a *cis* effect of an altered chromatin structure than to the lack of the  $H19$  RNA.

Different studies have led to propose a role for the  $H19$  RNA. First, maternal transmission of the  $H19^{A13}$  or  $\Delta^3$  deletion leads to modification of the methylation pattern of the  $Igf2$  gene DMRs in *cis* with hypermethylation of the maternal allele but also in *trans* with hypomethylation on the paternal allele (Forne *et al.*, 1997). Second,  $H19$  was found to be associated with polysomes in different human and mouse cell lines (Li *et al.*, 1998; Milligan *et al.*, 2000). Finally, after transfection of human DiHepG2 cells,  $H19$  was shown to participate in IGF2 repression, in part through regulation of  $Igf2$  transcription (Wilkin *et al.*, 2000).

Another clue for the function of the  $H19$  gene is its implication in tumorigenesis.  $H19$  cDNA transfection into a cancer cell line results in loss of clonogenicity of the cells and reduced tumorigenicity in *nude* mice (Hao *et al.*, 1993). Moreover,  $H19/Igf2$  locus imprinting is lost in Wilms' tumour resulting in  $Igf2$  biallelic expression and total repression of  $H19$ . This loss of imprinting is also more frequently found in Beckwith-Wiedemann syndrome patients with tumours (Cooper *et al.*, 2005). Other *in vitro* studies suggested that  $H19$  plays a role in tumorigenicity and is activated by the c-myc oncogene transcription factor (Barsyte-Lovejoy *et al.*, 2006). Finally, we recently showed that  $H19$  acts *in vivo* as a tumour suppressor in three different mouse models (Yoshimizu *et al.*, 2008). Targets of the tumour suppressor  $H19$  locus remain to be identified.

We decided to study the potential role of the  $H19$  RNA *in vivo*. We created a new  $H19$  transgenic Tg line and used the previously described YAC line (Ainscough *et al.*, 2000). We crossed these lines onto an  $H19^{\Delta^3}$  background. Our aim was to rescue the  $H19^{\Delta^3}$  phenotype with the transgenic  $H19$  in order to observe a *trans* effect of the RNA. The overgrowth phenotype observed in  $H19^{\Delta^3}$  animals compared to wild-type littermates is not present in  $H19^{\Delta^3}$  mice expressing the transgenes, suggesting a rescue due to ectopic expression of  $H19$  RNA. An extensive analysis of the expression of genes implicated in the imprinted gene network (IGN) previously described (Varrault *et al.*, 2006) was performed on wild type,  $H19^{\Delta^3}$  and  $H19^{\Delta^3}$  transgenic (Tg and YAC) mice. Our results suggest that  $H19$  is a *trans* regulator of the IGN.



**Figure 1: Ectopic expression of *H19*.** The transgenes were bred on an *H19*<sup>-/+</sup> background in order to detect transgenic *H19* expression alone. **A.** Expression in placenta (Pl), liver (Li) and posterior limb muscle (Mu) tissues from E14.5 embryos. **B.** Expression in brain (Br), lung (Lu), liver (L), kidney (Ki), tongue (To), heart (He) and skeletal muscle (Mu) tissues from P5 neonates. *H19* expression, measured by RT-qPCR, is expressed in arbitrary units, relative to *TBP* gene expression.

## **Results**

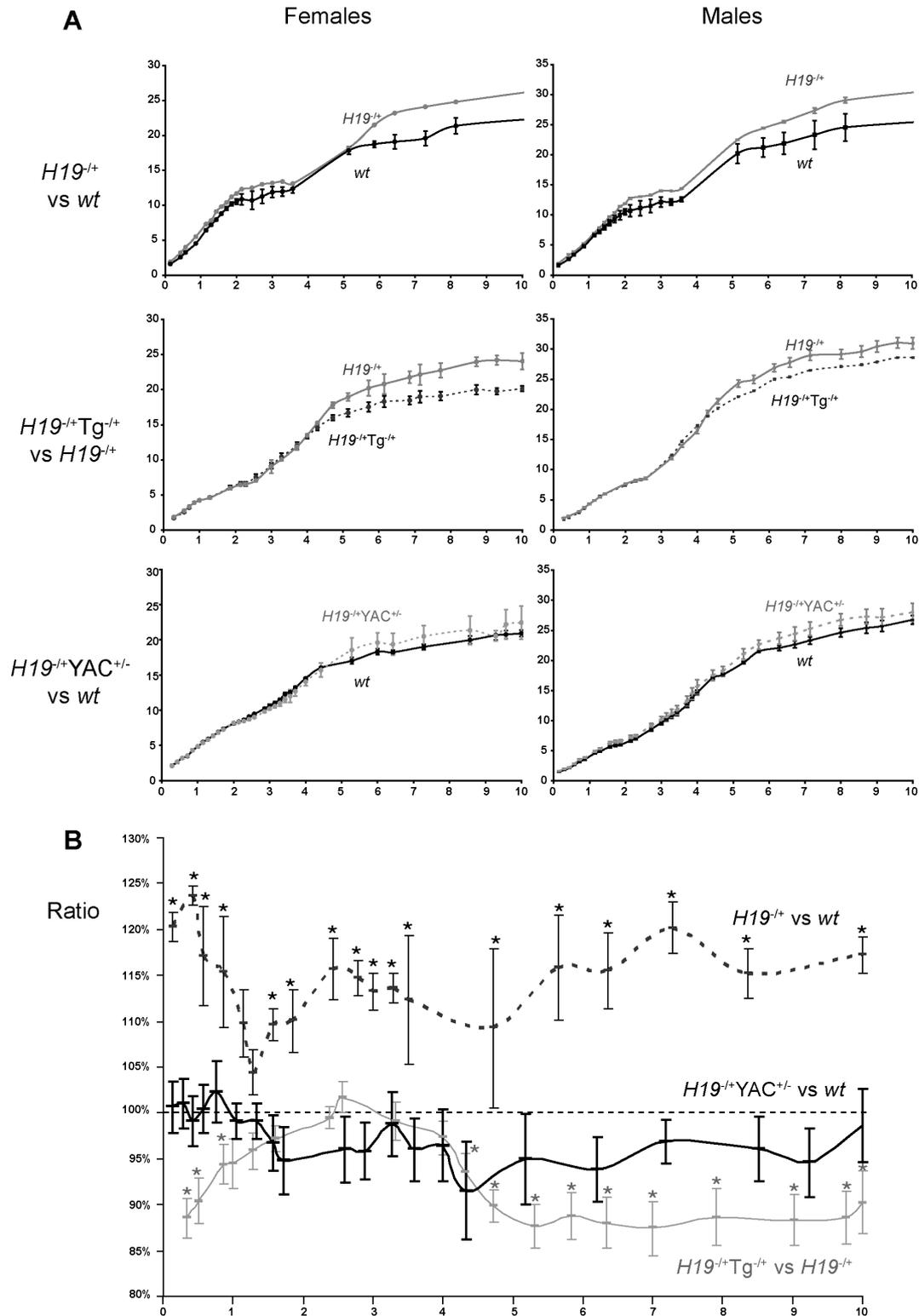
### **Pattern of expression of *H19* in transgenic animals**

We first performed an analysis of *H19* ectopic expression in order to ensure the *H19* expression from transgenic origin. We crossed *H19*<sup>Δ3</sup> females with Tg males and *H19*<sup>Δ3</sup>YAC females with wild-type (*wt*) males to obtain littermates with *H19* expression coming from transgenic origin only. E14.5 embryos were dissected and placenta, liver and muscle (posterior limbs) were harvested. P5 newborns were also dissected and brain, tongue, heart, lung, liver, kidney and muscle from posterior limbs were collected. RNA was prepared from these tissues, the samples of same genotype were pooled and a RT-qPCR was performed (Figure 1). In embryos, ectopic *H19* was expressed in both transgenic lines in all tissues studied: expression is comparable to *wt* in posterior limb muscle but lower in embryonic liver. In placenta, expression in the YAC line is identical to wild type, whereas the Tg line placenta only expressed low amounts of *H19* RNA. In newborn mice of the Tg line, *H19* is transcribed at a lower level than *wt* in liver, tongue and muscle and at a higher level in brain, lung, kidney and heart. The expression profile of the YAC transgenic line is consistent with previously described data (Ainscough *et al.*, 2000), with notably lack of expression in lung, kidney and heart and strong expression in liver, tongue and muscle.

### **Phenotype of the transgenic lines**

As both transgenic lines expressed ectopic *H19*, we closely observed their phenotype. The *H19*<sup>Δ3</sup> mutant displays an overgrowth phenotype. We therefore performed a meticulous analysis of growth in both transgenic lines to observe if ectopic expression of *H19* could rescue the phenotype of the *H19*<sup>Δ3</sup> mutants. Mice were weighed for 10 weeks and weights were plotted on growth curves (Figure 2A). The ratio between mutant and control genotype for each sex was calculated at each point. The average ratio for each cross is plotted on figure 2B. We first compared the growth curves of *H19*<sup>Δ3</sup> animals and *wt* littermates. At birth, the *H19*<sup>-/+</sup> mice have a 20% overgrowth phenotype, this overgrowth decreases to 4% at day 9 and then increases to 14% in the second week. Overall, the *H19*<sup>Δ3/+</sup> mice present a 14% overgrowth compared to their wild type littermates.

*H19*<sup>Δ3</sup> females were crossed with heterozygous Tg males in order to obtain *H19*<sup>-/+</sup> and *H19*<sup>-/+</sup>Tg animals in the same litter. Embryos at E12.5, E14.5, E16.5 and E18.5 were collected and



**Figure 2: Growth phenotype of the transgenic animals. A.** Growth curves for female (left) and male (right) littermates from  $H19^{-/-}$  x  $wt$ ,  $H19^{-/-}$  x  $Tg^{+/-}$  and  $H19^{-/-}YAC^{+/-}$  x  $wt$  crosses. One representative litter is given, as an example for each line. **B.** Ratios for each sex at each point have been calculated for each cross. The graph represents mean ratios for at least 4 littermates during 10 weeks. Error bars indicate standard error of the mean (SEM). 100% represents no difference of one genotype versus the other.

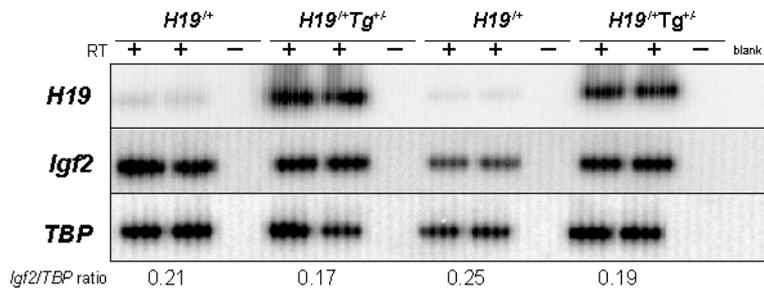
weighed. Additional newborn litters were weighed up to 10 weeks. We observed a growth restriction of  $H19^{-/+}$ Tg animals compared to their control  $H19^{-/+}$  littermates (Figure 2A). This phenotype is already present at E14.5 although the difference is not statistically significant. At birth, the  $H19^{-/+}$ Tg mice weight is decreased by 11% compared to control  $H19^{-/+}$  littermates. Interestingly, the  $H19^{-/+}$ Tg mice catch up and have a similar weight to control mice during the second week. The weight difference decreases again and after weaning the  $H19^{-/+}$ Tg mice present a 11% decrease in weight compared to control  $H19^{-/+}$  littermates. These opposite growth phenotypes suggest that the Tg ectopic  $H19$  expression is able to rescue the  $H19^{\Delta 3}$  overgrowth phenotype.

We then analysed the YAC transgenic line. As the YAC transgene is imprinted, it must be transmitted from the mother to obtain  $H19$  transgenic expression.  $H19^{-/+}YAC^{-/+}$  females were obtained in a first cross. These females were mated with *wt* males in order to obtain *wt*,  $H19^{-/+}$  and  $H19^{-/+}YAC^{+/-}$  littermates. Unfortunately, we observed a bias in the litter and obtained only few  $H19^{-/+}$  animals. Therefore only *wt* and  $H19^{-/+}YAC^{+/-}$  were used for the growth study.  $H19^{-/+}YAC^{+/-}$  mice are not different from *wt* littermates at birth. The ratio of the weight (Figure 2B) of  $H19^{-/+}YAC^{+/-}$  versus *wt* clearly shows that there is no weight difference between these animals. As  $H19^{-/+}$  mice are bigger than *wt* and as there is no weight difference between  $H19^{-/+}YAC^{+/-}$  and *wt*, we can assume that the YAC transgene expression on  $H19^{-/+}$  background rescues the knock-out overgrowth phenotype.

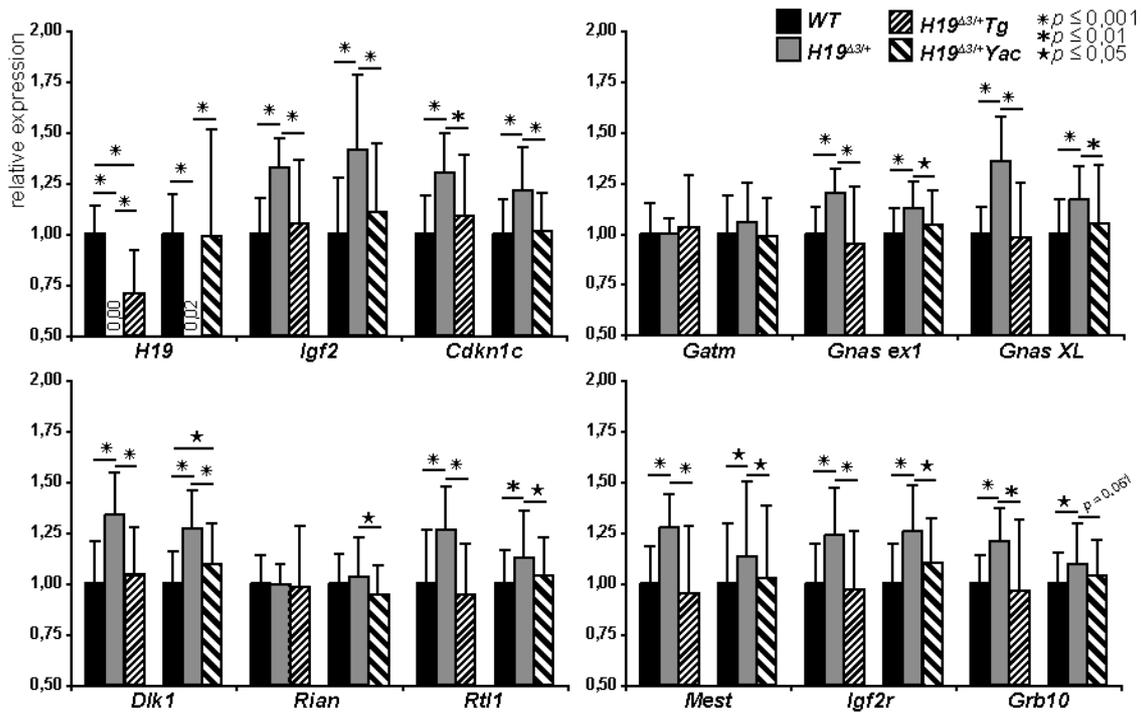
Taken together, these growth curve analyses show that ectopic  $H19$  expression leads to recovery of normal growth on a  $H19^{-/+}$  background. This suggests a role for the  $H19$  RNA itself on the control of growth factors.

### ***Igf2* expression analysis**

The overgrowth phenotype of  $H19^{\Delta 3}$  mice is attributed to the loss of imprinting and overexpression of *Igf2* in mesoderm derived tissues. The  $H19^{\Delta 3/+}$  P5 animals have a 25% increase of *Igf2* mRNA expression in muscle tissue compared to their wild-type littermates (Ripoche *et al.*, 1997). We therefore analysed the *Igf2* expression in the Tg line. We crossed  $H19^{-/-}$  females with Tg heterozygous males in order to obtain  $H19^{-/+}$ Tg and  $H19^{-/+}$  controls in the same litter. P5 neonates were dissected and RNA from muscle tissues was prepared for 2  $H19^{-/+}$ Tg and 2  $H19^{-/+}$  animals. After reverse transcription, a 20 cycle semi-quantitative PCR was performed for *H19*, *Igf2* and *TBP* (Figure 3). As expected, expression of  $H19$  RNA is



**Figure 3: *Igf2* gene expression in Tg line.** *H19*, *Igf2* and *TBP* gene expression were analysed by semi-quantitative RT-PCR. Transgenic animals on an *H19*<sup>-/-</sup> background were compared to control *H19*<sup>+/-</sup> mouse. *Igf2*/*TBP* ratios are indicated under the corresponding fragments.



**Figure 4: Imprinted gene network analysis.** The twelve genes with significant similar results between the two transgenic lines are represented. The expression level of *wt* mice was set at 1 and histograms show modifications relative to this level. For each gene, the three first histograms represent results obtained with *H19*<sup>-/-</sup> x *Tg*<sup>+/-</sup> and the three next ones the results obtained with *H19*<sup>+/-</sup>YAC<sup>+/-</sup> x *wt* crosses. Error bars indicate standard deviation. Student test *p* value significance is indicated with stars.

mainly seen in  $H19^{-/+}$ Tg. Two independent experiments were conducted and when the average  $Igf2/TBP$  expression ratio is set at  $1 \pm 0.07$  for  $H19^{-/+}$  animals (0.21 and 0.25), this ratio becomes  $0.73 \pm 0.10$  for  $H19^{-/+}$ Tg (0.17 and 0.19) indicating a ~25% decrease in  $Igf2$  mRNA expression in  $H19$  expressing mice. These results were further confirmed by RT-qPCR experiments (data not shown) and suggest that the  $H19$  RNA might act in *trans* on  $Igf2$  mRNA level.

### **Imprinted gene network (IGN) analysis**

A network of coexpressed imprinted genes that could regulate embryonic growth has been previously described (Varrault *et al.*, 2006). As  $H19$  and  $Igf2$  belong to this IGN, we decided to study whether loss of  $H19$  function can lead to dysregulation of the other genes of the IGN and whether this dysregulation is rescued by the  $H19$  transgenic expression. We therefore crossed  $H19^{\Delta3/+}$  females with Tg males and  $H19^{\Delta3/+}$ YAC<sup>+/-</sup> females with wild-type males in order to obtain *wt*,  $H19^{-/+}$  and  $H19^{-/+}$ Tg<sup>-/+</sup> littermates and *wt*,  $H19^{-/+}$  and  $H19^{-/+}$ YAC<sup>+/-</sup> littermates, respectively. E14.5 embryos were dissected and placenta, liver and posterior limb muscle tissues were collected. We focused on muscle tissue in which the expression of the transgenes is the strongest (Figure 1). The results of the RT-qPCR experiment have been analysed with a Student statistical test, histograms are shown in Figure 4 and a complete table is available in supplemental data. We were able to show that eight genes are significantly upregulated in  $H19^{-/+}$  mice and rescued to wild-type level in  $H19^{-/+}$ Tg and  $H19^{-/+}$ YAC mice:  $Igf2$  and  $Cdkn1c$ , which are located as  $H19$  on distal part of mouse chromosome 7,  $Gnas$ ,  $Dlk1$ ,  $Rtl1$ ,  $Mest$ ,  $Igf2R$  and  $Grb10$  (in  $H19^{-/+}$ YAC mice,  $Grb10$  rescue is at the limit of significance with  $p=0,051$ ). Interestingly, two genes are found not to be regulated by  $H19$  in both crosses:  $Gatm$  and  $Rian$ . The  $Gtl2$  gene expression is significantly modified in the Tg cross but not in the YAC one, suggesting a non specific effect. Moreover, three genes are found to be regulated in the second cross but were not studied in the first one:  $Dcn$ ,  $Peg3$  and  $Slc38a4$ . We also analysed placenta RNA from the YAC cross. It is interesting to note that a similar effect was not observed in the placenta (data not shown). These results demonstrate for the first time a *trans* effect of the  $H19$  RNA on the mRNA levels of at least eight genes. Moreover this reinforces the finding that a functional imprinting gene network (IGN) exists and suggests that  $H19$  is a *trans* regulator of the IGN.

**A.  $H19^{\Delta 3/+}$  females x Tg males**

	Litter 1*	Litter 2*	Litter 3		total	%	<i>H19</i> expression
<i>wt</i>	3	4	1		8	26%	1,06 ± 0.14 (100%)
<i>Tg<sup>+/-</sup></i>	1	2	1		4	13%	1,39 ± 0.41 (131%)
<i>H19<sup>+/-</sup></i>	5	4	1		10	32%	0,01 ± 0.00 (0%)
<i>H19<sup>+/-</sup>Tg<sup>+/-</sup></i>	2	2	5		9	29%	0,83 ± 0.11 (79%)
TOTAL	11	12	8		31		

**B.  $H19^{\Delta 3/+}$ YAC<sup>+/-</sup> females x wild-type males**

	Litter 1*	Litter 2	Litter 3	Litter 4*	total	%	<i>H19</i> expression
<i>wt</i>	4	3	2	8	17	36%	0.93 ± 0.15 (100%)
YAC <sup>+/-</sup>	0	0	0	1	1	2%	1.07 (115%)
<i>H19<sup>+/-</sup></i>	5	2	2	2	11	24%	0.00 ± 0.00 (0%)
<i>H19<sup>+/-</sup>YAC<sup>+/-</sup></i>	4	2	6	5	18	38%	0.82 ± 0.29 (88%)
TOTAL	13	8	9	16	47		

**Table 1: Bias in transgenic genotype inheritance.** Distribution of the four genotypes obtained in the three  $H19^{\Delta 3/+}$  x  $Tg^{+/-}$  litters and in the four  $H19^{\Delta 3/+}$ YAC<sup>+/-</sup> x *wt* litters. A bias in the transmission of the  $H19^{+/-}$ Tg<sup>+/-</sup> and  $H19^{+/-}$ YAC<sup>+/-</sup> was observed as it is expected to be 25%. The average *H19* expression is given for each genotype. The stars indicate the litter from which originate the animals used for the IGN analysis.

### **Control of *H19* expression in transgenic animals**

In our experiments, we observed a bias in the transgene transmission (Table 1). In the Tg line cross, we obtained only 4 Tg out of 31 animals from 3 litters, which corresponds to 13% instead of the expected 25%. In the YAC line, we obtained only 1 YAC out of 47 animals from 4 litters, which corresponds to 2% instead of the expected 25%. Moreover, the few transgenic animals obtained on the *H19*<sup>+/+</sup> background do not overexpress *H19* (30%) as expected by the expression on *H19*<sup>-/+</sup> background (80%). These results are in favour of a yet unexplained control of the *H19* RNA level during embryogenesis.

## **Discussion**

The *H19* gene was cloned in 1988 and it was soon proposed that it functioned as a non coding RNA (Brannan *et al.*, 1990). Although we have recently shown that the *H19* locus acts as a tumour suppressor, its precise function is still unknown (Yoshimizu *et al.* 2008). Our purpose was to use animal models to explore this function. Since the *H19* deletion displays a *cis* effect on the adjacent *Igf2* gene, we decided to produce transgenic animals to express the RNA outside of his own genomic context. We obtained a transgenic line using the *Necdin* gene promoter, which is mostly ubiquitous with a high expression in brain and central nervous system (Gerard *et al.*, 1999). Although we did not perform extensive behaviour analysis, we did not observe any obvious phenotype due to this ectopic brain expression. We also used the YAC transgenic line previously described (Ainscough *et al.*, 2000). Careful growth curve analysis clearly shows that the *H19* ectopic expression on an *H19*<sup>-/+</sup> background is able, in both transgenic lines, to rescue the *H19*<sup>-/+</sup> overgrowth phenotype, during embryogenesis and throughout adult life. The phenotype of the *H19*<sup>Δ13</sup> and *H19*<sup>Δ3</sup> animals has been explained by a *cis* effect of the *H19* locus on the *Igf2* gene, which becomes biallelically expressed (Leighton *et al.*, 1995; Ripoche *et al.*, 1997). This loss of imprinting results in *Igf2* overexpression, which can explain the overgrowth phenotype. Our results show that the *H19* gene does not only control growth via a *cis* effect on *Igf2* but that the *H19* RNA itself may be implicated. The lack of *H19* RNA may therefore be in part responsible for the *H19*<sup>Δ13</sup> and *H19*<sup>Δ3</sup> phenotypes. As a matter of fact, we observed that *Igf2* expression was reduced in *H19*<sup>-/+</sup>Tg<sup>-/+</sup> animals compared to *H19*<sup>-/+</sup>. This is the first direct proof of an effect of the *H19* RNA on the *Igf2* mRNA expression in *trans*, *in vivo*.

Another clue for *H19* implication in growth control is that it has been shown to belong to an imprinted gene network (IGN) that can control embryonic and post-natal growth (Varrault *et al.*, 2006). This IGN is composed of ~16 coexpressed imprinted genes and almost all of these induce growth phenotypes upon targeted deletion or overexpression. A recent study suggests that 11 of these genes, and among them *H19* and *Igf2*, are important for fetal and early post-natal growth and that their down-regulation contributes to growth deceleration in order to limit body size (Lui *et al.*, 2008). We closely analysed if the expression of the imprinted genes of this network was modified in *H19*<sup>-/+</sup> mice and if this expression could be rescued in *H19*<sup>-/+</sup>Tg<sup>-/+</sup> and *H19*<sup>-/+</sup>YAC<sup>+/-</sup> mice. We observed an upregulation of at least eight imprinted genes in *H19*<sup>-/+</sup> animals and a recovery of normal expression, comparable to *wt*, in transgenic *H19*

	Expression	KO : growth phenotype	Function	Reference
<i>Dlk1</i>	Pat	+/- : Growth restriction	Delta-like transmembrane and secreted protein, EGF-like family	(Moon <i>et al.</i> , 2002)
<i>Gnas</i>	Pat	+/- : Growth restriction	G $\alpha$ subunit of G-protein	(Yu <i>et al.</i> , 1998)
<i>Igf2</i>	Pat	+/- : Growth restriction (60% wt)	Embryonic growth factor	(DeChiara <i>et al.</i> , 1990)
<i>Mest</i>	Pat	+/- : Growth restriction	Presumptive hydrolase enzyme	(Lefebvre <i>et al.</i> , 1998)
<i>Peg3</i>	Pat	+/- : Growth restriction	ZnF protein	(Li <i>et al.</i> , 1999)
<i>Rtl1</i>	Pat	No growth phenotype	Retrotransposon-like protein	(Sekita <i>et al.</i> , 2008)
<i>Slc38a4</i>	Pat	Not done	Amino acid transporter	
<i>Cdkn1c</i>	Mat	No growth phenotype	Inhibition of cell cycle	(Zhang <i>et al.</i> , 1997)
<i>Dcn</i>	Mat	No growth phenotype	Proteoglycan	(Danielson <i>et al.</i> , 1997)
<i>Grb10</i>	Mat	-/+ : Disproportionate overgrowth	Adaptator protein in IGF pathway	(Charalambous <i>et al.</i> , 2003)
<i>Igf2R</i>	Mat	-/+ : overgrowth (140% wt)	<i>Igf2</i> clearance	(Wang <i>et al.</i> , 1994)

**Table 2: Regulated genes and their effect on growth according to targeted deletion experiments.** Seven of the regulated genes are paternally expressed and five of them have a positive growth effect. Four are maternally expressed and three of them have a negative effect on growth or proliferation.

expressing animals on  $H19^{-/+}$  background. These experiments reinforce the functional existence of the IGN. Interestingly, among the regulated genes, four are maternally expressed (*Cdkn1c*, *Igf2R*, *Grb10* and *Dcn*) and seven are paternally expressed (*Igf2*, *Dlk1*, *Rtl1*, *Mest*, *GnasXL*, *Peg3* and *Slc38a4*). In agreement with the parental conflict theory (Moore and Haig, 1991), the majority of paternally expressed regulated genes have a growth promoting effect whereas the majority of the maternally expressed regulated genes inhibit growth as shown by their targeted deletion phenotype (Table 2). Moreover, the deregulation of the genes is modest in magnitude, with a maximum significant increase of ~35% for *Igf2* and of ~15% for *Gnas* and *Grb10*. Overall, the phenotype of  $H19^{-/+}$  mice must therefore be the resultant of upregulation of these growth regulatory factors, which lead to overgrowth. These observations are in favour of a role for *H19* in fine-tuned regulation of the IGN.

Parthenogenetic embryos are not viable (McGrath and Solter, 1984; Surani *et al.*, 1984) and this was the basis of the discovery of imprinting. Data from the production of bi-maternal embryos has suggested a role for *H19* in this lethality. Production of  $ng^{H19\Delta13}/fg^{wt}$  bi-maternal embryos lead to a low survival rate (Kono *et al.*, 2004) and production of  $ng^{AIG-DMR}/fg^{wt}$  embryos lead to birth of non viable mice (Kawahara *et al.*, 2007a). The double  $ng^{H19\Delta13 \Delta IG-DMR}/fg^{wt}$  bi-maternal embryos are viable, with a survival rate close to *in vitro* fertilized embryos (Kawahara *et al.*, 2007b). These results suggested that parthenogenetic lethality is due to a lack of imprinting of the *H19/Igf2* and *Dlk1/Gtl2* loci, both belonging to the IGN (Varrault *et al.*, 2006 194). In particular, bi-allelic *H19* expression would be involved in lethality of  $ng^{AIG-DMR}/fg^{wt}$  embryos suggesting that *H19* transcripts play an important role in mouse development. Our present study brings new clues to better understanding of the role of *H19* as a regulator of a network of imprinted genes which is important for survival and growth of the embryo.

The *H19/Igf2* locus is implicated in different pathologies and particularly in Silver-Russell syndrome (SRS) (Gicquel *et al.*, 2005). In 60% of the SRS patients, hypomethylation of the ICR is observed (Netchine *et al.*, 2007) which could lead to loss of imprinting of the locus with loss of *Igf2* expression and *H19* overexpression. Nevertheless circulating IGF2 peptide levels are found within normal values. Therefore, *H19* overexpression itself could be involved in this syndrome. Other SRS (10%) patients show maternal disomy of chromosome 7, which carries two imprinted loci with *Grb10* and *Sgce*, *Peg10*, *Mest*, *Copγ2* genes respectively

(Eggermann *et al.*, 1997; Preece *et al.*, 1997). In our study we found that *H19* is able not only to control *Igf2* gene expression but also *Grb10* and *Mest* mRNA levels. The regulation of the IGN by *H19* might be a clue to understand the relationship existing between the three implicated imprinted loci in SRS. It would be interesting to explore *H19* RNA and the IGN genes expression in SRS patients.

The functional evidence of this IGN raises the question of the mechanism of regulation between these genes. *Zac1* transcription factor was described as being able to bind common enhancers of the *H19* and *Igf2* genes to promote their expression (Varrault *et al.*, 2006). A link between *Igf2* and *Cdkn1c* has also been reported, even though the mechanism is unknown: increase in circulating Igf2 peptide level in *Igf2R*<sup>+/+</sup> animals leads to decrease of *Cdkn1c* gene expression (Grandjean *et al.*, 2000). *Igf2*, *Igf2R* and *Grb10* belong to the same IGF signalling pathway. Igf2R is a scavenger receptor that leads to the clearance of circulating Igf2 (Wang *et al.*, 1994). Grb10 is an adaptor protein which is an inhibitor of insulin and IGF signalling pathway, although knock-out experiments show that its growth effect is independent of this pathway (Charalambous *et al.*, 2003; Smith *et al.*, 2006). We can also propose a similar mechanism in *H19*<sup>+/+</sup> mice in which *Igf2* expression increase may lead to *Igf2R* and *Grb10* overexpression as a response to maintain overall growth homeostasis. In our models, we show that *H19* can act in *trans* to regulate genes of the IGN, which suggests an effect of the RNA itself. This regulation can occur at transcriptional or post-transcriptional level. Two major features arise from evolutionary conservation study of the *H19* gene: the exon-intron structure of the gene and the precursor of the microRNA miR675 are conserved, suggesting a role for both the full-length spliced transcript and for the processed miRNA (Smits *et al.*, 2008). One mechanism of post-transcriptional control is regulation *via* full-length *H19* RNA and RNA-binding proteins interaction. As an example, *Igf2* mRNA-binding proteins IMP are able to bind to 5' leaders and 3'UTR of *Igf2* mRNAs and can also bind *H19* RNA (Runge *et al.*, 2000). IMP1 (CRD-BP) is also able to bind *c-Myc* RNA and protect it from endonucleotidic cleavage by a polysome associated protein (Sparanese and Lee, 2007). A cleavage also exists for *Igf2* mRNA (Nielsen and Christiansen, 1992; van Dijk *et al.*, 2000). Since *H19* RNA has been described to be associated with polysomes, it could therefore bring *Igf2* mRNA to polysomes via IMP1 interaction and favour *Igf2* mRNA cleavage. Other yet unknown RNA-binding proteins could also be involved in the mechanism by which *H19* RNA can control the IGN gene expression. Another mechanism implicates miR-675. By binding their mRNA targets in miRISC complex with an imperfect match, microRNAs are

known as translational repressors but may also lead to mRNA degradation or sequestration in P-bodies (Chu and Rana, 2007). MiR-675 could therefore act directly on the mRNA levels found to be regulated in our mouse models or indirectly control mRNA stability or translation of a common regulator. It would be interesting to analyse the protein levels of these 11 genes in our mouse models to investigate whether the protein regulation is stronger than mRNA regulation as expected for a miRNA. To distinguish between *H19* full-length and miR-675 effect, it would be conceivable to produce miR-675 transgenic mice and analyse whether this ectopic miRNA expression can rescue the *H19*<sup>A3</sup> phenotype.

In summary, using transgenic mouse models to explore the role of ectopic expression of *H19* RNA, we have shown that this non coding RNA can act as a *trans* regulator, not only on *Igf2*, but also on several other imprinted genes belonging to an interconnected network, involved in the control of growth and survival of mouse embryo. Although the direct or indirect mechanisms involved in this regulation are not yet understood, this opens new perspectives on regulatory aspects of genomic imprinting and its function in development and human pathologies.

**A. PCR primer sequences and annealing conditions**

	Gene amplified	Sense primer (5'-NNN-3')	Antisense primer (5'-NNN-3')	T <sup>a</sup>
Genotyping	<i>Necdin</i>	(N8) GCATCTTATTCATGAGAGAC	(N28) CTCGGTGGAGACCAGCAG	58°C
	<i>Necdin</i>	(N6) GCTCTCCATTTCTATTAGTC	(N16) GCAATATTTCTCTACAAGAC	52°C
	<i>Neo</i>	GTGTTCCGGCTGTCAGCGCA	GTCCTGATAGCGGTCCGCCA	66°C
	<i>LacZ</i>	CAGTTTACCCGCTCTGCTAC	GCGTTGGCAATTAACCGCC	55°C
Semi-quantitative PCR	<i>H19</i>	CCACTACACTACCTGCCTCAGAATCTGC	GGTGGTACTGGGCAGCATTG	60°C
	<i>Igf2</i>	GACGTGTCTACCTCTCAGGCCGTACTION	GGGTGCAATTGGGTTGTTTAGAGCCA	
	<i>TBP</i>	AACACACGGAGCCAGAAGTC	TACTGAACTGCTGGTGGGTC	
qPCR	<i>H19</i>	GGAGACTAGGCCAGGTCTC	GCCCATGGTGTCAAGAAGGC	60°C
	<i>TBP</i>	GCAATCAACATCTCAGCAACC	CGAAGTGCAATGGTCTTTAGG	

**B. Probe sequences**

	Gene	Sequence (5'-NNN-3')
Semi-quantitative PCR	<i>H19</i>	ACATTCATACGGAGAGACTC
	<i>Igf2</i>	ACACGCTTCAGTTTGTCTGT
	<i>TBP</i>	GAGTTGTGCAGAAGTTGGGC

**Table 3: Primers and probes used for PCR, sqPCR and qPCR experiments.**

## **Materials and methods**

### **Transgene construct**

Transgenic constructs for the *Necdin-H19* line consisted in a plasmid containing the *Ndn* promoter (a 1.3 kb *EcoRI-SacI* fragment, accession number D76440) associated with a conserved 2.5kb *EcoRI* fragment (Kuo *et al.*, 1995). This promoter region was placed upstream from the *DraIII-Sall* full genomic region of the *H19* gene (AF04091). Plasmid DNA was prepared by cesium chloride gradient purification and resuspended in transgenic injection buffer (10mM TrisHCl pH7.5, 0.1 mM EDTA pH8, 100mM NaCl) at 5-10ng/ $\mu$ l and microinjected into the pronucleus of C57BL/6xCBA fertilized eggs.

Founders were identified by Southern blot analysis after *HindIII* digestion using a *Ndn* PCR amplified probe (primers N8 and N28, Table 3). Two founders were bred with C57BL/6/CBA females to establish the transgenic lines. Ectopic *H19* expression was then checked by breeding transgenic males with *H19*<sup>A3</sup> females. Only one line, subsequently named Tg line, was kept, since the other displayed variegated expression as detected by *in situ* hybridisation. Expression was found not only in the central nervous system as expected from the expression pattern of the endogenous *Ndn* gene, but also to be relatively ubiquitous, suggesting that the promoter region did not contain all the regulatory control sequences present in the endogenous gene. Transgene copy number was determined to be 30 copies by Southern blot analysis.

*H19*<sup>A3</sup>, Tg and YAC lines were maintained on an outbred C57BL/6/CBA/129 background. *Wt* animals used in the experiments were C57BL/6/CBA mice.

### **Genotyping**

DNA was extracted from tail biopsies and PCR was done according to manufacturer's instructions (Invitrogen). For Tg line, primers N6 and N16, hybridizing to the *Necdin* promoter were used; a 1.5 kb transgenic fragment and a 2 kb *Necdin* endogenous fragment are obtained. For the YAC line, *lacZ* primers were used. *Neo* primers were used to genotype *H19*<sup>A3</sup> allele in heterozygous animals. Primer sequences and annealing temperature conditions are described in Table 3.

### **Gene expression analysis**

Collected tissues were disrupted using a MixerMill apparatus (Qiagen) and total RNA was extracted with TRizol reagent (Invitrogen) according to the manufacturer's instructions. Extracted RNA was RQ1-DNase treated (Promega) and then reextracted with phenol:chloroform and chloroform before ethanol precipitation.

For the expression profile analysis, we carried out reverse transcription with SuperScriptII (Invitrogen) on 500 ng of total RNA with random hexamer oligonucleotides. Quantitative PCR was performed for *H19* and *TBP* genes on 1 ng cDNA in 10 µl final volume with Fast-start MasterMix reagent (Roche) in a Light-Cycler 2.0 apparatus (Roche). For the semi-quantitative PCR, analysis was done on 50 ng cDNA and 20 cycles of amplification were chosen for *H19*, *Igf2* and *TBP* primers. Samples were separated on 1% agarose gel and blotted to Hybond N+ membranes (Amersham) in 0.4N NaOH. Membranes were rinsed with 50mM NaPi pH6.5. Hybridization was performed in Church buffer with an internal probe for each PCR product. Membranes were exposed in a PhosphoImager cassette and quantification was performed with ImageQuant Software (MP Biomedicals). The sequences of primers and probes used and annealing temperature conditions are given in Table 3.

For the imprinted genes network analysis, RNA was prepared from two animals per genotype from two litters; therefore four animals per genotype for each cross were analysed. We carried out 3 independent reverse transcription experiments for each sample on 1 µg of total RNA. qPCR was performed on 2 ng cDNA in a Light-Cycler 480 apparatus (Roche). The primers used to amplify the imprinted genes were described previously (Varrault *et al.*, 2006). The selection of appropriate housekeeping genes was performed with geNorm (Vandesompele *et al.*, 2002). The level of expression of each imprinted gene X was normalized to the geometric mean of the expression levels of several housekeeping genes (*Hprt2*, *Trfr1*, and *Tubb2* for Tg line; *Gus2*, *Mrpl32*, *Hprt2*, and *Tubb2* for YAC line), according to the formula:  $X/\text{geometric mean (R1, R2, R3)} = 2^{(\text{Ct}[X] - \text{arithmetic mean [Ct(R1), Ct(R2), Ct(R3)])}$ , where Ct is the threshold cycle, and R1, R2, R3 are the three reference genes. Student test statistical analysis was performed.

### **Growth curve analysis**

Newborn animals were marked at day 1 and weighed almost every day from P1 to 4 weeks. Males and females were separated at 4 weeks and the littermates of the same sex were placed in the same cage in order to keep the same environment for the different genotypes during the

rest of the experiment. The animals were then weighed twice a week up to 10 weeks. For each line, at least 4 litters were analysed. Weights were plotted on growth curves and a two-factors (litter and genotype) Anova statistical analysis using StatView software was performed for each sex at each point. The weight ratio of transgenic to control animals ( $H19^{-/+}$  or wild-type) was calculated. This ratio was expressed as the percentage of transgenic animal weight versus control.

### **Acknowledgements**

This work was supported by funding from the Ministère de la Recherche (ACI), Association de la Recherche contre le Cancer (ARC), Ligue contre le Cancer, Association Française contre les Myopathies (AFM), as well as fellowships from the MRT and ARC to A.G.

## Figure legends

**Figure 1: Ectopic expression of *H19*.** The transgenes were bred on an *H19*<sup>-/+</sup> background in order to detect transgenic *H19* expression alone. **A.** Expression in placenta (Pl), liver (Li) and posterior limb muscle (Mu) tissues from E14.5 embryos. **B.** Expression in brain (Br), lung (Lu), liver (L), kidney (Ki), tongue (To), heart (He) and skeletal muscle (Mu) tissues from P5 neonates. *H19* expression, measured by RT-qPCR, is expressed in arbitrary units, relative to *TBP* gene expression.

**Figure 2: Growth phenotype of the transgenic animals.** **A.** Growth curves for female (left) and male (right) littermates from *H19*<sup>-/+</sup> x *wt*, *H19*<sup>-/-</sup> x *Tg*<sup>+/-</sup> and *H19*<sup>-/+</sup>*YAC*<sup>+/-</sup> x *wt* crosses. One representative litter is given, as an example for each line. **B.** Ratios for each sex at each point have been calculated for each cross. The graph represents mean ratios for at least 4 littermates during 10 weeks. Error bars indicate standard error of the mean (SEM). 100% represents no difference of one genotype versus the other.

**Figure 3: *Igf2* gene expression in Tg line.** *H19*, *Igf2* and *TBP* gene expression were analysed by semi-quantitative RT-PCR. Transgenic animals on an *H19*<sup>-/+</sup> background were compared to control *H19*<sup>-/+</sup> mouse. *Igf2/TBP* ratios are indicated under the corresponding fragments.

**Figure 4: Imprinted gene network analysis.** The twelve genes with significant similar results between the two transgenic lines are represented. The expression level of *wt* mice was set at 1 and histograms show modifications relative to this level. For each gene, the three first histograms represent results obtained with *H19*<sup>-/-</sup> x *Tg*<sup>+/-</sup> and the three next ones the results obtained with *H19*<sup>-/+</sup>*YAC*<sup>+/-</sup> x *wt* crosses. Error bars indicate standard deviation. Student test *p* value significance is indicated with stars.

**Table 1: Bias in transgenic genotype inheritance.** Distribution of the four genotypes obtained in the three *H19*<sup>-/-</sup> x *Tg*<sup>+/-</sup> litters and in the four *H19*<sup>-/+</sup>*YAC*<sup>+/-</sup> x *wt* litters. A bias in the transmission of the *H19*<sup>+/+</sup>*Tg*<sup>+/-</sup> and *H19*<sup>+/+</sup>*YAC*<sup>+/-</sup> was observed as it is expected to be 25%. The average *H19* expression is given for each genotype. The stars indicate the litter from which originate the animals used for the IGN analysis.

Crosses	♀ <i>H19<sup>Δ3/+</sup></i> x ♂ <i>H19<sup>+/+</sup>Tg</i>										♀ <i>H19<sup>Δ3/+</sup>YAC</i> x ♂ <i>WT</i>											
	WT		<i>H19<sup>-/+</sup></i>				<i>H19<sup>-/+</sup>Tg</i>				WT		<i>H19<sup>-/+</sup></i>				<i>H19<sup>-/+</sup>YAC</i>					
Genotype	Average	sd	Average	sd	KO/WT	p vs WT	Average	sd	KO/Tg /WT	p vs WT	p vs KO	Average	sd	Average	sd	KO /WT	p vs. WT	Average	sd	KO Yac/ WT	p vs WT	p vs KO
<i>Cdkn1c</i>	29,33	5,55	38,14	5,80	1,30	<b>0,000</b>	31,966	8,745	1,09	0,218	<b>0,006</b>	57,52	10,01	70,09	12,15	1,22	<b>0,000</b>	58,35	10,69	1,01	0,463	<b>0,000</b>
<i>Dcn</i>	--	--	--	--	--	--	--	--	--	--	--	6,62	1,07	7,38	1,19	1,11	<b>0,005</b>	6,64	1,25	1,00	0,881	<b>0,000</b>
<i>Dlk1</i>	22,94	4,85	30,73	4,87	1,34	<b>0,000</b>	23,953	5,368	1,04	0,497	<b>0,000</b>	26,41	4,21	33,60	5,06	1,27	<b>0,000</b>	28,96	5,31	1,10	<b>0,015</b>	<b>0,000</b>
<i>Gatm</i>	0,42	0,06	0,42	0,03	1,00	0,979	0,434	0,108	1,03	0,590	0,530	0,50	0,09	0,53	0,09	1,06	0,189	0,50	0,09	0,99	0,860	0,083
<i>Gnas ex1</i>	16,02	2,15	19,25	1,97	1,20	<b>0,000</b>	15,227	4,491	0,95	0,437	<b>0,000</b>	26,03	3,34	29,26	3,44	1,12	<b>0,000</b>	27,27	4,35	1,05	0,117	<b>0,030</b>
<i>Gnas XL</i>	2,36	0,31	3,21	0,52	1,36	<b>0,000</b>	2,319	0,635	0,98	0,782	<b>0,000</b>	2,30	0,39	2,69	0,37	1,17	<b>0,001</b>	2,43	0,66	1,06	0,269	<b>0,010</b>
<i>Gtl2</i>	20,11	3,16	24,77	4,35	1,23	<b>0,000</b>	18,535	6,259	0,92	0,278	<b>0,000</b>	11,37	1,77	11,20	2,13	0,98	0,470	12,07	3,27	1,06	0,230	0,150
<i>Grb10</i>	6,19	0,88	7,49	1,01	1,21	<b>0,000</b>	5,976	2,181	0,97	0,654	<b>0,003</b>	8,13	1,23	8,89	1,65	1,09	<b>0,042</b>	8,44	1,46	1,04	0,137	0,051
<i>H19</i>	265,37	37,71	0,27	0,13	0,00	<b>0,000</b>	188,470	56,145	0,71	<b>0,000</b>	<b>0,000</b>	215,43	42,50	4,75	11,24	0,02	<b>0,000</b>	213,35	113,45	0,99	0,893	<b>0,000</b>
<i>Igf2</i>	102,29	17,93	135,93	14,81	1,33	<b>0,000</b>	107,926	32,066	1,06	0,456	<b>0,000</b>	64,49	17,86	91,28	24,00	1,42	<b>0,000</b>	71,38	21,93	1,11	0,073	<b>0,000</b>
<i>Igf2r</i>	1,72	0,34	2,13	0,40	1,24	<b>0,000</b>	1,670	0,491	0,97	0,704	<b>0,001</b>	0,66	0,13	0,84	0,15	1,26	<b>0,000</b>	0,73	0,14	1,11	0,052	<b>0,024</b>
<i>Mest</i>	10,37	1,91	13,26	1,73	1,28	<b>0,000</b>	9,896	3,426	0,95	0,557	<b>0,000</b>	7,29	2,19	8,28	2,70	1,14	<b>0,014</b>	7,48	2,62	1,03	0,215	<b>0,045</b>
<i>Ndn</i>	--	--	--	--	--	--	--	--	--	--	--	1,58	0,39	1,56	0,40	0,99	0,820	1,53	0,46	0,97	0,543	0,377
<i>Nnat</i>	--	--	--	--	--	--	--	--	--	--	--	10,09	2,00	9,31	1,71	0,92	<b>0,002</b>	8,97	1,75	0,89	<b>0,016</b>	0,521
<i>Peg3</i>	--	--	--	--	--	--	--	--	--	--	--	17,93	2,33	19,84	2,88	1,11	<b>0,001</b>	18,47	2,89	1,03	0,281	<b>0,034</b>
<i>Rian</i>	16,03	2,26	16,02	1,59	1,00	0,991	15,791	4,848	0,99	0,830	0,827	9,96	1,43	10,27	1,93	1,03	0,513	9,43	1,42	0,95	0,243	<b>0,010</b>
<i>Rtl1</i>	0,27	0,07	0,34	0,06	1,27	<b>0,000</b>	0,253	0,066	0,95	0,497	<b>0,000</b>	0,41	0,07	0,47	0,10	1,13	<b>0,003</b>	0,43	0,08	1,04	0,330	<b>0,023</b>
<i>Sgce</i>	--	--	--	--	--	--	--	--	--	--	--	0,10	0,04	0,11	0,04	1,02	0,424	0,09	0,04	0,84	0,075	<b>0,011</b>
<i>Slc38a4</i>	--	--	--	--	--	--	--	--	--	--	--	1,52	0,32	1,81	0,52	1,19	<b>0,004</b>	1,63	0,50	1,07	0,223	<b>0,001</b>
<i>Zac1</i>	--	--	--	--	--	--	--	--	--	--	--	5,80	1,32	6,28	1,84	1,08	0,117	5,52	1,64	0,95	0,291	<b>0,002</b>

**Supplemental Table 1: Complete results of the IGN network analysis.** Each value is the average of four samples of animals coming from two litters. RT-qPCR was performed with geNorm normalisation. Statistical Student *p* value test was performed; significant *p* values are marked in bold. -- indicates not done.

**Table 2: Regulated genes and their effect on growth according to targeted deletion experiments.** Seven of the regulated genes are paternally expressed and five of them have a positive growth effect. Four are maternally expressed and three of them have a negative effect on growth or proliferation.

**Table 3: Primers and probes used for PCR, sqPCR and qPCR experiments.**

**Supplemental Table 1: Complete results of the IGN network analysis.** Each value is the average of four samples of animals coming from two litters. RT-qPCR was performed with geNorm normalisation. Statistical Student *p* value test was performed; significant *p* values are marked in bold. -- indicates not done.

## References

- Ainscough, J. F., Dandolo, L. and Surani, M. A. (2000). Appropriate expression of the mouse *H19* gene utilises three or more distinct enhancer regions spread over more than 130 kb. *Mech Dev* 91, 365-8.
- Barsyte-Lovejoy, D., Lau, S. K., Boutros, P. C., Khosravi, F., Jurisica, I., Andrulis, I. L., Tsao, M. S. and Penn, L. Z. (2006). The c-Myc oncogene directly induces the *H19* noncoding RNA by allele-specific binding to potentiate tumorigenesis. *Cancer Res* 66, 5330-7.
- Bartolomei, M. S., Zemel, S. and Tilghman, S. M. (1991). Parental imprinting of the mouse *H19* gene. *Nature* 351, 153-5.
- Brannan, C. I., Dees, E. C., Ingram, R. S. and Tilghman, S. M. (1990). The product of the *H19* gene may function as an RNA. *Mol Cell Biol* 10, 28-36.
- Charalambous, M., Smith, F. M., Bennett, W. R., Crew, T. E., Mackenzie, F. and Ward, A. (2003). Disruption of the imprinted *Grb10* gene leads to disproportionate overgrowth by an *Igf2*-independent mechanism. *Proc Natl Acad Sci U S A* 100, 8292-7.
- Chu, C. Y. and Rana, T. M. (2007). Small RNAs: regulators and guardians of the genome. *J Cell Physiol* 213, 412-9.
- Cooper, W. N., Luharia, A., Evans, G. A., Raza, H., Haire, A. C., Grundy, R., Bowdin, S. C., Riccio, A., Sebastio, G., Blied, J., Schofield, P. N., Reik, W., Macdonald, F. and Maher, E. R. (2005). Molecular subtypes and phenotypic expression of Beckwith-Wiedemann syndrome. *Eur J Hum Genet* 13, 1025-32.
- DeChiara, T. M., Robertson, E. J. and Efstratiadis, A. (1991). Parental imprinting of the mouse *insulin-like growth factor II* gene. *Cell* 64, 849-59.
- Eggermann, T., Wollmann, H. A., Kuner, R., Eggermann, K., Enders, H., Kaiser, P. and Ranke, M. B. (1997). Molecular studies in 37 Silver-Russell syndrome patients: frequency and etiology of uniparental disomy. *Hum Genet* 100, 415-9.
- Forne, T., Oswald, J., Dean, W., Saam, J. R., Bailleul, B., Dandolo, L., Tilghman, S. M., Walter, J. and Reik, W. (1997). Loss of the maternal *H19* gene induces changes in *Igf2* methylation in both *cis* and *trans*. *Proc Natl Acad Sci U S A* 94, 10243-8.
- Gabory, A., Ripoché, M. A., Yoshimizu, T. and Dandolo, L. (2006). The *H19* gene: regulation and function of a non-coding RNA. *Cytogenet Genome Res* 113, 188-93.
- Gerard, M., Hernandez, L., Wevrick, R. and Stewart, C. L. (1999). Disruption of the mouse *Necdin* gene results in early post-natal lethality. *Nat Genet* 23, 199-202.
- Gicquel, C., Rossignol, S., Cabrol, S., Houang, M., Steunou, V., Barbu, V., Danton, F., Thibaud, N., Le Merrer, M., Burglen, L., Bertrand, A. M., Netchine, I. and Le Bouc, Y. (2005). Epimutation of the telomeric imprinting center region on chromosome 11p15 in Silver-Russell syndrome. *Nat Genet* 37, 1003-7.
- Grandjean, V., Smith, J., Schofield, P. N. and Ferguson-Smith, A. C. (2000). Increased IGF-II protein affects p57kip2 expression in vivo and in vitro: implications for Beckwith-Wiedemann syndrome. *Proc Natl Acad Sci U S A* 97, 5279-84.
- Hao, Y., Crenshaw, T., Moulton, T., Newcomb, E. and Tycko, B. (1993). Tumour-suppressor activity of *H19* RNA. *Nature* 365, 764-7.
- 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, 486-9.
- Juan, V., Crain, C. and Wilson, C. (2000). Evidence for evolutionarily conserved secondary structure in the *H19* tumor suppressor RNA. *Nucleic Acids Res* 28, 1221-7.
- Kawahara, M., Wu, Q., Ferguson-Smith, A. C. and Kono, T. (2007a). Appropriate expression of imprinted genes on mouse chromosome 12 extends development of bi-maternal embryos to term. *FEBS Lett* 581, 5178-84.
- Kawahara, M., Wu, Q., Takahashi, N., Morita, S., Yamada, K., Ito, M., Ferguson-Smith, A. C. and Kono, T. (2007b). High-frequency generation of viable mice from engineered bi-maternal embryos. *Nat Biotechnol* 25, 1045-50.
- Kono, T., Obata, Y., Wu, Q., Niwa, K., Ono, Y., Yamamoto, Y., Park, E. S., Seo, J. S. and Ogawa, H. (2004). Birth of parthenogenetic mice that can develop to adulthood. *Nature* 428, 860-4.

- Kuo, C. H., Uetsuki, T., Kim, C. H., Tanaka, H., Li, B. S., Taira, E., Higuchi, H., Okamoto, H., Yoshikawa, K. and Miki, N. (1995). Determination of a neocdin cis-acting element required for neuron specific expression by using zebra fish. *Biochem Biophys Res Commun* 211, 438-46.
- Leighton, P. A., Ingram, R. S., Eggenschwiler, J., Efstratiadis, A. and Tilghman, S. M. (1995). Disruption of imprinting caused by deletion of the *H19* gene region in mice. *Nature* 375, 34-9.
- Li, Y. M., Franklin, G., Cui, H. M., Svensson, K., He, X. B., Adam, G., Ohlsson, R. and Pfeifer, S. (1998). The *H19* transcript is associated with polysomes and may regulate *IGF2* expression in trans. *J Biol Chem* 273, 28247-52.
- Lui, J. C., Finkielstain, G. P., Barnes, K. M. and Baron, J. (2008). An imprinted gene network that controls mammalian somatic growth is down-regulated during postnatal growth deceleration in multiple organs. *Am J Physiol Regul Integr Comp Physiol* 295, R189-96.
- McGrath, J. and Solter, D. (1984). Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* 37, 179-83.
- Milligan, L., Antoine, E., Bisbal, C., Weber, M., Brunel, C., Forne, T. and Cathala, G. (2000). *H19* gene expression is up-regulated exclusively by stabilization of the RNA during muscle cell differentiation. *Oncogene* 19, 5810-6.
- Moore, T. and Haig, D. (1991). Genomic imprinting in mammalian development: a parental tug-of-war. *Trends Genet* 7, 45-9.
- Netchine, I., Rossignol, S., Dufourg, M. N., Azzi, S., Rousseau, A., Perin, L., Houang, M., Steunou, V., Esteva, B., Thibaud, N., Demay, M. C., Danton, F., Petriczko, E., Bertrand, A. M., Heinrichs, C., Carel, J. C., Loeuille, G. A., Pinto, G., Jacquemont, M. L., Gicquel, C., Cabrol, S. and Le Bouc, Y. (2007). 11p15 imprinting center region 1 loss of methylation is a common and specific cause of typical Russell-Silver syndrome: clinical scoring system and epigenetic-phenotypic correlations. *J Clin Endocrinol Metab* 92, 3148-54.
- Nielsen, F. C. and Christiansen, J. (1992). Endonucleolysis in the turnover of *insulin-like growth factor II* mRNA. *J Biol Chem* 267, 19404-11.
- Pachnis, V., Brannan, C. I. and Tilghman, S. M. (1988). The structure and expression of a novel gene activated in early mouse embryogenesis. *EMBO J* 7, 673-81.
- Poirier, F., Chan, C. T., Timmons, P. M., Robertson, E. J., Evans, M. J. and Rigby, P. W. (1991). The murine *H19* gene is activated during embryonic stem cell differentiation in vitro and at the time of implantation in the developing embryo. *Development* 113, 1105-14.
- Preece, M. A., Price, S. M., Davies, V., Clough, L., Stanier, P., Trembath, R. C. and Moore, G. E. (1997). Maternal uniparental disomy 7 in Silver-Russell syndrome. *J Med Genet* 34, 6-9.
- Ripoche, M. A., Kress, C., Poirier, F. and Dandolo, L. (1997). Deletion of the *H19* transcription unit reveals the existence of a putative imprinting control element. *Genes Dev* 11, 1596-604.
- Runge, S., Nielsen, F. C., Nielsen, J., Lykke-Andersen, J., Wewer, U. M. and Christiansen, J. (2000). *H19* RNA binds four molecules of insulin-like growth factor II mRNA-binding protein. *J Biol Chem* 275, 29562-9.
- Smith, F. M., Garfield, A. S. and Ward, A. (2006). Regulation of growth and metabolism by imprinted genes. *Cytogenet Genome Res* 113, 279-91.
- Smits, G., Mungall, A. J., Griffiths-Jones, S., Smith, P., Beury, D., Matthews, L., Rogers, J., Pask, A. J., Shaw, G., VandeBerg, J. L., McCarrey, J. R., the SAVOIR consortium, Renfree, M. B., Reik, W. and Dunham, I. (2008). Conservation of the *H19* non-coding RNA and *H19-Igf2* imprinting mechanism in therians. *Nat Genet*.
- Sparanese, D. and Lee, C. H. (2007). CRD-BP shields c-myc and MDR-1 RNA from endonucleolytic attack by a mammalian endoribonuclease. *Nucleic Acids Res* 35, 1209-21.
- 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, 548-50.
- 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 Dev* 12, 3693-702.
- van Dijk, E. L., Sussenbach, J. S. and Holthuizen, P. E. (2000). Distinct RNA structural domains cooperate to maintain a specific cleavage site in the 3'-UTR of IGF-II mRNAs. *J Mol Biol* 300, 449-67.

- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3.
- Varrault, A., Gueydan, C., Delalbre, A., Bellmann, A., Houssami, S., Aknin, C., Severac, D., Chotard, L., Kahli, M., Le Digarcher, A., Pavlidis, P. and Journot, L. (2006). *Zac1* regulates an imprinted gene network critically involved in the control of embryonic growth. *Dev Cell* 11, 711-22.
- Wang, Z. Q., Fung, M. R., Barlow, D. P. and Wagner, E. F. (1994). Regulation of embryonic growth and lysosomal targeting by the imprinted *Igf2/Mpr* gene. *Nature* 372, 464-7.
- Wilkin, F., Paquette, J., Ledru, E., Hamelin, C., Pollak, M. and Deal, C. L. (2000). *H19* sense and antisense transgenes modify *insulin-like growth factor-II* mRNA levels. *Eur J Biochem* 267, 4020-7.
- Yoshimizu, T., Miroglio, A., Ripoché, M. A., Gabory, A., Vernucci, M., Riccio, A., Colnot, S., Godard, C., Terris, B., Jammes, H. and Dandolo, L. (2008). The *H19* locus acts *in vivo* as a tumour suppressor. *Proc Natl Acad Sci U S A* 105, 12417-22.
- Zemel, S., Bartolomei, M. S. and Tilghman, S. M. (1992). Physical linkage of two mammalian imprinted genes, *H19* and *insulin-like growth factor 2*. *Nat Genet* 2, 61-5.
- Zhang, Y. and Tycko, B. (1992). Monoallelic expression of the human *H19* gene. *Nat Genet* 1, 40-4.

## **Conclusions**

Nos résultats montrent dans deux lignées transgéniques indépendantes un effet négatif de l'ARN *H19* sur la croissance des souris. Cet effet de croissance s'accompagne d'une répression de l'expression de l'ARNm de onze gènes soumis à l'empreinte appartenant au réseau de gènes soumis à l'Empreinte (IGN) impliqué dans la régulation de la croissance (Varrault *et al.*, 2006). Ces données montrent pour la première fois un effet régulateur en *trans* de l'ARN *H19* sur le niveau d'expression d'autres gènes. Parmi les gènes régulés, sept sont d'expression paternelle : *Igf2*, *Dlk1*, *Rtl1*, *Gnas*, *Mest*, *Peg3* et *Slc38a4* et quatre sont d'expression maternelle : *Cdkn1c*, *Grb10*, *Igf2R* et *Dcn*. Par ailleurs, cinq ont un effet positif sur la croissance embryonnaire et trois ont un effet négatif sur la croissance ou la prolifération cellulaire. Les phénotypes de surcroissance des souris *H19*<sup>-/+</sup> et de croissance similaire à celle des animaux sauvages pour les *H19*<sup>-/+</sup>Tg ou les *H19*<sup>-/+</sup>YAC sont donc dus à la résultante de la régulation de ces gènes. L'ARN *H19* aurait ainsi une fonction de régulation fine de la croissance *via* l'IGN.

Nos résultats apportent une confirmation de l'existence fonctionnelle de ce réseau IGN. Ces données permettent des avancées dans la compréhension de la fonction du gène *H19*, au cours de l'embryogenèse et en tant que gène suppresseur de tumeur, d'une part et de la régulation des gènes soumis à l'empreinte d'autre part.

La question du mécanisme d'action de l'ARN *H19* reste posée. La régulation de l'expression de ces gènes peut être due, au niveau post-transcriptionnel, à l'ARN pleine longueur, *via* des interactions ARN-protéines, ou à miR-675, un microARN produit à partir de l'ARN *H19*. Par ailleurs, l'exploration des mutants des autres gènes de l'IGN pourrait être importante pour comprendre la hiérarchie de régulation des gènes de ce réseau. Nos résultats ouvrent ainsi de nouvelles perspectives dans la compréhension du rôle des gènes soumis à l'Empreinte dans le développement embryonnaire et dans les pathologies humaines liées à l'Empreinte parentale.

## Article 4

### *H19 and Igf2 imprinting relaxation as a marker for adult stem cells*

Manuscrit en préparation.

## **Introduction**

L'Empreinte Parentale est un mécanisme épigénétique qui conduit à l'expression monoallélique d'un petit groupe de gènes en fonction de l'origine parentale de l'allèle. Les gènes *H19* et *Igf2* ont été parmi les premiers gènes soumis à l'empreinte décrits. Ils constituent un locus soumis à l'empreinte et ont une expression opposée : *H19* étant exprimé à partir de l'allèle maternelle et *Igf2* à partir de l'allèle paternel. Le facteur *Igf2* a été impliqué dans la myogenèse (Charge and Rudnicki, 2004) et l'ARN *H19* avait été cloné dans la même banque de facteurs myogéniques que *MyoD* et appelé *MyoH* (Davis *et al.*, 1987). Par ailleurs, ces gènes sont induits, avec cinq autres gènes soumis à l'Empreinte, au cours du processus de régénération musculaire parallèlement à l'induction de la *Myogenin* (Yan *et al.*, 2003). Ces données nous ont conduit à poser la question du rôle des gènes *H19* et *Igf2* dans la myogenèse, en particulier chez l'adulte. La régénération musculaire, qui permet de réparer le muscle lésé en produisant de nouvelles fibres musculaires, est un processus de différenciation myogénique à partir des cellules souches musculaires, appelées cellules satellites (Le Grand and Rudnicki, 2007; Sambasivan and Tajbakhsh, 2007). Nous avons donc porté un intérêt particulier à ces cellules.

Peu de données sont disponibles quant à la régulation de l'Empreinte dans les cellules souches adultes. La première observation est que, en accord avec le cycle de l'Empreinte parentale, l'Empreinte est effacée dans les cellules germinales primordiales (PGC), qui sont les cellules souches gamétiques, et les gènes *H19* et *Igf2* s'y expriment de façon biallélique (Szabo and Mann, 1995a). La seconde donnée est que chez l'humain le gène *Igf2* s'exprime de façon monoallélique dans le sang circulant et biallélique dans la moelle osseuse, qui est riche en cellules souches hématopoïétiques (HSC) (Morison *et al.*, 2000). Par ailleurs, le gène *Igf2* est également exprimé à partir des deux allèles dans le sang de cordon (Giannoukakis *et al.*, 1993). Ces données obtenues dans les PGC et les HSC suggèrent que l'expression biallélique des gènes soumis à l'Empreinte pourrait être un marqueur de cellules souches adultes.

Nous avons utilisé la cellule satellite pour vérifier cette hypothèse dans un nouveau modèle de cellule souche adulte. Nous avons réalisé ces études avec trois lignées de souris : *H19*<sup>-/-</sup>, *Igf2*<sup>-/-</sup> et SD7, qui porte un fragment *spretus* sur la partie distale du chromosome 7 et contient donc des polymorphismes au niveau des gènes *H19* et *Igf2*. Les femelles *H19*<sup>-/-</sup> ont

été croisées avec des mâles sauvages afin d'obtenir des descendants hétérozygotes maternels  $H19^{-/+}$ , les mâles  $Igf2^{-/-}$  ont été croisés avec des femelles sauvages afin d'obtenir des descendants hétérozygotes paternels  $Igf2^{+/-}$ . L'expression respectivement du gène  $H19$  et  $Igf2$  dans les cellules satellites de ces souris hétérozygotes indique une expression à partir d'un allèle normalement réprimé et suggère donc une expression biallélique de ces gènes. Nous avons ensuite utilisé les animaux SD7 afin de nous affranchir d'un éventuel dialogue entre les allèles conduisant à une réactivation de l'allèle normalement réprimé. Par ailleurs, nous avons voulu étudier les cellules satellites dans un état de quiescence proche de celui qu'elles ont sur la fibre musculaire. Pour cela, en collaboration avec D. Montarras, nous avons croisé les souris SD7 avec des souris  $Pax3^{GFP}$ , ce gène étant un marqueur des cellules satellites et nous avons pu faire un tri par FACS pour obtenir une population quiescente pure. Nous avons ainsi observé dans ces cellules triées l'expression biallélique du gène  $Igf2$ , suggérant un relâchement de l'Empreinte dans les cellules satellites quiescentes.

## ***H19* and *Igf2* imprinting relaxation as a marker for adult stem cells**

Anne Gabory<sup>1</sup>, Georgia Pallafacchina<sup>2</sup>, Marie-Anne Ripoché<sup>1</sup>, Christel Picard<sup>3</sup>, Edith Heard<sup>3</sup>, Guy Cathala<sup>4</sup>, Thierry Forné<sup>4</sup>, Hélène Jammes<sup>1</sup>, Didier Montarras<sup>2</sup> and Luisa Dandolo<sup>1§</sup>

<sup>1</sup>Genetics and Development Department, Inserm U567, CNRS UMR 8104, University Paris Descartes, Institut Cochin, Paris, France

<sup>2</sup>Molecular Genetics of Development Department, URA CNRS 2578, Institut Pasteur, Paris, France

<sup>3</sup>Mammalian Developmental Epigenetics Group, CNRS UMR 218, Institut Curie, Paris, France

<sup>4</sup>Institut de Génétique Moléculaire, CNRS UMR 5535, University Montpellier II, Montpellier, France

§Corresponding author:

Luisa Dandolo

Department of Genetics and Development, Institut Cochin

24 rue Fbg St Jacques, Paris 75014, France

Phone 33 1 44 41 24 55

Fax 33 1 44 41 24 48

[dandolo@cochin.inserm.fr](mailto:dandolo@cochin.inserm.fr)

Running title: *H19-Igf2* Imprinting in Adult Stem Cells

Keywords: *H19*, *Igf2*, mouse models, genomic imprinting, satellite cells

14 pages, 4 Figures.

## Abstract

Genomic imprinting is an epigenetic mechanism which controls gene expression and leads to parent-of-origin specific expression. It is established during gametogenesis and is erased at each generation in the primordial germ cells, which are the gamete stem cells. Imprinting consists of a primary mark which is apposed on specific regulatory sequences called imprinting centres (ICR) and results in differential DNA methylation. This mark is then read in somatic tissues and leads to monoallelic expression of the imprinted genes. These genes are mostly involved in growth control and proliferation of the embryo. The *H19* and *Igf2* genes have been extensively studied to define the regulatory mechanisms linked with this epigenetic control. We investigated the imprinted status of these two genes in adult stem cells and chose to study myogenic stem cells, because there is evidence that these two genes play a role in myogenesis and muscle regeneration. Using different mouse models, we were able to show that expression of these two genes was biallelic in quiescent and activated satellite cells. Expression remains biallelic upon differentiation of these cells *in vitro*. We also showed that the *H19* ICR is differentially methylated in these stem cells, suggesting that the primary imprinting mark is present but that the reading is then relaxed or perhaps ignored in these undifferentiated cells.

## Introduction

Genomic imprinting is a complex multistep process that results in parent-of-origin dependent expression of one allele of certain genes. CpG methylation is one of the key features involved in mammalian imprinting, although little is known about the nature of the imprint and how it is recognized. Imprinted genes are characterised by differentially methylated regions (DMR), with some of these DMRs being considered as Imprinting Control Regions (ICR) such as the *H19* DMR in the *H19-Igf2* locus. Upon deletion of this ICR, both *H19* and *Igf2* expression are affected, showing that this 2kb region has a central role in the control of this imprinted locus (Thorvaldsen et al. 1998). This region has also been defined as the target for a primary imprint, set during gametogenesis and absolutely essential for maintaining the imprinted status in the embryo. The other DMRs which have been identified in this region are linked to the *Igf2* gene, DMR1 and 2, and are considered as secondary DMRs, with the establishment of their differential methylation occurring after fertilisation (Lopes et al. 2003).

The imprinting mark which is responsible for monoallelic expression must be erased at each generation in the primordial germ cells (PGCs), which are the gamete stem cells, and reset on each parental genome during gametogenesis (Ferguson-Smith et al. 2003). The erasure of imprinting has been defined by complete demethylation of DMR sequences and should result in biallelic expression of these genes. This has indeed been observed in PGCs for genes such as *Igf2* and its receptor *Igf2R* (Mann 2001).

Since imprinting is lacking in germ cells, the question of a possible link between stemness and imprinting was raised. Adult stem cells were thus analysed and it was shown that *Igf2* is biallelically expressed in human cord blood (Giannoukakis et al. 1996) and in haematopoietic stem cells (HSCs) of human bone marrow, whereas only paternal specific expression was observed in peripheral blood cells (Morison et al. 2000). The results obtained in PGCs and HSCs therefore suggested that lack (or absence) of imprinting is a potential general marker of stem cells.

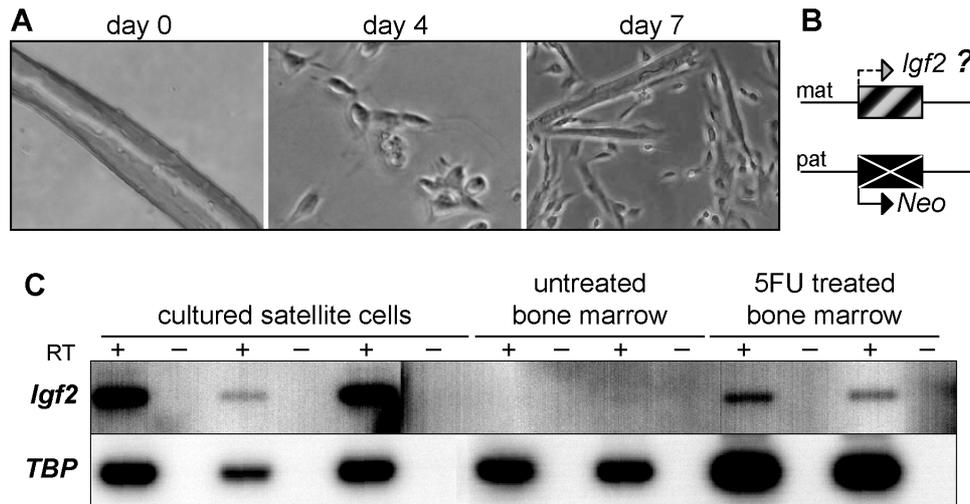
The *H19* gene produces a non-coding RNA and is exclusively expressed from the maternal allele (Gabory et al. 2006). It is strongly expressed throughout embryogenesis and is shut down after birth, with the exception of skeletal muscle (and heart) in which it remains expressed in adults (Poirier et al. 1991). *Igf2* (insulin-like growth factor-2) is a paternally expressed growth factor, highly expressed during embryogenesis (DeChiara et al. 1990). This gene is also involved in muscle growth and differentiation and is linked to the myogenic factors (MRFs) such as MyoD (Charge and Rudnicki 2004).

Because *H19* remains strongly expressed in skeletal muscle cells, we were interested in studying the expression of this *H19-Igf2* locus in satellite cells, considered as myogenic stem cells. These cells are present in the mouse embryo from day 16.5 on and are then found mostly after birth in two sets of populations. The first one appears around p2 to p7 and consists of self-renewing, dividing cells (Sambasivan and Tajbakhsh 2007). The second population is found in adults (or at least after p14). These cells are quiescent and are found on the myotubes (Beauchamp et al. 2000). These satellite cells will be activated and recruited in a regeneration process and participate in new formation of myoblasts and myotubes.

Interestingly, a microarray analysis performed on injured muscles undergoing regeneration and producing a burst of these satellite cells revealed an increase in the expression of several imprinted genes among which *H19* and *Igf2* (Yan et al. 2003). The question was raised if this was linked to simple induction of expression or to modification of some epigenetic marks in the satellite cells.

Our aim was to analyse the expression of these two genes in muscle satellite cells and to determine if biallelic expression (or lack of imprinting) could be a marker for adult stem cells. From several mouse crosses, either using null mutants of *H19* or *Igf2* or crosses with SD7 mice, we analysed allele specific expression and methylation status of these two imprinted genes in activated satellite cells as well as in quiescent satellite cells.

We concluded that imprinted expression of these genes is lost in satellite cells. Methylation analysis of the *H19-Igf2* locus showed that the *H19* ICR remained differentially methylated, thus suggesting that biallelic expression is linked to a lack of reading of the imprinted signals and not to a modification or an erasure of the primary imprint in these stem cells. Finally we also addressed the status of imprinting during differentiation of these cells in culture and found that absence of imprinting was maintained in culture.



### Figure 1

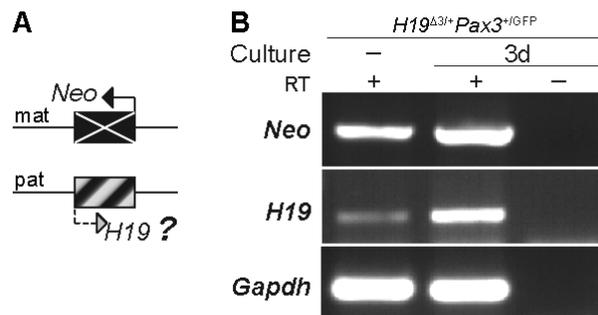
**A** Satellite cell culture obtained from individual muscle fibres. Fibres are incubated on Matrigel coated plates. By day 4, small round satellite cells can be seen. By day 7, these cells have differentiated into myotubes.

**B** Schematic representation of the Pat Het *Igf2*<sup>+/-</sup> alleles carrying a *neoR* gene inserted into the *Igf2* coding region.

**C** Maternal expression of *Igf2* in activated satellite cells.

RT-PCR was performed on RNA extracted from day 4 cultured satellite cells from three individual experiments using *Igf2* primers and TBP primers as control.

Bone marrow was collected from untreated and 5FU treated Pat Het mice (2 mice, one male and one female, for each sample) and the same RT-PCR was performed on RNA from these cells. The *Igf2* fragment which is detected is expressed from the normally silent maternal allele.



### Figure 2

**A** Schematic representation of the *H19*<sup>Δ3/+</sup> Mat Het alleles. The *neoR* gene (*Neo*) has replaced the *H19* 3kb transcription unit.

**B** Expression of maternal and paternal alleles in *Pax3*<sup>+GFP</sup> *H19*<sup>Δ3/+</sup> satellite cells. GFP+ FACS sorted cells (without culture) and a fraction of these cells cultured for 3 days (3d) were used for expression analysis. Detection of maternally expressed *neoR* gene and of the normally silent paternal *H19* allele was performed by RT-PCR on RNA from these two samples. *Gapdh* was used as a control.

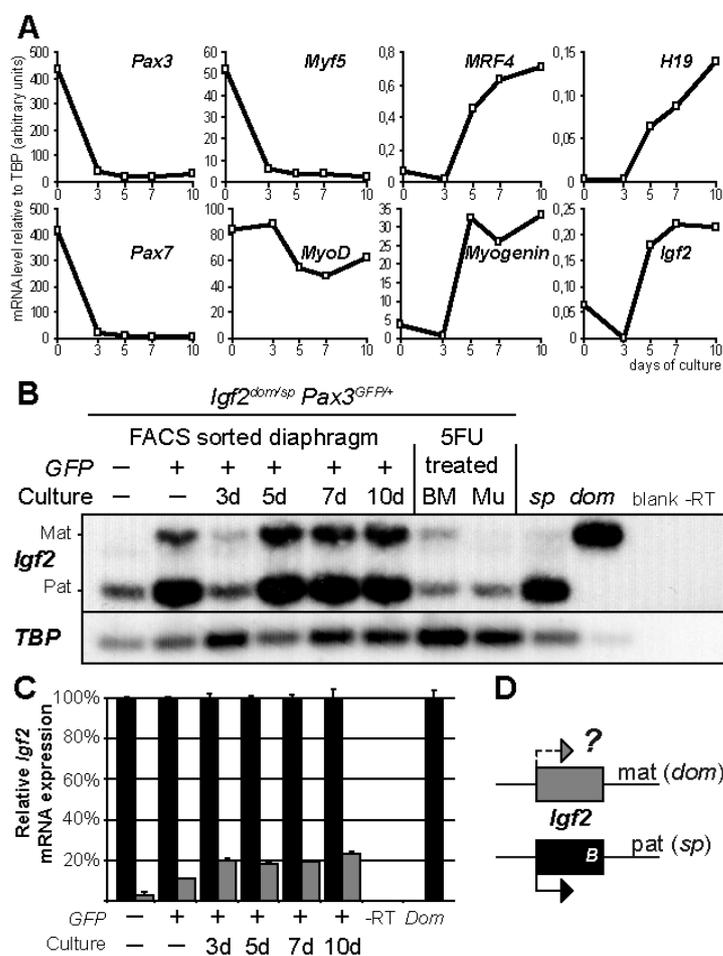
## Results

### *Igf2* expression in activated satellite cells

In order to investigate the allelic status of expression of imprinted genes, we started by establishing satellite cell cultures from heterozygous mouse crosses. Female *wt* mice were bred with male *Igf2*<sup>-/-</sup> mice to produce *Pat Het Igf2*<sup>+/-</sup> progeny. This allows to assess the parental origin of the transcripts, since *Igf2* mRNA can only be transcribed from the maternal allele (Figure 1B). Cell cultures were set up from isolated muscle fibres of these mice (Figure 1A). The muscles from the hind limb were dissected out and fibres were obtained after collagenase digestion. These fibres were put into culture on Matrigel coated plates and satellite cells dropping off of these fibres were collected after 4 days. These cells are defined as activated satellite cells. At day 7, these cells have differentiated into myotubes and a large number of these can be seen in the culture. These cells were collected, RNA was prepared and RT-PCR was performed to investigate the presence of (normally silent) maternal *Igf2* expression (Figure 1B). This expression was detected in these cultured satellite cells in three different experiments (Figure 1C). This was the first evidence for unusual biallelic expression of the imprinted *Igf2* gene in these satellite cells.

We also detected a high level of *H19* expression in these satellite cells, although the allelic origin could not be identified in this experiment. Since *H19* expression is essentially embryonic and is strongly down-regulated in adult mice except for skeletal muscle, this high expression in satellite cells is an indication for a potential role of *H19* in myogenesis.

In parallel, we injected 10 weeks-old *Igf2*<sup>+/-</sup> Pat Het mice with 5FU to enrich the bone marrow for undifferentiated haematopoietic stem cells. These cells were collected from hind limb bones and the same RT-PCR analysis was performed. Bone marrow cells from non injected adult mice were used as controls. In this non-enriched population, *Igf2* expression could not be detected (Figure 1C), showing that no expression arises from the maternal allele, and that paternal monoallelic expression is the only *Igf2* mRNA detected in *wt* context. In contrast, in the 5FU enriched population, the maternal *Igf2* allele can be detected. Thus *Igf2* is biallelically expressed in murine HSCs confirming the previously published data in human (Morison et al. 2000) The interpretation of these results in HSC would be that loss of biallelic expression occurs upon differentiation *in vivo*, and *Igf2* expression becomes monoallelic in fully differentiated cells.



### Figure 3

**A** Expression of myogenic markers. *Pax3*<sup>+GFP</sup> mice were crossed with SD7 mice. GFP<sup>+</sup> cells from the diaphragm of these mice were sorted by FACS. RNA was extracted either directly or after maintenance in culture for 3, 5, 7 and 10 days. Markers of muscle differentiation, as well as *H19* and *Igf2* were analysed for each time point by quantitative RT-PCR. *TBP* expression was used as control.

**B** Biallelic expression of *Igf2* in satellite cells. The paternal *Spretus* allele from the SD7 mice carries a *BsaA1* sequence polymorphism in the *Igf2* coding region. The same RNA samples were used for RT-PCR followed by *BsaA1* digestion. Progeny from the same cross were treated with 5-FU and RNA was prepared from bone-marrow cells from hind limbs as well as surrounding muscle. Analysis was performed on GFP<sup>-</sup> cells, on GFP<sup>+</sup> d0 cells, on GFP<sup>+</sup> cells maintained in culture (3 to 10d), on bone marrow cells (BM) and on adult muscle (M) in order to detect *Igf2* expression. *Sp* is a *Spretus* muscle RNA sample, *Dom* is a *domesticus* muscle RNA sample (5 day-neonate). The paternal (Pat) fragment corresponds to the *BsaA1* digested *Spretus* allele and the maternal (Mat) fragment is the normally silent allele.

**C** The same samples obtained after FACS sorting were analysed in parallel by quantitative allele-specific RT-PCR. Expression of the paternal allele is normalized to 100%, showing that there is a constant 20% level of maternal *Igf2* in the GFP<sup>+</sup> satellite cell samples. *Dom* corresponds to 5day-neonate muscle used as control for *domesticus* allele detection.

**D** Schematic representation of the maternal *domesticus* (*Dom*) and paternal *Spretus* *Igf2* alleles. B corresponds to the *Spretus* *BsaA1* polymorphism.

### **Imprinted expression in quiescent satellite cells**

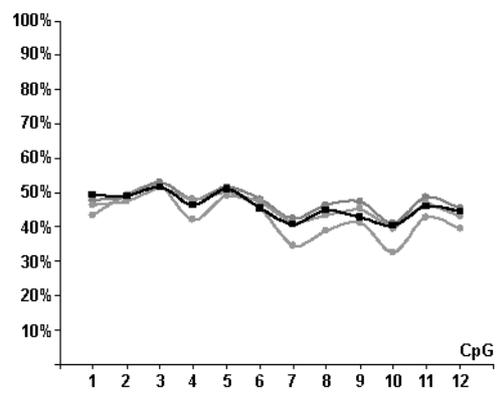
We then investigated if the absence of imprinting identified in activated satellite cells also exists in quiescent satellite cells. These cells are identified as *Pax3* and *Pax7* positive cells (Beauchamp et al. 2000; Buckingham 2006). Crosses were set up between *Igf2*<sup>-/-</sup> and *H19*<sup>-/-</sup> mice with *Pax3*<sup>+/*GFP*</sup> mice, thus allowing the possibility of a direct FACS selection of quiescent satellite cells labelled with GFP (Montarras et al. 2005).

In the experiment shown in Figure 2, female *H19*<sup>Δ3/+</sup> mutants were crossed with *Pax3*<sup>+/*GFP*</sup> males (Figure 2A). The *Pax3*<sup>+/*GFP*</sup> mutation is essentially expressed in satellite cells from the diaphragm and intercostal muscles. Diaphragms from the progeny of this cross were dissected and a cell suspension was prepared from which the *Pax3*<sup>+/*GFP*</sup> quiescent satellite cells were selected after FACS sorting. Cells were then put into culture as previously described (Montarras et al. 2005) and harvested after 3 days of culture. Cells collected directly after cell sorting are considered as quiescent, whereas after 3 days of culture, they are defined as activated satellite cells. Further myogenic differentiation is observed as culture is continued up to 10 days, when myotubes are clearly visible.

RT-PCR results (Figure 2B) show the expected expression of the neomycin resistance gene which was inserted into the *H19* locus to produce the deleted allele. Interestingly, the normally silent paternal allele of *H19* is also expressed in quiescent and 3 day activated cells. This result suggests that *H19* is also biallelically expressed in satellite cells.

### **Imprinted expression using *Spretus* polymorphisms**

Finally, since the use of knock-out mutants could introduce a possible bias in the expression patterns of *H19* or *Igf2*, we performed another cross between *Pax3*<sup>+/*GFP*</sup> females and males from the SD7 mouse strain, which carries polymorphisms of the distal region of chromosome 7 (Figure 3D). After FACS sorting and culture of the GFP positive cells up to 10 days, we performed real-time qRT-PCR on the samples from d0 to d10 and evaluated the relative levels of *Pax3*, *Pax7*, *myf5*, *MyoD*, *myogenin* and *MRF4*, as well as *H19* and *Igf2* (Figure 3A). As expected, *Pax3*, *Pax7* and *Myf5* decrease between d0 and d3, confirming that the cells which have been collected are quiescent satellite cells which are undergoing activation. Induction of *myogenin* and *MRF4* follow the differentiation pattern of these cells, whereas the levels of *MyoD* decrease progressively. Finally, *H19* expression increases with differentiation and *Igf2* is present in quiescent cells then repressed in activated satellite cells to then continue rising upon differentiation.



**Figure 4**

Methylation analysis of *H19* ICR. DNA samples from FACS sorted GFP+ satellite cells cultured for 5 day (grey curves) and control muscle (black curve) were bisulphite treated, PCR amplified and analysed by pyrosequencing for the presence of methylated CpGs.

We then evaluated the allelic expression of *Igf2* in these cells and were able to show absence of strict imprinting in these cells (Figure 3B). RT-PCR experiments followed by *BsaAI* digestion showed expression of both maternal and paternal alleles in GFP positive cells but not in GFP negative cells, nor in muscle samples (Mu) where only the paternal allele was detected. The reduced expression of *Igf2* at day 3 observed by qRT-PCR was confirmed and remains to be explained. 5FU treated bone marrow cells, enriched in HSC, confirmed our previous result showing biallelic expression in these stem cells.

Allele-specific RT-PCR was also performed to confirm and quantify precisely these results (Figure 3C). This was done with oligos carrying a single polymorphic nucleotide and resulted in allele specific amplification, as previously published (Weber et al. 2003). This quantification showed that the maternal *Igf2* allele was expressed between 10 and 20% of the full paternal allele. Interestingly, the biallelic status continues to be observed as differentiation proceeds.

### **Methylation status of the DMRs**

DNA methylation is a characteristic of imprinted domains: the *H19-Igf2* locus contains an imprinting control region or ICR which is differentially methylated on the paternal and maternal alleles and controls the imprinted status of the locus. This ICR contains binding sites for the boundary factor called CTCF, which can only bind to the unmethylated maternal allele (Hark et al. 2000). We investigated the pattern of methylation of this ICR in the quiescent satellite cells from the *Pax3<sup>+GFP</sup>* mice. This was done by bisulphite treatment of the DNA, followed by PCR amplification and pyrosequencing. The results showed no difference in the differentially methylated ICR sequences between the DNA extracted from the GFP positive quiescent satellite cells, from satellite cells cultured for 5 days or from control muscle (Figure 4). We therefore concluded that although relaxation of imprinting is observed at the level of expression, the primary imprint set on the *H19* ICR is not modified in these adult myogenic stem cells.

## Discussion

Genomic imprinting concerns a small number of genes involved in growth and proliferation control. This specific regulation is characterised by two important features: one is that the primary mark which is responsible for imprinting is set during gametogenesis and requires erasure of this mark at each new generation; the second is that this mark is then read by chromatin remodelling factors in somatic cells and results in monoallelic expression in differentiated tissues.

It was known that imprinting was modified in two populations of embryonic stem cells. One is the PGC population of gametic stem cells in which imprinting is totally erased to then be reapposed during gametogenesis (Szabo and Mann 1995b) and the second is the ES cell population derived from the early blastocyst, in which biallelic expression of certain imprinted genes was found (Szabo and Mann 1995a).

The focus of our study was to investigate the status of imprinting in adult stem cells. Previously published data had shown that *Igf2* was biallelically expressed in human HSCs (Morison et al. 2000). We performed experiments using mutant mouse models which allow to distinguish allelic origin of imprinted genes and were first able to show that biallelic expression of *Igf2* was also observed in murine HSCs.

We then set up experiments to investigate the status of imprinting in myogenic stem cells. Several imprinted genes, among which *H19* and *Igf2*, are expressed in muscle and in satellite cells, strongly suggesting that these genes, are involved in muscle differentiation and/or regeneration (Poirier et al. 1991; Yan et al. 2003). We therefore chose these two genes to investigate their allelic status in satellite cells. These experiments were done in two different stages, quiescent and activated satellite cells. We were able to show biallelic expression of *H19* and *Igf2* in both types of cells. This is the first time that loss (or relaxation) of imprinting has been identified in myogenic precursor cells. This result opens many perspectives about gene regulation and differentiation.

Interestingly, biallelic expression was observed to be at a level of 20% for the normally silent allele. This raises the question of identifying if this corresponds to all cells expressing only 20% of this allele, or if 20% of the cells in the selected population express 100% of each allele. This could be answered by either an analysis done on individual clones of satellite cells or by RNA FISH analysis on the FACS selected population of satellite cells. Obtaining this additional information may allow to identify two different populations in the global stem cell population, as has been suggested recently (Kuang et al. 2006).

In order to investigate if this a general feature of imprinted genes or strictly restricted to the *H19-Igf2* locus, it would be of interest to investigate if other imprinted genes are also biallelically expressed and more particularly those found overexpressed in the satellite cells, such as *Cdkn1c*, *Mest*, *Peg3* or *Zac1* (Yan et al. 2003).

We then investigated the methylation status of the ICR of the *H19-Igf2* locus. This ICR is considered as the sequence responsible for the primary imprint which is set during gametogenesis. This primary imprint resists the wave of demethylation which occurs during early development and remains constantly differentially methylated. Other DMRs such as the *Igf2* DMR1 and 2 are considered as secondary imprinted marks, and their differential methylation is often only observed after fertilisation or even after implantation (Lopes et al. 2003). Interestingly, the ICR was differentially methylated, suggesting that the primary imprint is present in these multipotent adult stem cells. Our results imply that the primary imprinting mark although present, is ignored in these cells, and leads to absence of reading and thus biallelic expression of *H19*. If the *Igf2* DMRs are not differentially methylated, this would confirm that these DMRs are linked to the reading procedure occurring in somatic cells.

One interesting observation was that biallelic expression was maintained during culture of these cells. This suggests that potential factors involved in resetting imprinting are missing in cultured cells compared to *in vivo* situations. It would be of interest to investigate the status of expression of chromatin remodelling factors such as the Dnmt proteins or members of the Polycomb/ Trithorax family of epigenetic regulators.

Dynamics which are required to restrict expression to monoallelism must not be present *in vitro*, whereas they must exist as differentiation proceeds *in situ* in the organs. Cellular therapy experiments have been proposed using satellite cells and it would be important to ensure that genes are correctly regulated after retransplantation of these cells in injured muscles.

In summary, we have been able to show for the first time that two imprinted genes are not strictly imprinted in quiescent and activated adult satellite cells. Relaxation or loss of imprinting could potentially be used as a marker of adult stem cells with the possibility of being able to maintain these cells in constant renewal or at least enrich for stem cell populations.

## Materials and Methods

### Mice strains and genotyping

The *H19*<sup>A3</sup> strain carries a 3kb deletion of the *H19* transcription unit and was initially established on a C57BL/6/CBA background (Ripoche et al. 1997). The *Igf2*<sup>-/-</sup> strain was a kind gift of A. Efstratiadis (DeChiara et al. 1990). The *Pax3* *GFP* mice were a kind gift of F. Relaix and M. Buckingham (Relaix et al. 2005). The *SD7* strain is a C57BL/6/CBA strain carrying the distal part of *Mus spretus* chromosome 7 and was a kind gift from W. Reik.

The protocol of animal handling and treatment was performed in accordance with the guidelines of the animal ethics committee of the Ministère de l'Agriculture of France.

### Activated satellite cell culture

To prepare muscle fibres, *tibialis*, *EDL* and *soleus* muscles were collected from 8-10 weeks old mice. The muscles were dissociated in 2% Collagenase D (Roche) in DMEM at 37°C for 60 min. Intact fibres were gently disrupted using a Pasteur pipette and rinsed several times in DMEM medium. 100 to 200 fibres/ cm<sup>2</sup> were cultured on Matrigel coated plates in DMEM medium supplemented with 20% FCS, 2% Ultrosor (Biosepra) and a cocktail of insulin-transferrin-selenium (Gibco) (Rosenblatt et al. 1995). After 4 days of culture, round satellite cells have dropped off from the fibres and are visible as small colonies. After 7 days of culture, myotubes can be seen in the plates.

### Quiescent satellite cells

*H19*<sup>A3/+</sup> or *SD7* females were bred with *Pax3*<sup>+/*GFP*</sup> mice. Diaphragms from 6 to 8 weeks old mice were dissected and prepared as previously described (Montarras et al. 2005).

### RNA preparation and analysis

Total RNA was extracted from cells using a RNeasy microkit (Qiagen). For RT-PCR analysis, DNase I treated RNA (0.5mg) was reverse-transcribed with Superscript II and random hexamer primers (Invitrogen). 5-10 ng of cDNA were then amplified using gene-specific primers and Taq DNA polymerase (Invitrogen) during 35 cycles.

Detection of *Igf2* transcripts derived from the *SD7* cross was performed using primers: forward 5'-GACGTGTCTACCTCTCAGGCCGTA CTT-3' and reverse 5'-GGGTGTCAATTGGGTTGTTTAGAGCCA-3'. The 517 bp product was digested with *BsaA1*, yielding a 241 bp paternal fragment detected by the internal primer 5'-

TCAAATTTGGTTTTTTAGAA-3'. RT-PCR products were separated on 1% agarose gels and transferred onto Hybond N+ membranes in 0.4M NaOH. Blots were probed with  $\gamma$ P<sup>32</sup>ATP kinase-labelled primers at 42°C in Church buffer. Membranes were washed in 0.4xSSC, 0.5% SDS at 42°C and results were quantified using Phosphorimager analysis and ImageQuant software.

Quantitative real-time PCR was performed using a Light Cycler system and Sybr Green PCR kits (Roche Diagnostics, Mannheim, Germany). 1 to 5 ng of cDNA were amplified in duplicate using primers for *H19*, *Igf2*, *Pax3*, *Pax7*, *MyoD*, *myf5*, *myogenin* and *MRF4*. Relative expression was normalised to the expression of *GAPDH* and *TBP*.

### **Methylation analysis**

Genomic DNA (1 µg) was treated with sodium bisulphite as previously described (Dupont et al. 2004). Bisulphite-treated DNA was used as PCR template. Pyrosequencing was done using the pyrosequencing PSQ 96Masystem with the PyroGold SQA reagent kit according to manufacturer's instructions (Pyrosequencing AB, Uppsala, Sweden) (Tost et al. 2006).

## Figure Legends

### Figure 1

**A** Satellite cell culture obtained from individual muscle fibres. Fibres are incubated on Matrigel coated plates. By day 4, small round satellite cells can be seen. By day 7, these cells have differentiated into myotubes.

**B** Schematic representation of the Pat Het *Igf2*<sup>+/-</sup> alleles carrying a *neoR* gene inserted into the *Igf2* coding region.

**C** Maternal expression of *Igf2* in activated satellite cells.

RT-PCR was performed on RNA extracted from day 4 cultured satellite cells from three individual experiments using *Igf2* primers and TBP primers as control.

Bone marrow was collected from untreated and 5FU treated Pat Het mice (2 mice, one male and one female, for each sample) and the same RT-PCR was performed on RNA from these cells. The *Igf2* fragment which is detected is expressed from the normally silent maternal allele.

### Figure 2

**A** Schematic representation of the *H19*<sup>A3/+</sup> Mat Het alleles. The *neoR* gene (*Neo*) has replaced the *H19* 3kb transcription unit.

**B** Expression of maternal and paternal alleles in *Pax3*<sup>+/*GFP*</sup> *H19*<sup>A3/+</sup> satellite cells. GFP+ FACS sorted cells (without culture) and a fraction of these cells cultured for 3 days (3d) were used for expression analysis. Detection of maternally expressed *neoR* gene and of the normally silent paternal *H19* allele was performed by RT-PCR on RNA from these two samples. *Gapdh* was used as a control.

### Figure 3

**A** Expression of myogenic markers. *Pax3*<sup>+/*GFP*</sup> mice were crossed with SD7 mice. GFP+ cells from the diaphragm of these mice were sorted by FACS. RNA was extracted either directly or after maintenance in culture for 3, 5, 7 and 10 days. Markers of muscle differentiation, as well as *H19* and *Igf2* were analysed for each time point by quantitative RT-PCR. *TBP* expression was used as control.

**B** Biallelic expression of *Igf2* in satellite cells. The paternal *Spretus* allele from the SD7 mice carries a *BsaAI* sequence polymorphism in the *Igf2* coding region. The same RNA samples

were used for RT-PCR followed by *BsaAI* digestion. Progeny from the same cross were treated with 5-FU and RNA was prepared from bone-marrow cells from hind limbs as well as surrounding muscle. Analysis was performed on GFP- cells, on GFP+ d0 cells, on GFP+ cells maintained in culture (3 to 10d), on bone marrow cells (BM) and on adult muscle (M) in order to detect *Igf2* expression. *Sp* is a *Spretus* muscle RNA sample, *Dom* is a *domesticus* muscle RNA sample (5 day-neonate). The paternal (Pat) fragment corresponds to the *BsaAI* digested *Spretus* allele and the maternal (Mat) fragment is the normally silent allele.

**C** The same samples obtained after FACS sorting were analysed in parallel by quantitative allele-specific RT-PCR. Expression of the paternal allele is normalized to 100%, showing that there is a constant 20% level of maternal *Igf2* in the GFP+ satellite cell samples. *Dom* corresponds to 5day-neonate muscle used as control for *domesticus* allele detection.

**D** Schematic representation of the maternal *domesticus* (*Dom*) and paternal *Spretus Igf2* alleles. B corresponds to the *Spretus BsaAI* polymorphism.

#### **Figure 4**

Methylation analysis of *H19* ICR. DNA samples from FACS sorted GFP+ satellite cells cultured for 5 day (grey curves) and control muscle (black curve) were bisulphite treated, PCR amplified and analysed by pyrosequencing for the presence of methylated CpGs.

## Bibliography

- Beauchamp, J.R., Heslop, L., Yu, D.S., Tajbakhsh, S., Kelly, R.G., Wernig, A., Buckingham, M.E., Partridge, T.A., and Zammit, P.S. 2000. Expression of CD34 and Myf5 defines the majority of quiescent adult skeletal muscle satellite cells. *J Cell Biol* **151**(6): 1221-1234.
- Buckingham, M. 2006. Myogenic progenitor cells and skeletal myogenesis in vertebrates. *Curr Opin Genet Dev* **16**(5): 525-532. .
- Charge, S.B. and Rudnicki, M.A. 2004. Cellular and molecular regulation of muscle regeneration. *Physiol Rev* **84**(1): 209-238.
- DeChiara, T.M., Efstratiadis, A., and Robertson, E.J. 1990. A growth deficiency phenotype in heterozygous mice carrying an Insulin-like growth factor II gene disrupted by gene targeting. *Nature* **345**: 78-80.
- Dupont, J.M., Tost, J., Jammes, H., and Gut, I.G. 2004. De novo quantitative bisulfite sequencing using the pyrosequencing technology. *Anal Biochem* **333**(1): 119-127.
- Ferguson-Smith, A., Lin, S.P., Tsai, C.E., Youngson, N., and Tevendale, M. 2003. Genomic imprinting-insights from studies in mice. *Semin Cell Dev Biol* **14**(1): 43-49.
- Gabory, A., Ripoché, M.A., Yoshimizu, T., and Dandolo, L. 2006. The H19 gene: regulation and function of a non-coding RNA. *Cytogenet Genome Res* **113**(1-4): 188-193.
- Giannoukakis, N., Deal, C., Paquette, J., Kukuvtis, A., and Polychronakos, C. 1996. Polymorphic functional imprinting of the human IGF2 gene among individuals, in blood cells, is associated with H19 expression. *Biochem Biophys Res Commun* **220**(3): 1014-1019.
- 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**(6785): 486-489.
- Kuang, S., Charge, S.B., Seale, P., Huh, M., and Rudnicki, M.A. 2006. Distinct roles for Pax7 and Pax3 in adult regenerative myogenesis. *J Cell Biol* **172**(1): 103-113.
- Lopes, S., Lewis, A., Hajkova, P., Dean, W., Oswald, J., Forne, T., Murrell, A., Constancia, M., Bartolomei, M., Walter, J., and Reik, W. 2003. Epigenetic modifications in an imprinting cluster are controlled by a hierarchy of DMRs suggesting long-range chromatin interactions. *Hum Mol Genet* **12**(3): 295-305.
- Mann, J.R. 2001. Imprinting in the germ line. *Stem Cells* **19**(4): 287-294.
- Montarras, D., Morgan, J., Collins, C., Relaix, F., Zaffran, S., Cumano, A., Partridge, T., and Buckingham, M. 2005. Direct isolation of satellite cells for skeletal muscle regeneration. *Science* **309**(5743): 2064-2067. Epub 2005 Sep 20061.
- Morison, I.M., Eccles, M.R., and Reeve, A.E. 2000. Imprinting of insulin-like growth factor 2 is modulated during hematopoiesis. *Blood* **96**(9): 3023-3028.
- Poirier, F., Chan, C.-T.J., Timmons, P.M., Robertson, E.J., Evans, M.J., and Rigby, P.W.J. 1991. The murine H19 gene is activated during embryonic stem cell differentiation *in vitro* and at the time of implantation in the developing embryo. *Development* **113**: 1105-1114.
- Relaix, F., Rocancourt, D., Mansouri, A., and Buckingham, M. 2005. A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature* **435**(7044): 948-953.
- Ripoché, M.A., Kress, C., Poirier, F., and Dandolo, L. 1997. Deletion of the H19 transcription unit reveals the existence of a putative imprinting control element. *Genes Dev* **11**: 1596-1604.
- Rosenblatt, J.D., Lunt, A.I., Parry, D.J., and Partridge, T.A. 1995. Culturing satellite cells from living single muscle fiber explants. *In Vitro Cell Dev Biol Anim* **31**(10): 773-779.
- Sambasivan, R. and Tajbakhsh, S. 2007. Skeletal muscle stem cell birth and properties. *Semin Cell Dev Biol* **18**(6): 870-882.
- Szabo, P.E. and Mann, J.R. 1995a. Allele-specific expression and total expression levels of imprinted genes during early mouse development: Implications for imprinting mechanisms. *Genes Dev* **9**(24): 3097-3108.
- Szabo, P.E. and Mann, J.R. 1995b. Biallelic expression of imprinted genes in the mouse germ line: Implications for erasure, establishment, and mechanisms of genomic imprinting. *Genes Dev* **9**(15): 1857-1868.

- 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 Dev* **12**(23): 3693-3702.
- Tost, J., El Abdalaoui, H., and Gut, I.G. 2006. Serial pyrosequencing for quantitative DNA methylation analysis. *Biotechniques* **40**(6): 721-722, 724, 726.
- Weber, M., Hagege, H., Lutfalla, G., Dandolo, L., Brunel, C., Cathala, G., and Forne, T. 2003. A real-time polymerase chain reaction assay for quantification of allele ratios and correction of amplification bias. *Anal Biochem* **320**(2): 252-258.
- Yan, Z., Choi, S., Liu, X., Zhang, M., Schageman, J.J., Lee, S.Y., Hart, R., Lin, L., Thurmond, F.A., and Williams, R.S. 2003. Highly coordinated gene regulation in mouse skeletal muscle regeneration. *J Biol Chem* **278**(10): 8826-8836.

## **Conclusions**

Nous avons tout d'abord observé l'expression biallélique du gène *Igf2* dans des cellules de moelle de souris traité à la 5FU, traitement qui permet d'enrichir la moelle en cellules souches hématopoïétiques (HSC). Ces résultats confirment ceux précédemment obtenus chez l'humain (Morison *et al.*, 2000). Nous avons également observé une expression biallélique des gènes *H19* et *Igf2* dans les cellules satellites quiescentes et activées par la culture alors qu'elle est monoallélique dans les tissus musculaires différenciés. De plus nous avons observé que l'expression biallélique du gène *Igf2* persiste avec la différenciation des cellules musculaires *in vitro*. Par ailleurs, une analyse de la méthylation du centre d'Empreinte du locus *H19/Igf2* a montré que celui-ci présente bien une méthylation différentielle. Dans l'ensemble, ces résultats avec ceux précédemment obtenus dans les PGC et les HSC, nous conduisent à proposer que le relâchement de l'Empreinte du gène *Igf2* peut être un marqueur de cellules satellites.

Il serait important de vérifier cette proposition dans d'autres cellules souches adultes. Par ailleurs, le taux d'expression de l'allèle normalement réprimé est de 20% par rapport à l'allèle normalement exprimé. Cette observation peut être due à 20% de cellules exprimant de façon équivalente les deux allèles, ce qui serait en conformité avec l'hypothèse qu'il existe différentes populations de cellules satellites. Il est également possible que toutes les cellules satellites expriment l'allèle maternel du gène *Igf2* à un taux de 20%, ce qui indiquerait plutôt un relâchement de la répression de cet allèle spécifique des cellules satellites.

L'observation que l'expression du gène *Igf2* ne devient pas monoallélique avec la différenciation myogénique *in vitro* est en accord avec d'autres études montrant que la manipulation des cellules, notamment la culture peut altérer le contrôle des gènes soumis à l'empreinte. Par ailleurs, la dérégulation de nombreux gènes, comme *H19* et *Igf2*, a été observée dans différents types de tumeurs. Aussi, il serait intéressant de vérifier que l'empreinte peut être rétablie à partir de ces cellules *in vivo*.

Nous nous proposons de compléter ces données par :

- 1) une étude détaillée de la méthylation de toutes les DMR du locus *H19/Igf2* par traitement de l'ADN au bisulfite et pyroséquençage ; l'ICR qui est une DMR gamétique et les DMR1, DMR2, et le promoteur du gène *H19* qui sont des DMR secondaires.
- 2) l'extension de notre étude par l'analyse de l'expression allélique et de l'état de méthylation des gènes soumis à l'Empreinte également présents dans la partie distale du chromosome 7 murin (à l'aide de la lignée SD7) : *Cdkn1c*, *Kvlqt1*, *Lit1* ... Ceci permettrait de mettre en évidence que l'absence d'Empreinte est une caractéristique générale des gènes soumis à l'Empreinte et non restreinte aux seuls gènes *H19* et *Igf2*.
- 3) l'évaluation de la capacité d'établissement de l'Empreinte dans les cellules satellites retransplantées. Nous nous proposons de sélectionner des cellules satellites SD7 *Pax3<sup>GFP</sup>* et de les réinjecter dans le muscle lésé de souris *Igf2<sup>-/-</sup>*. Nous pourrions ainsi suivre l'expression mono ou biallélique des cellules satellites qui vont participer à la réparation du muscle lésé.

Les données obtenues à partir de toutes ces expériences pourraient aboutir à une meilleure identification des cellules souches adultes, à une meilleure compréhension des mécanismes d'établissement et de maintien de l'Empreinte parentale dans ces cellules, ainsi qu'à une meilleure connaissance de la biologie des cellules souches adultes en général qui, à terme, seront très certainement utilisées en thérapie cellulaire.

# *Discussion*

## 1) Un rôle de régulateur d'un réseau pour le gène *H19*.

Le gène *H19* est atypique pour différentes raisons. Tout d'abord, ce gène est très abondamment exprimé au cours de l'embryogenèse ; il l'est d'ailleurs à un niveau comparable à la  $\beta$ -actine en hybridation *in situ* et à un niveau très supérieur aux autres gènes étudiés dans nos études de PCR quantitative (Poirier *et al.*, 1991). Par ailleurs, ce gène produit un ARN non codant ; il ne possède pas de cadre ouvert de lecture conservé (Brannan *et al.*, 1990). De plus, ce fut l'un des premiers gènes décrit comme étant soumis à l'Empreinte parentale, il est en effet d'expression monoallélique maternelle (Bartolomei *et al.*, 1991). Enfin, le gène *H19* est voisin du gène *Igf2* sur les génomes murins et humains, ce qui a constitué le premier locus soumis à l'Empreinte décrit (Zemel *et al.*, 1992). Pour ces différentes raisons, plusieurs laboratoires se sont intéressés à *H19* et *Igf2* et plus d'une trentaine de modèles murins ont été établis pour déchiffrer la régulation de l'Empreinte et les fonctions de ces gènes (Arney, 2003). D'autre part l'implication de ces gènes dans la Pathologie, particulièrement dans les tumeurs de Wilms et le syndrome de Beckwith-Wiedemann, a permis de bâtir l'hypothèse d'un rôle suppresseur de tumeur pour le gène *H19* (Hao *et al.*, 1993; Cui *et al.*, 1997).

Notre laboratoire s'est attaché à étudier la fonction de l'ARN *H19*, partant de l'hypothèse qu'un ARN aussi abondamment transcrit au cours de l'embryogenèse devait avoir un rôle important pour le développement. Un modèle murin d'inactivation du gène par recombinaison homologue chez la souris a été développé (Ripoche *et al.*, 1997). Malgré la forte expression embryonnaire, les souris invalidées pour le gène *H19* sont viables et fertiles et présentent un phénotype subtil de surcroissance somatique de 8% par rapport aux animaux sauvages. En raison de l'hypothèse d'un rôle suppresseur de tumeur, nous avons observé attentivement ces souris *H19*<sup>-/-</sup> mais nous n'avons jamais détecté de développement de tumeurs spontanées. La comparaison avec un autre modèle murin invalidé pour une plus grande région génomique autour du gène *H19* (Leighton *et al.*, 1995) a permis de mettre en évidence l'implication de ce domaine dans la régulation de l'Empreinte du gène *Igf2*. Malheureusement ces modèles n'ont pas permis d'apporter de réponses claires quant au rôle du gène *H19*.

A mon arrivée au laboratoire, deux projets ont été développés. Le premier a porté sur l'étude du rôle suppresseur de tumeur du gène *H19*. La littérature a créé une controverse autour de ce gène, proposant à la fois un rôle de suppresseur de tumeur mais aussi un rôle oncogénique (Hao *et al.*, 1993; Cui *et al.*, 1997; Berteaux *et al.*, 2005; Barsyte-Lovejoy *et al.*,

2006; Matouk *et al.*, 2007). Les différences observées dans les diverses études peuvent s'expliquer par les modèles et les techniques utilisés. L'originalité de notre étude a été de proposer trois modèles murins, pour lesquels les tumeurs sont induites *in vivo*. Nous avons considéré que l'animal était un contexte plus physiologique que les cultures cellulaires et, dans cette optique, était plus proche de l'humain développant un cancer. Grâce à ces trois modèles, tératocarcinomes, hépatocarcinomes et tumeurs coliques, nous avons pu corrélérer l'absence d'expression du gène *H19* avec l'augmentation de la croissance, de la vitesse d'apparition et du nombre des tumeurs. Ces résultats sont donc vraiment en faveur d'un rôle suppresseur de tumeur pour le gène *H19*.

Le second projet, que j'ai mené plus particulièrement, a porté sur l'étude de la fonction du gène *H19* au cours du développement de la souris. Dans ce but, des animaux transgéniques, exprimant *H19* de façon ectopique, ont été développés. Cette construction génétique nous permettait de nous affranchir d'un effet en *cis*, détectée chez les KO, pour observer un effet en *trans*, c'est à dire de l'ARN *H19* lui-même. L'utilisation des séquences régulatrices du gène *Necdin* pour contrôler l'expression du transgène a été choisie dans l'optique d'exprimer *H19* dans un tissu où il ne l'est pas normalement : le cerveau. Quatre lignées ont été développées. Malheureusement, l'une d'entre elle avait une expression mosaïque et deux autres avaient une expression du transgène très faible. Aussi, seule une lignée (appelée lignée Tg) a été retenue pour notre étude. Nous avons par ailleurs utilisé la lignée YAC déjà établie dans le laboratoire d'Azim Surani (Ainscough *et al.*, 2000) afin d'avoir une seconde lignée qui nous permettrait d'exclure un effet insertionnel.

Une première observation que nous avons faite concerne la régulation de l'expression du gène *H19*. En effet, dans nos lignées, le transgène placé sur fond *H19*<sup>-/+</sup> a un taux d'expression de 75% pour la lignée Tg et similaire à l'endogène d'une souris sauvage pour la lignée YAC. En revanche, sur un fond *H19*<sup>+/+</sup>, la surexpression de l'ARN *H19* atteint au maximum 30%. Des résultats impliquant une dérégulation de l'endogène en présence d'un transgène ont déjà été rapportés pour le gène *Igf2* (Sun *et al.*, 1997) : dans les cellules ES, les transgènes sont exprimés au même niveau que l'endogène. Mais, au cours de l'embryogenèse, les transgènes *Igf2* sont réprimés au cours du développement et l'endogène est suractivé de façon concomitante à la déméthylation de la DMR2 du transgène et à l'hyperméthylation de l'endogène. Ces données sont en faveur de la titration d'un répresseur par l'allèle transgénique qui favoriserait l'expression de l'allèle endogène. Dans notre lignée Tg, il n'est pas possible de distinguer l'allèle endogène de l'allèle transgénique. Par ailleurs, un biais est observé dans la

	Expression	KO : phénotype de croissance	Fonction	Référence
<i>Dlk1</i>	Pat	+/- : retard de croissance	Protéine Delta-like transmembraire et sécrétée, famille EGF-like.	(Moon <i>et al.</i> , 2002)
<i>Gnas</i>	Pat	+/- : retard de croissance	Sous-unité G $\alpha$ s des protéines G	(Yu <i>et al.</i> , 1998)
<i>Igf2</i>	Pat	+/- : retard de croissance	Facteur de croissance embryonnaire	(DeChiara <i>et al.</i> , 1990)
<i>Mest</i>	Pat	+/- : retard de croissance	Enzyme hydrolase potentielle	(Lefebvre <i>et al.</i> , 1998)
<i>Peg3</i>	Pat	+/- : retard de croissance	Protéine à doigt de zinc (ZnF)	(Li <i>et al.</i> , 1999)
<i>Rtl1</i>	Pat	Pas de phénotype de croissance	Protéine retrotransposon-like	(Sekita <i>et al.</i> , 2008)
<i>Slc38a4</i>	Pat	?	Transporteur d'acides aminés	-
<i>Cdkn1c</i>	Mat	Pas de phénotype de croissance	Inhibition du cycle cellulaire	(Zhang <i>et al.</i> , 1997)
<i>Dcn</i>	Mat	Pas de phénotype de croissance	Protéoglycane	(Danielson <i>et al.</i> , 1997)
<i>Grb10</i>	Mat	-/+ : surcroissance disproportionnée	Protéine adaptatrice dans la voie IGF	(Charalambous <i>et al.</i> , 2003)
<i>Igf2R</i>	Mat	-/+ : surcroissance	Clairance d'Igf2	(Wang <i>et al.</i> , 1994)

### Tableau 6 : Effet de l'inactivation de gènes soumis à l'Empreinte sur la croissance.

Phénotype de croissance des animaux invalidés pour un des gènes appartenant au réseau de gènes soumis à l'Empreinte Parentale et régulés par le gène *H19* dans notre étude.

transmission de ce génotype : la proportion de souris transgéniques sur fond sauvage est très inférieure aux 25% attendus. Ces résultats s'ajoutent à de précédentes observations qui impliquaient un contrôle de l'expression du gène et proposaient un effet délétère d'une expression excessive du gène (Yoo-Warren *et al.*, 1988; Ainscough *et al.*, 2000). Aussi, une première cible de l'ARN *H19* pourrait être le gène *H19* lui-même. De ce fait, nous avons étudié l'effet du transgène sur fond *H19*<sup>-/+</sup> en le comparant avec les souris *H19*<sup>-/+</sup>.

Au niveau phénotypique, l'expression ectopique du gène *H19* dans le cerveau n'entraîne visiblement pas de modification, bien que nous n'ayons pas réalisé d'analyse comportementale poussée, suggérant que l'absence d'expression reflète une absence de fonction dans cet organe. Nous avons en revanche observé un phénotype de réduction de croissance en miroir avec le phénotype des animaux *H19*<sup>-/+</sup> (+8%), les animaux *H19*<sup>-/+</sup>Tg ayant une réduction de taille et de poids. Enfin, l'expression du transgène YAC sur fond *H19*<sup>-/+</sup> permet de récupérer une courbe de croissance superposable à celle des sauvages.

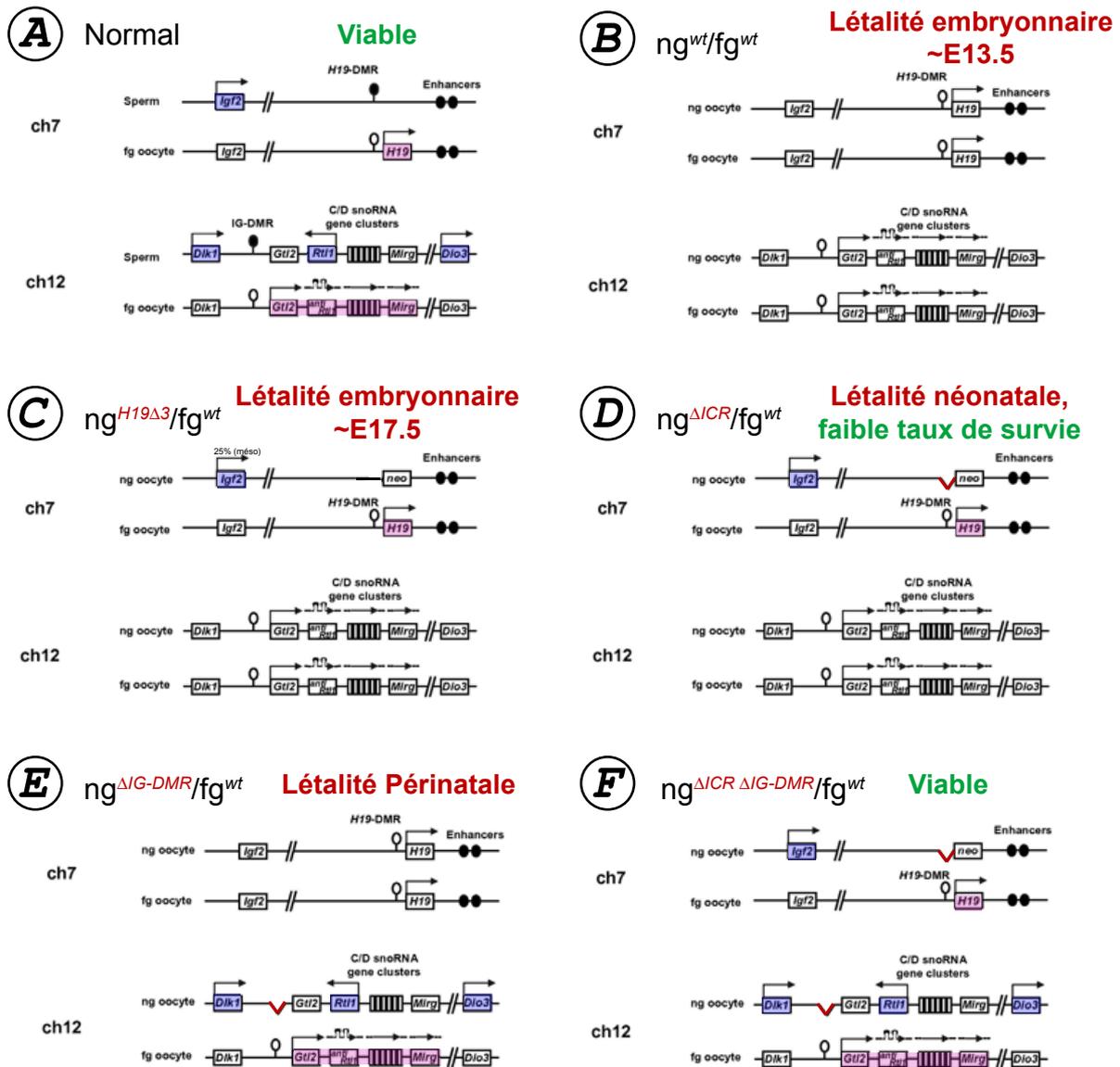
Nous avons dans un premier temps analysé l'expression du gène *Igf2* et observé une régulation négative du niveau d'ARNm *Igf2* par l'ARN *H19* dans notre modèle. La publication d'un réseau de gènes soumis à l'Empreinte (IGN pour Imprinted Gene Network) auquel appartiennent les gènes *H19* et *Igf2* (Varrault *et al.*, 2006) nous a conduit à élargir notre analyse, en collaboration avec l'équipe de Laurent Journot, aux gènes de ce réseau. Pour cette étude, l'obtention des trois génotypes d'intérêt (sauvage, *H19*<sup>-/+</sup> et *H19*<sup>-/+</sup> transgénique *H19*) dans une même portée nous a permis d'observer un sauvetage du phénotype des KO. En effet, onze gènes du réseau sont surexprimés dans les souris *H19*<sup>-/+</sup> par rapport aux souris sauvages et leur expression est restaurée chez les souris *H19*<sup>-/+</sup>Tg<sup>-/+</sup> et *H19*<sup>-/+</sup>YAC<sup>+/-</sup>. Parmi ces onze gènes, sept sont d'expression paternelle : *Igf2*, *Dlk1*, *Rtl1*, *Gnas*, *Mest*, *Peg3* et *Slc38a4* ; les cinq premiers ont un effet positif sur la croissance somatique au vue de l'analyse des souris KO (Tableau 6). Quatre gènes sont d'expression maternelle : *Cdkn1c*, *Grb10*, *Igf2R* et *Dcn* ; les trois premiers ont un effet négatif sur la croissance somatique et la prolifération cellulaire. De façon intéressante, cette régulation n'a pas été observée dans le placenta. Ces données nous permettent de **proposer que l'ARN *H19* est un *trans* régulateur du réseau de gènes soumis à l'Empreinte dans l'embryon.**

## **1-1) Implications de ces résultats**

### **1-1-1) Confirmation de l'existence d'un réseau fonctionnel**

Notre modèle était idéal pour étudier le réseau de gènes soumis à l'Empreinte. Ce réseau a été bâti sur une analyse bioinformatique des données de micro-arrays. La première preuve fonctionnelle a été apportée par l'analyse des mutants du gène *Zac1*, dans lesquels quatre gènes sont significativement modulés : *H19*, *Igf2*, *Cdkn1c* et *Dlk1* (Varrault *et al.*, 2006). Le facteur de transcription *Zac1* se fixe aux enhanceurs endodermiques communs des gènes *H19* et *Igf2* pour accroître leur expression. Un autre article récent montre que onze gènes de ce réseau sont réprimés simultanément après la naissance dans plusieurs organes de façon coordonnée avec le ralentissement de la croissance somatique, suggérant un lien de cause à effet (Lui *et al.*, 2008). Nos résultats apportent des évidences supplémentaires pour l'existence fonctionnelle de ce réseau. L'ARN *H19* est capable de contrôler négativement l'expression de plusieurs gènes de l'IGN. De façon intéressante cela touche à la fois des gènes qui favorisent la croissance et des gènes qui la répriment. Ces résultats sont concordants avec la théorie de l'utilité d'un tel réseau : la compensation des effets de l'altération de l'expression de l'un de ces gènes par les autres permet de maintenir une homéostasie globale, notamment en ce qui concerne la croissance. La surcroissance somatique de 8% des animaux *H19*<sup>+/+</sup> serait donc la résultante de la dérégulation de ces gènes et expliquerait le phénotype modeste.

Par ailleurs, le fait que nous observions cette régulation dans le muscle embryonnaire mais pas dans le placenta permet également de suggérer deux systèmes de régulation différents. En effet, il a déjà été observé des différences entre les tissus embryonnaire et placentaire en ce qui concerne l'Empreinte (Wagschal and Feil, 2006). Tout d'abord, un certain nombre de gènes ne s'exprime de façon monoallélique que dans le placenta, comme *Ascl2* et les gènes à la périphérie du locus *Cdkn1c*. De plus, le mode de régulation de l'Empreinte apparaît différent dans le placenta : en particulier, on observe surtout des modifications post-traductionnelles des histones alors que la méthylation de l'ADN semble secondaire (Lewis *et al.*, 2004; Umlauf *et al.*, 2004). Ainsi le programme d'expression des gènes soumis à l'Empreinte et leur régulation diffèrent entre le placenta et l'embryon. Nos résultats suggèrent également une différence de ce type en ce qui concerne le réseau de gènes soumis à l'Empreinte. Il serait donc intéressant d'étudier si l'IGN existe dans cette annexe extra embryonnaire et d'explorer son mode de régulation.



**Figure 37 : Résumé des travaux sur les embryons bi-maternels.**

Les rectangles roses indiquent une expression appropriée des gènes normalement exprimés à partir de l'allèle maternel, les rectangles bleus indiquent une expression appropriée des gènes paternels. Les délétions des centres d'Empreinte sont indiquées par des traits rouges.

**A. Embryon normal.** Expression monoallélique des gènes *Igf2* et *H19* et des gènes *Dlk1*, *Rtl1*, *Dio3* et des ARN non codants maternels. Génotype viable.

**B. Embryon bimaternel à partir d'un ovocyte ng wt.** Expression bi-allélique des ARN non codants, répression des gènes codant des protéines aux loci *H19/Igf2* et *Dlk1/Gtl2/Dio3*. Génotype létal à E13.5.

**C. Embryon bimaternel à partir d'un ovocyte ng H19Δ3.** Expression appropriée au locus *H19/Igf2*, mais dérégulation au locus *Dlk1/Gtl2/Dio3*. Génotype létal à E17.5.

**D. Embryon bimaternel à partir d'un ovocyte ng ΔICR H19/Igf2.** Expression appropriée au locus *H19/Igf2*, mais dérégulation au locus *Dlk1/Gtl2/Dio3*. Génotype à létalité néonatale avec faible taux de survie.

**E. Embryon bimaternel à partir d'un ovocyte ng ΔIG-DMR Dlk1/Gtl2/Dio3.** Expression appropriée au locus *Dlk1/Gtl2/Dio3*, mais dérégulation au locus *H19/Igf2*. Génotype à létalité périnatale.

**F. Embryon bimaternel à partir d'un ovocyte ng ΔICR ΔIG-DMR.** Expression appropriée aux loci *H19/Igf2* et *Dlk1/Gtl2/Dio3*. Génotype viable.

(Adapté de Kawahara *et al.* 2007).

### 1-1-2) *H19*, réseau et parthénogenèse

Les résultats de l'équipe de T. Kono sur la production de souris bi-maternelles ont suggéré un rôle important pour le gène *H19* dans l'embryogenèse. En effet, deux régions principales semblent être importantes dans le blocage parthénogénétique chez la souris : les loci *H19/Igf2* et *Dlk1/Gtl2*. Les centres d'Empreinte de ces deux domaines sont des DMR gamétiques méthylées sur l'allèle paternel, respectivement dénommés ICR et IG-DMR. Les souris bi-maternelles  $ng^{\Delta ICR \Delta IG-DMR}/fg^{wt}$  ont un taux de survie équivalent aux souris produites par fécondation *in vitro* (Figure 37). En revanche les souris  $ng^{\Delta IG-DMR}/fg^{wt}$  meurent à la naissance et les souris  $ng^{\Delta ICR}/fg^{wt}$  ont un très faible taux de survie périnatale (Kono *et al.*, 2004; Kawahara *et al.*, 2007a; Kawahara *et al.*, 2007b). Dans l'ensemble, ces résultats suggèrent que la létalité périnatale des animaux  $ng^{\Delta IG-DMR}/fg^{wt}$  est due à un défaut d'Empreinte au locus *H19/Igf2*, qui conduit à la surexpression de l'ARN *H19*, l'expression de ce dernier étant biallélique dans les embryons  $ng^{wt}/fg^{wt}$ .

Les niveaux d'expression de différents gènes ont été comparés entre embryons bi-maternels  $ng^{\Delta IG-DMR}/fg^{wt}$  et embryons normaux. Les taux d'ARNm *Igf2* sont quasiment nuls tandis que l'ARN *H19* est fortement augmenté. Les niveaux d'ARN *Gtl2* sont identiques à ceux des embryons sauvages, en revanche les niveaux d'ARNm *Dlk1* sont plus faibles que chez les sauvages. Ce dernier résultat peut s'expliquer par la technique de production de ces embryons qui ne reflète pas un contexte complètement sauvage. Cependant, il pourrait être également possible que cela soit dû à un effet de l'ARN *H19* sur le taux d'ARNm *Dlk1*. En effet, dans le contexte de notre étude, l'ARN *H19* est un élément de contrôle du niveau d'expression des gènes du réseau de gènes soumis à l'Empreinte et notamment de *Dlk1*.

L'impossibilité de parthénogenèse naturelle chez les mammifères est à la base même de la notion d'Empreinte Parentale (McGrath and Solter, 1984; Surani *et al.*, 1984). La notion de ce réseau de gènes soumis à l'Empreinte apporte aujourd'hui une autre perspective quant au fonctionnement de ce contrôle parthénogénétique.

### 1-1-3) Implications pour la pathologie humaine

Les gènes *H19/Igf2* sont impliqués dans deux pathologies humaines que sont les syndromes de Beckwith-Wiedemann (BWS) et de Silver-Russell (SRS) (Gicquel *et al.*, 2005; Enklaar *et al.*, 2006). Ces syndromes présentent des signes cliniques et une atteinte

épigénétique du locus *H19/Igf2* en miroir avec une surcroissance somatique associée à une hyperméthylation de l'ICR dans le BWS et un retard de croissance intra-utérin associé à une hypométhylation pour le SRS. D'après le modèle de compétition des enhancers et d'insulation par CTCF développé à partir des analyses des modèles murins, les deux épimutations conduiraient à une perte d'Empreinte du locus : une surexpression du gène *Igf2* et une répression du gène *H19* dans le cas de l'hyperméthylation ; une répression du gène *Igf2* et une surexpression du gène *H19* en cas d'hypométhylation.

Cependant dans le cas du SRS, les taux de peptides IGF2 circulant dans le sang des enfants atteints sont comparables aux taux normaux (Netchine *et al.*, 2007). Cette donnée suggère un rôle d'IGF2 au niveau paracrine et non endocrine ou un rôle pour la surexpression du gène *H19*. Par ailleurs, l'épimutation de l'ICR sur le locus chromosomique 11p15 explique 60% des cas de SRS. Pour 10% des patients, une unidisomie maternelle du chromosome 7 est également observée, suggérant une atteinte de l'Empreinte des gènes de ce chromosome (Eggermann *et al.*, 1997; Preece *et al.*, 1997). Deux loci soumis à l'Empreinte existent sur le chromosome 7 humain, le premier comporte le gène maternel *GRB10* et le second comporte le gène paternel *MEST*, qui font tous les deux partie de l'IGN décrit chez la souris. Il serait intéressant d'étudier si un réseau similaire existe également chez l'humain. Pour cela, il est envisageable, dans la mesure où il serait possible d'utiliser un tissu pathologique, d'étudier s'il existe une surexpression de l'ARN *H19* chez les patients SRS avec hypométhylation de l'ICR et d'explorer l'expression des gènes du réseau qui sont dérégulés dans notre étude. Ces résultats permettraient de mieux comprendre l'association de ces trois loci (*H19/IGF2*, *GRB10* et *MEST*) soumis à l'Empreinte dans la physiopathologie du syndrome de Silver-Russell. Cela pourrait peut-être apporter également des pistes sur de possibles atteintes d'autres loci soumis à l'Empreinte afin d'expliquer la pathologie dans les cas non associés au locus 11p15 ou au chromosome 7.

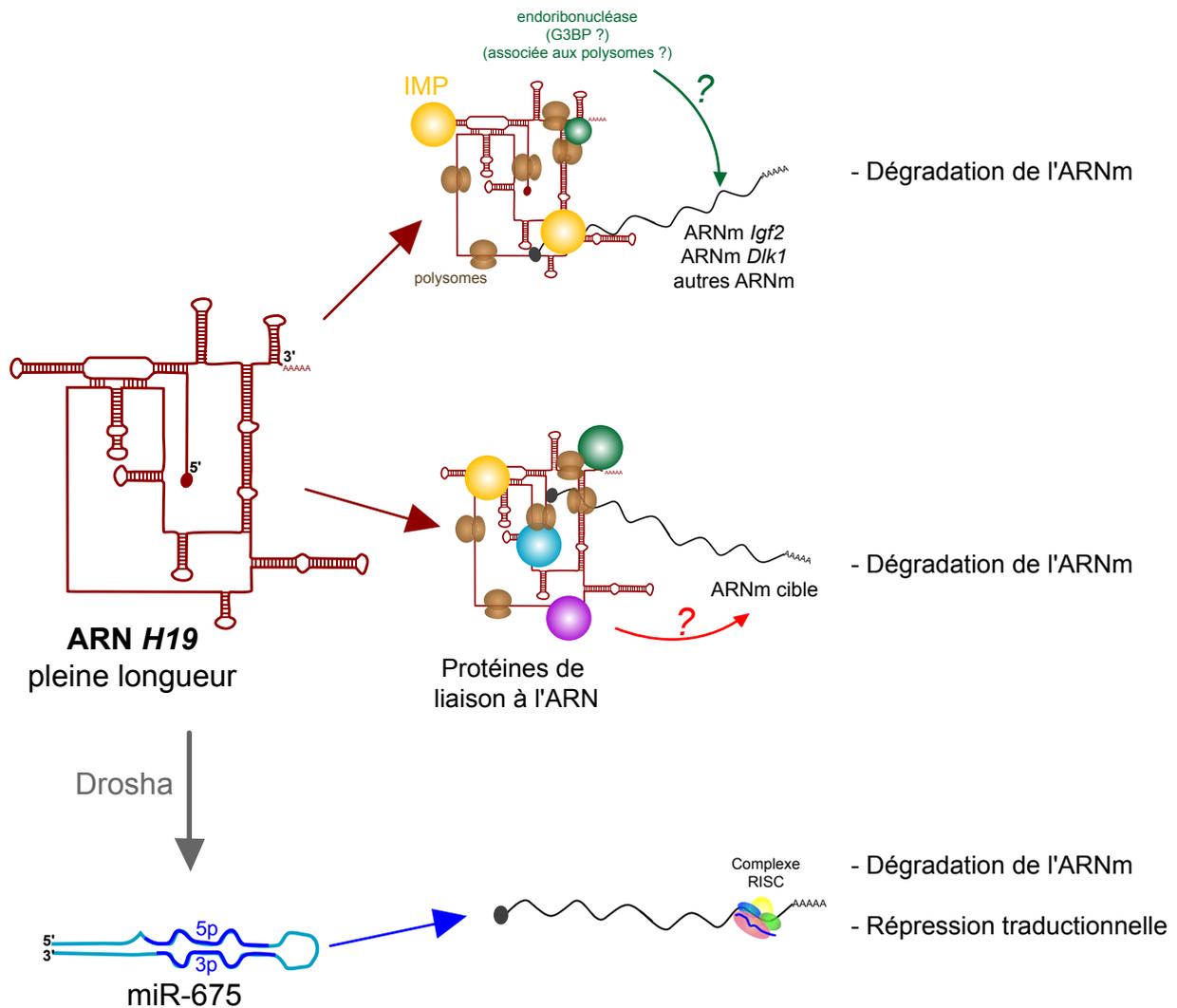
## 1-2) Mécanismes d'action de l'ARN *H19*

La régulation des gènes du réseau par l'ARN *H19* pose la question du mécanisme d'action de cet ARN non codant. L'étude bioinformatique de la séquence de l'ARN *H19* a permis de proposer une structure, conservée entre la souris, le rat et l'humain, complexe, composée de quatorze tige-boucles et de plusieurs duplex (Juan *et al.*, 2000). Par ailleurs un microARN, miR-675, a été décrit comme produit à partir de l'exon 1 de l'ARN *H19* (Mineno *et al.*, 2006; Cai and Cullen, 2007). Une analyse du gène au cours de l'évolution a montré que deux particularités sont conservées : la structure en 5 exons dont les 3 centraux sont de courte taille séparés par de courts introns, suggérant une fonction pour l'ARN pleine longueur épissé, et le microARN dans l'exon 1 qui est conservé depuis au moins 148 millions d'années, suggérant un rôle particulier dans la régulation d'une fonction biologique importante (Smits *et al.*, 2008). Dans notre étude, nous ne pouvons pas distinguer l'une et l'autre possibilité et il faut donc évaluer les mécanismes associés à chacune d'elles (Figure 38).

### 1-2-1) Mécanisme associé à miR-675

La fonction canonique décrite pour les microARN est l'inhibition de la traduction. Cependant, de plus en plus de données indiquent qu'ils pourraient être impliqués dans d'autres phénomènes, notamment la séquestration des ARN dans les P-bodies ou la dégradation des ARNm cibles. Dans notre étude, nous avons observé une réduction du niveau d'ARNm de différents gènes dans les animaux *H19*<sup>-/+</sup>, restaurés à un niveau similaire aux souris sauvages dans les animaux *H19*<sup>-/+</sup>Tg<sup>-/+</sup> et *H19*<sup>-/+</sup>YAC<sup>+/-</sup>. Il est donc possible que miR-675 soit responsable de cette régulation par la voie de dégradation des ARNm. Il serait intéressant d'analyser au niveau protéique les gènes régulés, le taux de protéines pouvant refléter la diminution de l'ARNm ou être encore plus importante, suggérant alors une régulation traductionnelle supplémentaire.

Pour distinguer un effet du microARN d'un effet de l'ARN pleine longueur, des lignées transgéniques surexprimant miR-675 uniquement pourraient être développées et croisées avec les souris *H19*<sup>-/+</sup>. L'analyse de l'expression des gènes du réseau permettrait alors de comprendre la part de la régulation due à miR-675. Cependant ces expériences sont lourdes et l'expression stable du microARN dans tous les tissus murins paraît incertaine. Il peut être envisagé de développer un système de culture de myoblastes *in vitro* à partir de souris *H19*<sup>+</sup> ou *H19*<sup>-</sup> d'une même portée et de comparer l'expression des gènes du réseau.



### Figure 38 : Mécanismes potentiels d'action de l'ARN *H19*.

L'ARN *H19* pourrait agir sous sa forme pleine longueur via des interactions ARN-protéine. Ainsi la liaison avec les IMP pourrait jouer un rôle dans le contrôle de l'expression de l'ARNm *Igf2* pour permettre l'action d'endoribonucléases associées aux polysomes ou liant l'ARN *H19*. Ce mécanisme pourrait aussi être impliqué dans la régulation d'autres ARNm. Par ailleurs, l'interaction avec d'autres protéines pourraient également jouer sur la stabilité et la dégradation d'ARNm cibles de l'ARN *H19*. La structure de l'ARN *H19* présentée sur ce schéma est celle proposée par Juan *et al.* 2000.

Les microARN miR-675 5p et 3p sont produits à partir de l'ARN *H19*. Ainsi, par la voie des miRNA, la stabilité et la traductibilité des ARNm cibles peuvent être altérées.

L'utilisation de myoblastes semble appropriée pour différentes raisons. Premièrement, nous avons montré que l'Empreinte du locus *H19/Igf2* est stable dans ce modèle. Deuxièmement, le gène *H19* endogène s'exprime dans ces cellules. Troisièmement, notre étude du réseau a été réalisée sur le tissu musculaire de l'embryon. Enfin, il a été montré que 7 gènes soumis à l'Empreinte, dont *H19*, *Igf2*, *Cdkn1c*, *Zac1*, *Mest* et *Peg3*, qui font partie de l'IGN, sont activés lors de la régénération musculaire, suggérant un rôle de ces gènes dans la différenciation de ce tissu. La transfection d'un vecteur d'expression du microARN dans les myoblastes pourrait alors permettre d'évaluer la restauration de l'expression normale des gènes et ainsi la régulation due à miR-675.

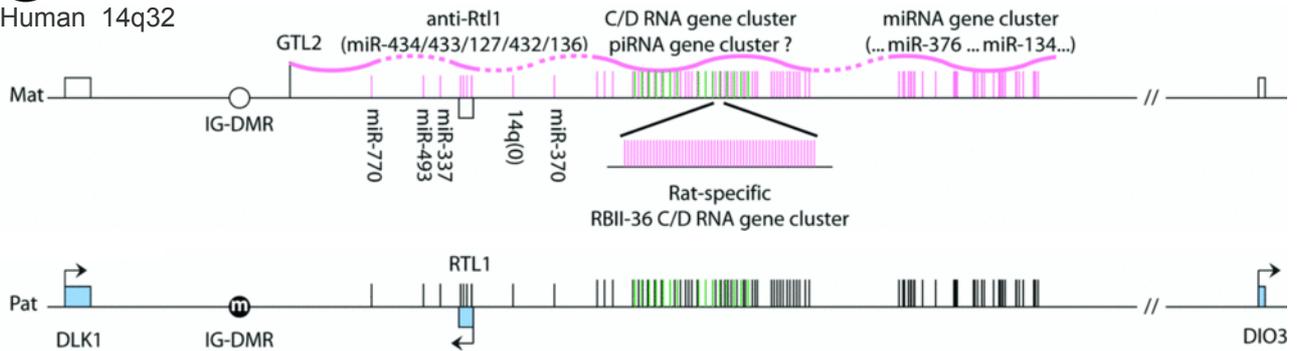
### **1-2-2) Mécanisme pour l'ARN pleine longueur**

L'ARN *H19* pourrait réguler l'expression des autres gènes du réseau via l'interaction avec des protéines de liaison à l'ARN (RBP pour RNA-binding protein). Les seules protéines connues pour se fixer à cet ARN sont les IMP (*Igf2* mRNA-binding proteins) (Runge *et al.*, 2000). Comme leur nom l'indique, ces protéines peuvent également se fixer sur l'ARNm *Igf2* (Nielsen *et al.*, 1999). Cette famille de protéines compte trois membres : IMP1, également appelé CRD-BP (coding-region determinant-binding protein), IMP2 et IMP3. Elles comportent comme motifs importants deux domaines RRM (RNA recognition motif) et quatre domaines KH (heterogenous nuclear ribonucleoprotein K homolog) qui sont des domaines de reconnaissance des ARN. Ces protéines ont été impliquées dans le trafic cellulaire des ARN et, dans le cytoplasme, IMP1 est associée aux ribonucléoparticules. La protéine IMP1 possède plusieurs sites de fixation dans le 5'UTR-leader3 de l'ARNm *Igf2* (issu de la transcription à partir du promoteur P3) et dans le 3'UTR. Elle peut aussi se lier à l'ARNm *c-Myc* et à celui de la  $\beta$ -actine. Dans ce cas, IMP1 protégerait l'ARNm *c-Myc* d'une attaque par une endoribonucléase associée aux polysomes et permettrait ainsi d'augmenter sa stabilité (Sparanese and Lee, 2007). Un tel clivage existe pour l'ARNm *Igf2* (Nielsen and Christiansen, 1992; van Dijk *et al.*, 2000).

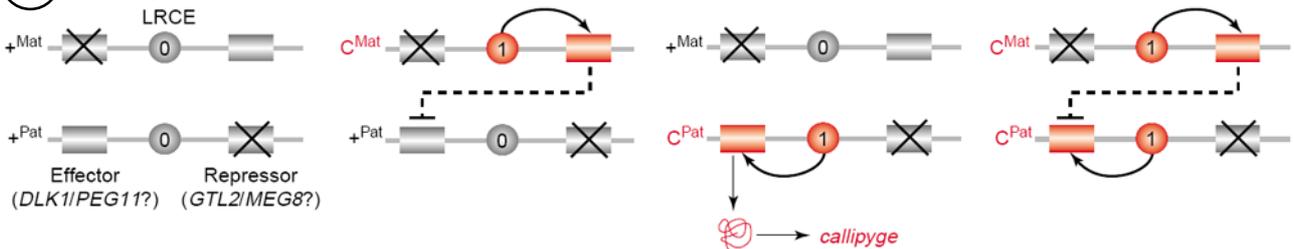
Il a été montré qu'une autre endoribonucléase, la G3BP (RasGAP-Associated endoribonuclease) est capable de cliver l'ARN *c-Myc* et l'ARN *Igf2* (Tourriere *et al.*, 2001; Zekri *et al.*, 2005). Par ailleurs, dans les fibroblastes embryonnaires de souris invalidées pour le gène *G3BP*, cinq ARN de gènes soumis à l'Empreinte sont significativement dérégulés : *Igf2*, *Dlk1*, *H19*, *Cdkn1c* et *Grb10* (Zekri *et al.*, 2005). Par ailleurs, un clivage a été évoqué pour l'ARNm *Dlk1* (Paulsen *et al.*, 2001) et l'article de Zekri *et al.* mentionne la présence de

**A**

Human 14q32

**B**

Mouton



**Figure 39 : Locus *Dlk1/Gtl2/Dio3*.**

**A.** Transcrits au locus 14q32 humain. A partir de l'allèle maternel sont transcrits les ARN non codants *GTL2*, *anti-RTL1*, des snoRNA et des piRNA (transcrit *Rian* murin) et des miRNA (transcrit *Mirg* murin). A partir de l'allèle paternel sont exprimés trois gènes codant des protéines : *DLK1*, *RTL1* et *DIO3*. Le centre d'Empreinte IG-DMR est méthylé uniquement sur l'allèle paternel. Les ARN maternels pourraient contrôler l'expression des gènes paternels. (Adapté de Royo and Cavaille, 2008)

**B.** Modèle de "Polar overdominance" chez le mouton : transmission de la mutation callipyge (allèle C). Lorsque la mutation est transmise par la mère, les ARN non codants maternels sont transcrits mais le phénotype n'est pas observé. En cas de transmission paternelle, elle entraîne l'expression persistante du gène *Dlk1* qui conduit à l'hypertrophie musculaire. En cas de transmission par les deux parents, les ARN non codants maternels opposent un contrôle post-transcriptionnel ou traductionnel qui empêche l'expression persistante de *Dlk1*, le phénotype d'hypertrophie n'est pas observé. (Adapté de Georges *et al.*, 2003)

trois sites pour G3BP. Cependant aucune évidence expérimentale n'est, à ma connaissance, disponible quant à la réalité et à la fonction de ce clivage.

Un modèle d'action de l'ARN via la liaison avec les IMP pourrait être le suivant. L'ARN *H19* est lié aux polysomes et lie plusieurs molécules d'IMP. De cette manière l'ARNm *Igf2* est en contact via IMP1 avec des polysomes où un clivage par une endoribonucléase associée aux polysomes pourrait avoir lieu. Ce modèle est bien sûr théorique et il serait intéressant d'étudier si l'association de l'ARN *H19* aux polysomes se fait via les IMP en utilisant par exemple le modèle de souris *IMP1<sup>-/-</sup>*. Sinon, il pourrait être pertinent d'obtenir des données biochimiques sur les affinités respectives de l'ARN *H19* pour les polysomes et pour les IMP. De plus il serait utile de savoir si la fixation des ARN *H19* et *Igf2* est mutuellement exclusive, ce qui serait plutôt en faveur d'un mécanisme de compétition. Par ailleurs, si le clivage de l'ARNm *Dlk1* est confirmé et est du même type que celui de l'ARNm *Igf2*, cela pourrait impliquer un rôle de l'ARN *H19* dans ces clivages. Cependant, au locus *Dlk1/Gtl2*, de nombreux petits ARN non codants sont impliqués dans la régulation de l'expression des gènes et il semble plus plausible, selon le modèle de "polar overdominance", qu'ils soient responsables d'un tel clivage (Figure 39). Ainsi pour ce locus, les microARN maternels produits à partir de l'ARN *AntiRtl1* entraînent la dégradation par clivage de l'ARNm *Rtl1* et il est possible qu'un mécanisme similaire participe à la régulation de l'ARNm *Dlk1* (Davis *et al.*, 2004; Davis *et al.*, 2005).

D'autres protéines pourraient se lier à l'ARN *H19* et pourraient intervenir dans la régulation de l'expression des ARN des gènes du réseau. Pour les identifier, la technique de triple-hybride pourrait être utilisée (SenGupta *et al.*, 1996). Dans cette technique adaptée du double hybride, trois vecteurs sont nécessaires. Le premier exprime la protéine MS2 fusionnée à un domaine de liaison à l'ADN, le second exprime l'ARN fusionné à une séquence de reconnaissance par MS2 et le dernier exprime la protéine candidate pour la liaison à l'ARN fusionné à un domaine d'activation transcriptionnelle. Ainsi, en cas de liaison entre l'ARN d'intérêt et la protéine candidate, la transcription d'un gène rapporteur est activée. En utilisant une banque d'expression de protéines de fusion, un criblage triple-hybride pourrait être réalisé pour identifier les protéines interagissant avec l'ARN *H19*. La difficulté de réalisation de cette approche réside dans le choix de la séquence d'ARN utilisée comme appât, puisque la taille de l'ARN *H19* est de 2,7 kb.

Une autre technique envisageable est la technique de Northwestern blot avec une sonde ARN *H19*. Cette première approche permettrait de savoir si l'ARN est capable d'interagir avec une ou plusieurs protéines. Le criblage d'une banque d'expression par la sonde ARN *H19* pourrait alors permettre de définir les protéines de liaison. Ces données pourraient être validées par la fixation des IMP sur *H19* et pourraient permettre de mieux comprendre le mécanisme par lequel l'ARN *H19* participe au réseau de gènes soumis à l'Empreinte.

### **1-2-3) ARN *H19* et méthylation de l'ADN ?**

Les mécanismes d'actions proposés ci-dessus sont post-transcriptionnels. Il est également possible que l'ARN *H19* agisse au niveau transcriptionnel. En effet, de plus en plus d'études suggèrent un contrôle transcriptionnel par les petits ARN. Il a récemment été montré le rôle d'un cluster de microARN dans la méthylation par les Dnmt *de novo* dans les cellules ES murines (Sinkkonen *et al.*, 2008). Par ailleurs Ago-1 pourrait être impliqué dans le recrutement des protéines Dnmt3a, HDAC1 et KMT6 (Ezh2) aux promoteurs via l'interaction de petits ARN avec l'ARN naissant ou l'ADN simple brin au moment de l'élongation. La méthylation de l'ADN et des histones et la désacétylation des histones entraîneraient alors une compaction de la chromatine et une répression transcriptionnelle (Hawkins and Morris, 2008).

### **1-2-4) Mécanisme d'action indirecte**

Nos résultats montrent que l'altération de l'expression de l'ARN *H19* entraîne la dérégulation de onze autres gènes du réseau. Le fonctionnement d'un tel réseau implique que ces gènes ou leurs produits interagissent les uns avec les autres. Aussi, la modification d'expression d'un gène pourrait entraîner la dérégulation d'autres gènes et ainsi conduire à un effet indirect de l'altération de l'ARN *H19*. Il est possible que la surexpression d'*Igf2* puisse entraîner la surexpression de régulateurs négatifs de la voie IGF comme *Igf2R* ou *Grb10*, bien qu'aucune étude n'ait exploré cette hypothèse. En revanche, un lien existe entre *Igf2* et *Cdkn1c*. En effet il a été rapporté qu'un taux élevé de peptide *Igf2* circulant chez la souris *Igf2R*<sup>+/+</sup> conduit à la réduction de l'expression du gène *Cdkn1c* (Grandjean *et al.*, 2000). Il serait intéressant d'explorer la dérégulation des différents gènes régulés par l'ARN *H19*, en utilisant les différents modèles d'inactivation et de surexpression disponibles afin d'établir les

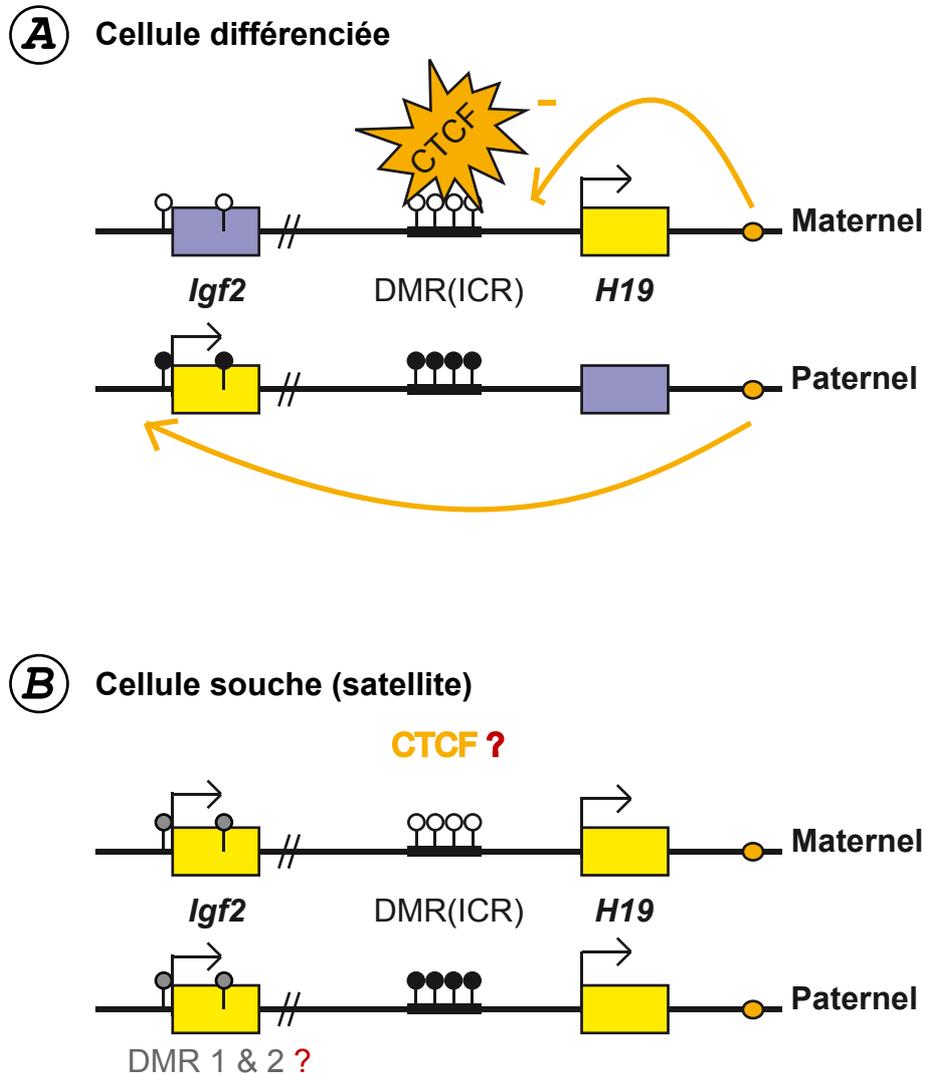
liens entre ces gènes. Ainsi, dans la souris invalidée pour *Zac1*, l'expression des gènes *H19* et *Igf2* sont modulés par l'absence de la fixation du facteur de transcription *Zac1* au niveau de leurs enhancers communs (Varrault *et al.*, 2006). Dans notre modèle, l'altération de l'expression du gène *H19* n'affecte pas l'expression du gène *Zac1*, ce qui permettrait de placer *H19* en aval de *Zac1* dans le réseau.

## 2) Empreinte et cellules souches adultes.

Le gène *H19* est très fortement exprimé dans les cellules embryonnaires. Après la naissance, l'expression décroît rapidement dans tous les tissus et le gène ne reste fortement exprimé à l'âge adulte que dans les tissus musculaires squelettiques (Poirier *et al.*, 1991). Par ailleurs, le gène *H19* avait été cloné dans la même banque que *MyoD* et appelé *MyoH*, suggérant un rôle important pour la différenciation musculaire (Davis *et al.*, 1987). Nous nous sommes donc intéressées à cet organe, avec l'idée que l'ARN *H19* pourrait y jouer un rôle particulier. La régénération musculaire est un mécanisme par lequel le muscle est réparé en cas de lésion. Ce processus fait intervenir les cellules satellites qui vont être activées, proliférer et se différencier selon une séquence d'expression génique similaire à celle observée au cours de l'embryogenèse. Ainsi, la cellule satellite semblait un bon modèle d'étude.

La cellule satellite est considérée comme la cellule souche du muscle adulte (Le Grand and Rudnicki, 2007; Sambasivan and Tajbakhsh, 2007). Il existe peu de données quant à la régulation de l'Empreinte dans les cellules souches adultes. Une étude sur des tissus humains montre que le gène *Igf2* est exprimé de façon monoallélique dans les cellules du sang circulant. En revanche, dans la moelle osseuse, riche en cellules souches hématopoïétiques, l'expression est biallélique. Ces résultats suggèrent que le gène *Igf2* s'exprime de façon biallélique dans les cellules souches et de façon monoallélique dans les cellules différenciées (Morison *et al.*, 2000). Un autre modèle de cellules souches étudié est celui des cellules germinales primordiales qui sont les cellules souches gamétiques. Dans ces cellules, la méthylation d'Empreinte est effacée et les gènes *H19*, *Igf2*, *Igf2R* et *Snrpn* sont exprimés de façon biallélique (Szabo and Mann, 1995a). Nous avons donc voulu explorer l'hypothèse que cette caractéristique pouvait également exister pour d'autres cellules souches chez l'adulte et choisi la cellule satellite comme modèle d'étude.

Dans un premier temps, nous avons prélevé le muscle de souris *Igf2*<sup>+/-</sup> et mis en culture des fibres musculaires. A partir de ces fibres, les cellules satellites présentes sous la membrane basale se détachent et commencent à se diviser en culture puis à se différencier. Nous avons observé que le gène *Igf2* est exprimé dans ces cellules. Cette expression provient de l'allèle maternel normalement réprimé, suggérant ainsi une expression biallélique du gène. En miroir, nous avons prélevé les cellules satellites de souris *H19*<sup>+/-</sup> et observé l'expression du gène *H19* à partir de l'allèle paternel, suggérant de nouveau l'expression biallélique du gène.



**Figure 40 : Résumé des résultats et questions sur les cellules souches adultes.**

Les rectangles jaunes indiquent un allèle exprimé, les rectangles bleus indiquent un allèle silencieux. Les ronds noirs indiquent la méthylation de l'ADN, un rond blanc indique l'ADN non méthylié. Les ronds jaunes symbolisent les enhancers communs des deux gènes.

**A.** Dans les cellules différenciées, tissus musculaires adultes et petites cellules non satellites du diaphragme, l'expression des gènes *H19* et *Igf2* est monoallélique.

**B.** Dans les cellules satellites musculaires, l'expression des gènes *H19* et *Igf2* est biallélique. L'ICR est convenablement différenciellement méthylié. L'expression biallélique du gène *Igf2* pourrait être due à une méthylation des DMR somatiques 1 et 2 ou à l'absence de fixation du facteur CTCF, entre autres hypothèses.

Nous avons cependant considéré que l'utilisation de souris mutantes hétérozygotes pouvait introduire un biais d'expression, en raison de la présence d'un allèle nul. Afin de s'affranchir d'un effet de l'allèle délété sur l'allèle homologue, nous avons utilisé la lignée SD7, qui porte un fragment *spretus* sur la partie distale du chromosome 7, qui comprend le locus *H19/Igf2*. Par ailleurs, nous avons voulu étudier les cellules satellites dans l'état de quiescence le plus proche possible de celui qu'elles ont dans la fibre musculaire. Pour cela, nous avons utilisé un modèle de souris dans lequel le gène de la GFP remplace le gène *Pax3*, qui est un marqueur de cellules satellites. Les tissus musculaires, en particulier le diaphragme, peuvent être triés au FACS et les cellules satellites isolées grâce à l'expression de la GFP (Montarras *et al.*, 2005). Des souris *Pax3<sup>GFP/+</sup>* ont été croisées avec les mâles SD7 et les cellules satellites du diaphragme de la descendance ont été isolées.

Nous avons analysé l'expression du gène *Igf2*, en utilisant un polymorphisme de restriction et montré l'expression biallélique dans les cellules de moelle de souris traitées avec la 5FU, qui permet d'accroître la proportion de cellules souches de la moelle, confirmant ainsi les résultats publiés dans la littérature (Morison *et al.*, 2000). Par ailleurs, nous avons pu montrer pour la première fois que le gène *Igf2* est également d'expression biallélique dans les cellules satellites alors qu'il est monoallélique dans le tissu musculaire différencié. Cette absence de lecture de l'Empreinte pourrait être due à des modifications de méthylation au locus. En effet, la méthylation différentielle est apposée sur les DMR gamétiques chez la souris au cours de la spermatogenèse et de la maturation ovocytaire. L'Empreinte est ensuite maintenue au cours des divisions cellulaires qui ont lieu pendant l'embryogenèse et peut être lue dans les cellules somatiques (Figure 22). Il existe une différence entre la présence des marques de méthylation différentielle et l'expression monoallélique. Ainsi, dans la masse cellulaire interne de l'embryon de souris, la méthylation est effectivement différentielle mais le gène *Igf2* s'exprime de façon biallélique (Szabo and Mann, 1995b). Nous avons donc analysé l'ICR du locus *H19/Igf2* et avons observé qu'il est correctement différentiellement méthylé dans les cellules satellites (Figure 40). Cette donnée suggère que les marques d'Empreinte sont présentes dans les cellules souches mais que l'Empreinte est "ignorée". L'analyse des DMR somatiques au niveau du gène *Igf2* est en cours et ces expériences vont ouvrir de nouvelles perspectives quant à un découplage méthylation/lecture de l'Empreinte dans les cellules satellites. De plus, il serait intéressant d'analyser l'expression de la protéine CTCF dans les cellules satellites afin de voir si elle peut y jouer son rôle d'insulateur. Il serait également important de pouvoir étudier les modifications post-traductionnelles des histones par Immunoprécipitation de la chromatine (ChIP). Cependant, les cellules satellites sont peu nombreuses dans les tissus musculaires et trier ces cellules pour étudier différentes marques

d'histones nécessiterait l'utilisation d'un très grand nombre d'animaux. Des expériences, menées dans d'autres laboratoires, sont en cours pour optimiser les analyses de CHIP sur de petits échantillons.

## 2-1) Implication de ces résultats pour l'étude des cellules satellites

Les questions posées en ce qui concerne les cellules satellites sont de connaître l'origine de ces cellules et d'étudier si elles constituent une population unique ou si plusieurs populations aux caractéristiques différentes existent (Le Grand and Rudnicki, 2007; Sambasivan and Tajbakhsh, 2007). Cette deuxième question est le fruit de la constatation que toutes les cellules satellites n'expriment pas les mêmes marqueurs, notamment certaines cellules expriment *Myf5* et d'autres non (Kuang *et al.*, 2007).

Le gène *Igf2* s'exprime de façon biallélique dans la masse cellulaire interne du blastocyste et est ensuite monoallélique dans les tissus endodermiques et mésodermiques (Szabo and Mann, 1995b). Ceci suggère que l'expression monoallélique apparaît avec la différenciation des tissus au cours de l'embryogenèse. Les données d'embryologie expérimentale montrent qu'au cours de la myogenèse, un pool de cellules souches,  $Pax3^+Pax7^+$ , migre depuis le dermomyotome pour assurer un apport continu de cellules progénitrices au myotome (Relaix *et al.*, 2005). L'expression biallélique du gène *Igf2* dans les cellules satellites serait également en faveur d'un stock de cellules qui gardent des caractéristiques "souches" malgré leur intégration dans les tissus différenciés. Par ailleurs, nos résultats indiquent que l'allèle normalement réprimé s'exprime à un taux de 15 à 20% par rapport à l'allèle normalement exprimé. Cela peut indiquer soit que 15 à 20% des cellules expriment les gènes de façon biallélique sans biais allélique (à 100%), soit que dans chaque cellule satellite, un relâchement de l'empreinte, permettant une expression de 15 à 20%, est observé pour ce locus. La confirmation de la première hypothèse serait en faveur de l'existence de populations distinctes de cellules satellites, certaines étant déjà plus déterminées que les autres vers le destin myogénique. Une étude de l'expression au locus *H19/Igf2* par ARN FISH sur les cellules triées au FACS pourrait permettre de résoudre cette question. Grâce à cette technique, une figure en un spot est observée en cas d'expression monoallélique et en deux spots en cas d'expression biallélique. Cette technique est très utilisée pour l'étude de l'ARN *Xist*, responsable de l'inactivation du chromosome X chez la femelle (Chaumeil *et al.*, 2004). Cette analyse est en cours en collaboration avec l'équipe d'E. Heard. Nous explorerons également l'expression de l'ARN *Xist* afin de voir si le même type de régulation existe pour cet ARN dans ces cellules souches, chez la femelle.

## 2-2) Cellules souches et thérapie cellulaire

Les cellules souches adultes sont un sujet de recherche très dynamique grâce à l'espoir que suscite l'utilisation de ces cellules en thérapie cellulaire. Dans notre étude, nous avons montré que l'expression biallélique du gène *Igf2* n'est pas modifiée en expression monoallélique au cours de la différenciation musculaire *in vitro*. Si l'expression biallélique n'est trouvée que dans une sous-population de cellules satellites, alors cette observation peut indiquer qu'en culture, l'auto-renouvellement de cette population est assurée. Ainsi, l'expression biallélique refléterait la présence et le maintien de ces cellules à caractéristiques souches dans la culture de cellules musculaires en différenciation.

Cependant, la littérature fait de plus en plus référence à l'altération de l'Empreinte due aux manipulations des cellules. Ainsi, chez les bovins et les ovins, il a été rapporté que la culture des embryons peut entraîner un syndrome de surcroissance (LOS pour Large Offspring Syndrome), résultant de la perte d'Empreinte du locus *Igf2/H19* ou du locus *Igf2R* (Feil, 2001). Par ailleurs, les embryons issus de fécondation *in vitro* chez la souris ont une altération de la méthylation au locus *H19/Igf2*, en comparaison à des embryons issus de fécondation *in vivo* (Fauque *et al.*, 2007). Dans ce contexte, nos résultats suggèrent que malgré la présence correcte des marques de méthylation d'Empreinte, l'expression monoallélique ne peut pas être établie *in vitro*. Cette donnée doit être prise en compte dans l'optique de l'utilisation de ces cellules pour la thérapie cellulaire. En effet, si la manipulation des cellules satellites entraîne des défauts d'établissement de l'Empreinte, le contrôle de l'expression des gènes soumis à l'Empreinte pourrait être altéré. Ces gènes ont pour une grande partie été impliqués dans la tumorigenèse, notamment les gènes *H19/Igf2*, *Zac1*, *Cdkn1c*, etc... D'autres gènes et d'autres types cellulaires devront être étudiés afin de mieux comprendre la régulation de l'Empreinte dans les cellules souches chez l'adulte. Par ailleurs il est essentiel d'analyser la différenciation de ces cellules *in vivo* et de vérifier que la régulation de l'Empreinte est correctement établie dans un contexte physiologique de réparation de tissu lésé. Pour cela, nous envisageons de réinjecter, dans le muscle lésé de souris *Igf2<sup>-/-</sup>*, les cellules satellites triées au FACS à partir de descendants du croisement *Pax3<sup>GFP</sup>* par SD7. Une RT-PCR allèle spécifique pourrait ainsi nous permettre de contrôler que les muscles régénérés à l'aide de ces cellules expriment le gène *Igf2* de façon monoallélique paternelle. Ainsi, cela montrerait clairement que l'Empreinte Parentale n'est pas lue dans ces cellules souches adultes mais que la marque primaire d'empreinte, toujours présente, conduit à l'établissement d'une expression monoallélique dans un tissu différencié.

### 3) Conclusions générales et perspectives

Les questions que nous avons posées pour ce travail de thèse étaient de comprendre le rôle et la régulation de l'expression du gène *H19*. Cet ARN non codant soumis à l'Empreinte a été découvert il y a vingt ans et sa fonction restait encore inconnue. L'originalité de nos travaux a été l'utilisation de modèles murins afin d'explorer les différents thèmes de recherche autour de ce gène. Dans un premier temps, trois modèles de tumorigenèse *in vivo* ont permis de confirmer le rôle suppresseur de tumeur du gène *H19*. Puis l'analyse de souris transgéniques a permis de révéler un rôle de *trans* régulateur d'un réseau de gènes soumis à l'Empreinte, impliqué dans la croissance embryonnaire. L'ARN *H19* contrôle ainsi finement le taux d'expression de plusieurs autres gènes soumis à l'Empreinte. Enfin, l'expression des gènes du locus *H19/Igf2* est biallélique dans les cellules satellites et les cellules souches hématopoïétiques, ce qui implique une régulation de l'Empreinte particulière pour les cellules souches adultes.

Les questions qui se posent à présent sont de comprendre par quels mécanismes l'ARN *H19* peut agir et de pouvoir différencier le rôle de l'ARN pleine longueur de celui du microARN miR-675. De plus, comprendre l'implication du gène *H19* dans le processus de différenciation/régénération musculaire pourrait permettre également de nouvelles avancées quant à la fonction de cet ARN non codant.

# *Bibliographie*

**A**

Aapola, U., Kawasaki, K., Scott, H. S., Ollila, J., Vihinen, M., Heino, M., Shintani, A., Minoshima, S., Krohn, K., Antonarakis, S. E., Shimizu, N., Kudoh, J. and Peterson, P. (2000). Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family. *Genomics* 65, 293-8.

Abu-Amero, S., Monk, D., Frost, J., Preece, M., Stanier, P. and Moore, G. E. (2008). The genetic aetiology of Silver-Russell syndrome. *J Med Genet* 45, 193-9.

Ainscough, J. F., Koide, T., Tada, M., Barton, S. and Surani, M. A. (1997). Imprinting of *Igf2* and *H19* from a 130 kb YAC transgene. *Development* 124, 3621-32.

Ainscough, J. F., Dandolo, L. and Surani, M. A. (2000a). Appropriate expression of the mouse *H19* gene utilises three or more distinct enhancer regions spread over more than 130 kb. *Mech Dev* 91, 365-8.

Ainscough, J. F., John, R. M., Barton, S. C. and Surani, M. A. (2000b). A skeletal muscle-specific mouse *Igf2* repressor lies 40 kb downstream of the gene. *Development* 127, 3923-30.

Allis, C. D., Berger, S. L., Cote, J., Dent, S., Jenuwien, T., Kouzarides, T., Pillus, L., Reinberg, D., Shi, Y., Shiekhhattar, R., Shilatifard, A., Workman, J. and Zhang, Y. (2007). New nomenclature for chromatin-modifying enzymes. *Cell* 131, 633-6.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* 215, 403-10.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M. J., Kuramochi-Miyagawa, S., Nakano, T., Chien, M., Russo, J. J., Ju, J., Sheridan, R., Sander, C., Zavolan, M. and Tuschl, T. (2006). A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* 442, 203-7.

Arney, K. L. (2003). *H19* and *Igf2*--enhancing the confusion? *Trends Genet* 19, 17-23.

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H. F., John, R. M., Gouti, M., Casanova, M., Warnes, G., Merckenschlager, M. and Fisher, A. G. (2006). Chromatin signatures of pluripotent cell lines. *Nat Cell Biol* 8, 532-8.

**B**

Bajard, L., Relaix, F., Lagha, M., Rocancourt, D., Daubas, P. and Buckingham, M. E. (2006). A novel genetic hierarchy functions during hypaxial myogenesis: Pax3 directly activates Myf5 in muscle progenitor cells in the limb. *Genes Dev* 20, 2450-64.

Banerjee, S. and Smallwood, A. (1995). A chromatin model of *IGF2/H19* imprinting. *Nat Genet* 11, 237-8.

Barlow, D. P., Stoger, R., Herrmann, B. G., Saito, K. and Schweifer, N. (1991). The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. *Nature* 349, 84-7.

- Barsyte-Lovejoy, D., Lau, S. K., Boutros, P. C., Khosravi, F., Jurisica, I., Andrulis, I. L., Tsao, M. S. and Penn, L. Z. (2006). The c-Myc oncogene directly induces the *H19* noncoding RNA by allele-specific binding to potentiate tumorigenesis. *Cancer Res* 66, 5330-7.
- Bartolomei, M. S., Zemel, S. and Tilghman, S. M. (1991). Parental imprinting of the mouse *H19* gene. *Nature* 351, 153-5.
- Bartolomei, M. S., Webber, A. L., Brunkow, M. E. and Tilghman, S. M. (1993). Epigenetic mechanisms underlying the imprinting of the mouse *H19* gene. *Genes Dev* 7, 1663-73.
- Barton, S. C., Surani, M. A. and Norris, M. L. (1984). Role of paternal and maternal genomes in mouse development. *Nature* 311, 374-6.
- Beauchamp, J. R., Heslop, L., Yu, D. S., Tajbakhsh, S., Kelly, R. G., Wernig, A., Buckingham, M. E., Partridge, T. A. and Zammit, P. S. (2000). Expression of CD34 and Myf5 defines the majority of quiescent adult skeletal muscle satellite cells. *J Cell Biol* 151, 1221-34.
- Bell, A. C., West, A. G. and Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98, 387-96.
- Bell, A. C. and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* 405, 482-5.
- Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S. L. and Lander, E. S. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315-26.
- Bernstein, E., Caudy, A. A., Hammond, S. M. and Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 409, 363-6.
- Berteaux, N., Lottin, S., Monte, D., Pinte, S., Quatannens, B., Coll, J., Hondermarck, H., Curgy, J. J., Dugimont, T. and Adriaenssens, E. (2005). *H19* mRNA-like noncoding RNA promotes breast cancer cell proliferation through positive control by E2F1. *J Biol Chem* 280, 29625-36.
- Bestor, T., Laudano, A., Mattaliano, R. and Ingram, V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *J Mol Biol* 203, 971-83.
- Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes Dev* 16, 6-21.
- 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, 1760-4.
- 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, 2536-9.

Brannan, C. I., Dees, E. C., Ingram, R. S. and Tilghman, S. M. (1990). The product of the *H19* gene may function as an RNA. *Mol Cell Biol* 10, 28-36.

Breiling, A., Sessa, L. and Orlando, V. (2007). Biology of polycomb and trithorax group proteins. *Int Rev Cytol* 258, 83-136.

Buckingham, M. (2006). Myogenic progenitor cells and skeletal myogenesis in vertebrates. *Curr Opin Genet Dev* 16, 525-32.

## C

Cai, X., Hagedorn, C. H. and Cullen, B. R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* 10, 1957-66.

Cai, X. and Cullen, B. R. (2007). The imprinted *H19* noncoding RNA is a primary microRNA precursor. *RNA* 13, 313-6.

Callis, T. E., Chen, J. F. and Wang, D. Z. (2007). MicroRNAs in skeletal and cardiac muscle development. *DNA Cell Biol* 26, 219-25.

Casola, S., Ungaro, P., Pedone, P. V., Lazzaro, D., Fattori, E., Ciliberto, G., Zarrilli, R., Bruni, C. B. and Riccio, A. (1995). Loss of heterozygosity of imprinted genes in SV40 t/T antigen-induced hepatocellular carcinomas. *Oncogene* 11, 711-21.

Caspary, T., Cleary, M. A., Perlman, E. J., Zhang, P., Elledge, S. J. and Tilghman, S. M. (1999). Oppositely imprinted genes *p57(Kip2)* and *igf2* interact in a mouse model for Beckwith-Wiedemann syndrome. *Genes Dev* 13, 3115-24.

Cattanach, B. M. and Kirk, M. (1985). Differential activity of maternally and paternally derived chromosome regions in mice. *Nature* 315, 496-8.

Cattanach, B. M., Beechey, C. V. and Peters, J. (2006). Interactions between imprinting effects: summary and review. *Cytogenet Genome Res* 113, 17-23.

Cavaille, J., Buiting, K., Kiefmann, M., Lalande, M., Brannan, C. I., Horsthemke, B., Bachellerie, J. P., Brosius, J. and Huttenhofer, A. (2000). Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. *Proc Natl Acad Sci U S A* 97, 14311-6.

Cavaille, J., Seitz, H., Paulsen, M., Ferguson-Smith, A. C. and Bachellerie, J. P. (2002). Identification of tandemly-repeated C/D snoRNA genes at the imprinted human 14q32 domain reminiscent of those at the Prader-Willi/Angelman syndrome region. *Hum Mol Genet* 11, 1527-38.

Charalambous, M., Smith, F. M., Bennett, W. R., Crew, T. E., Mackenzie, F. and Ward, A. (2003). Disruption of the imprinted *Grb10* gene leads to disproportionate overgrowth by an *Igf2*-independent mechanism. *Proc Natl Acad Sci U S A* 100, 8292-7.

Charalambous, M., Menheniott, T. R., Bennett, W. R., Kelly, S. M., Dell, G., Dandolo, L. and Ward, A. (2004). An enhancer element at the *Igf2/H19* locus drives gene expression in both imprinted and non-imprinted tissues. *Dev Biol* 271, 488-97.

- 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. *Curr Opin Endocrinol Diabetes Obes* 14, 3-12.
- Charge, S. B. and Rudnicki, M. A. (2004). Cellular and molecular regulation of muscle regeneration. *Physiol Rev* 84, 209-38.
- Charlier, C., Segers, K., Karim, L., Shay, T., Gyapay, G., Cockett, N. and Georges, M. (2001). The callipyge mutation enhances the expression of coregulated imprinted genes in cis without affecting their imprinting status. *Nat Genet* 27, 367-9.
- Charvet, C., Houbron, C., Parlakian, A., Giordani, J., Lahoute, C., Bertrand, A., Sotiropoulos, A., Renou, L., Schmitt, A., Melki, J., Li, Z., Daegelen, D. and Tuil, D. (2006). New role for serum response factor in postnatal skeletal muscle growth and regeneration via the interleukin 4 and insulin-like growth factor 1 pathways. *Mol Cell Biol* 26, 6664-74.
- Chaumeil, J., Okamoto, I. and Heard, E. (2004). X-chromosome inactivation in mouse embryonic stem cells: analysis of histone modifications and transcriptional activity using immunofluorescence and FISH. *Methods Enzymol* 376, 405-19.
- Chaumeil, J., Le Baccon, P., Wutz, A. and Heard, E. (2006). A novel role for *Xist* RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes Dev* 20, 2223-37.
- Chen, J. F., Mandel, E. M., Thomson, J. M., Wu, Q., Callis, T. E., Hammond, S. M., Conlon, F. L. and Wang, D. Z. (2006). The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nat Genet* 38, 228-33.
- Choo, J. H., Kim, J. D., Chung, J. H., Stubbs, L. and Kim, J. (2006). Allele-specific deposition of macroH2A1 in imprinting control regions. *Hum Mol Genet* 15, 717-24.
- Chu, C. Y. and Rana, T. M. (2007). Small RNAs: regulators and guardians of the genome. *J Cell Physiol* 213, 412-9.
- Cirio, M. C., Ratnam, S., Ding, F., Reinhart, B., Navara, C. and Chaillet, J. R. (2008). Preimplantation expression of the somatic form of *Dnmt1* suggests a role in the inheritance of genomic imprints. *BMC Dev Biol* 8, 9.
- Clop, A., Marcq, F., Takeda, H., Pirottin, D., Tordoir, X., Bibe, B., Bouix, J., Caiment, F., Elsen, J. M., Eychenne, F., Larzul, C., Laville, E., Meish, F., Milenkovic, D., Tobin, J., Charlier, C. and Georges, M. (2006). A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nat Genet* 38, 813-8.
- Colnot, S., Niwa-Kawakita, M., Hamard, G., Godard, C., Le Plenier, S., Houbron, C., Romagnolo, B., Berrebi, D., Giovannini, M. and Perret, C. (2004). Colorectal cancers in a new mouse model of familial adenomatous polyposis: influence of genetic and environmental modifiers. *Lab Invest* 84, 1619-30.
- Conboy, M. J., Karasov, A. O. and Rando, T. A. (2007). High incidence of non-random template strand segregation and asymmetric fate determination in dividing stem cells and their progeny. *PLoS Biol* 5, e102.

Constancia, M., Dean, W., Lopes, S., Moore, T., Kelsey, G. and Reik, W. (2000). Deletion of a silencer element in *Igf2* results in loss of imprinting independent of *H19*. *Nat Genet* 26, 203-6.

Constancia, M., Hemberger, M., Hughes, J., Dean, W., Ferguson-Smith, A., Fundele, R., Stewart, F., Kelsey, G., Fowden, A., Sibley, C. and Reik, W. (2002). Placental-specific IGF-II is a major modulator of placental and fetal growth. *Nature* 417, 945-8.

Cooper, W. N., Luharia, A., Evans, G. A., Raza, H., Haire, A. C., Grundy, R., Bowdin, S. C., Riccio, A., Sebastio, G., Blik, J., Schofield, P. N., Reik, W., Macdonald, F. and Maher, E. R. (2005). Molecular subtypes and phenotypic expression of Beckwith-Wiedemann syndrome. *Eur J Hum Genet* 13, 1025-32.

Costa, F. F. (2007). Non-coding RNAs: lost in translation? *Gene* 386, 1-10.

Cui, H., Hedborg, F., He, L., Nordenskjold, A., Sandstedt, B., Pfeifer-Ohlsson, S. and Ohlsson, R. (1997). Inactivation of *H19*, an imprinted and putative tumor repressor gene, is a preneoplastic event during Wilms' tumorigenesis. *Cancer Res* 57, 4469-73.

## D

Danielson, K. G., Baribault, H., Holmes, D. F., Graham, H., Kadler, K. E. and Iozzo, R. V. (1997). Targeted disruption of decorin leads to abnormal collagen fibril morphology and skin fragility. *J Cell Biol* 136, 729-43.

Dao, D., Walsh, C. P., Yuan, L., Gorelov, D., Feng, L., Hensle, T., Nisen, P., Yamashiro, D. J., Bestor, T. H. and Tycko, B. (1999). Multipoint analysis of human chromosome 11p15/mouse distal chromosome 7: inclusion of *H19/IGF2* in the minimal WT2 region, gene specificity of *H19* silencing in Wilms' tumorigenesis and methylation hyper-dependence of *H19* imprinting. *Hum Mol Genet* 8, 1337-52.

Davies, K., Bowden, L., Smith, P., Dean, W., Hill, D., Furuumi, H., Sasaki, H., Cattanaach, B. and Reik, W. (2002). Disruption of mesodermal enhancers for *Igf2* in the minute mutant. *Development* 129, 1657-68.

Davis, E., Jensen, C. H., Schroder, H. D., Farnir, F., Shay-Hadfield, T., Kliem, A., Cockett, N., Georges, M. and Charlier, C. (2004). Ectopic expression of DLK1 protein in skeletal muscle of padumnal heterozygotes causes the callipyge phenotype. *Curr Biol* 14, 1858-62.

Davis, E., Caiment, F., Tordoir, X., Cavaille, J., Ferguson-Smith, A., Cockett, N., Georges, M. and Charlier, C. (2005). RNAi-mediated allelic trans-interaction at the imprinted *Rtl1/Peg11* locus. *Curr Biol* 15, 743-9.

Davis, R. L., Weintraub, H. and Lassar, A. B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* 51, 987-1000.

Davis, T. L., Trasler, J. M., Moss, S. B., Yang, G. J. and Bartolomei, M. S. (1999). Acquisition of the *H19* methylation imprint occurs differentially on the parental alleles during spermatogenesis. *Genomics* 58, 18-28.

- Davis, T. L., Yang, G. J., McCarrey, J. R. and Bartolomei, M. S. (2000). The *H19* methylation imprint is erased and re-established differentially on the parental alleles during male germ cell development. *Hum Mol Genet* 9, 2885-94.
- DeChiara, T. M., Efstratiadis, A. and Robertson, E. J. (1990). A growth-deficiency phenotype in heterozygous mice carrying an *insulin-like growth factor II* gene disrupted by targeting. *Nature* 345, 78-80.
- DeChiara, T. M., Robertson, E. J. and Efstratiadis, A. (1991). Parental imprinting of the mouse *insulin-like growth factor II* gene. *Cell* 64, 849-59.
- Delaval, K., Govin, J., Cerqueira, F., Rousseaux, S., Khochbin, S. and Feil, R. (2007). Differential histone modifications mark mouse imprinting control regions during spermatogenesis. *EMBO J* 26, 720-9.
- Deltour, L., Montagutelli, X., Guenet, J. L., Jami, J. and Paldi, A. (1995). Tissue- and developmental stage-specific imprinting of the mouse proinsulin gene, *Ins2*. *Dev Biol* 168, 686-8.
- Doench, J. G., Petersen, C. P. and Sharp, P. A. (2003). siRNAs can function as miRNAs. *Genes Dev* 17, 438-42.
- Drewell, R. A., Brenton, J. D., Ainscough, J. F., Barton, S. C., Hilton, K. J., Arney, K. L., Dandolo, L. and Surani, M. A. (2000). Deletion of a silencer element disrupts *H19* imprinting independently of a DNA methylation epigenetic switch. *Development* 127, 3419-28.
- Drewell, R. A., Arney, K. L., Arima, T., Barton, S. C., Brenton, J. D. and Surani, M. A. (2002a). Novel conserved elements upstream of the *H19* gene are transcribed and act as mesodermal enhancers. *Development* 129, 1205-13.
- Drewell, R. A., Goddard, C. J., Thomas, J. O. and Surani, M. A. (2002b). Methylation-dependent silencing at the *H19* imprinting control region by MeCP2. *Nucleic Acids Res* 30, 1139-44.
- Dupont, J. M., Tost, J., Jammes, H. and Gut, I. G. (2004). De novo quantitative bisulfite sequencing using the pyrosequencing technology. *Anal Biochem* 333, 119-27.
- Duvillie, B., Cordonnier, N., Deltour, L., Dandoy-Dron, F., Itier, J. M., Monthieux, E., Jami, J., Joshi, R. L. and Bucchini, D. (1997). Phenotypic alterations in insulin-deficient mutant mice. *Proc Natl Acad Sci U S A* 94, 5137-40.

## E

- Efstratiadis, A. (1998). Genetics of mouse growth. *Int J Dev Biol* 42, 955-76.
- Eggermann, T., Wollmann, H. A., Kuner, R., Eggermann, K., Enders, H., Kaiser, P. and Ranke, M. B. (1997). Molecular studies in 37 Silver-Russell syndrome patients: frequency and etiology of uniparental disomy. *Hum Genet* 100, 415-9.
- Eggermann, T., Meyer, E., Obermann, C., Heil, I., Schuler, H., Ranke, M. B., Eggermann, K. and Wollmann, H. A. (2005). Is maternal duplication of 11p15 associated with Silver-Russell syndrome? *J Med Genet* 42, e26.

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. *Hum Mol Genet* 15, 2945-54.

Enklaar, T., Zabel, B. U. and Prawitt, D. (2006). Beckwith-Wiedemann syndrome: multiple molecular mechanisms. *Expert Rev Mol Med* 8, 1-19.

## F

Fauque, P., Jouannet, P., Lesaffre, C., Ripoche, M. A., Dandolo, L., Vaiman, D. and Jammes, H. (2007). Assisted Reproductive Technology affects developmental kinetics, H19 Imprinting Control Region methylation and H19 gene expression in individual mouse embryos. *BMC Dev Biol* 7, 116.

Fedoriw, A. M., Stein, P., Svoboda, P., Schultz, R. M. and Bartolomei, M. S. (2004). Transgenic RNAi reveals essential function for CTCF in *H19* gene imprinting. *Science* 303, 238-40.

Feil, R. (2001). Early-embryonic culture and manipulation could affect genomic imprinting. *Trends Mol Med* 7, 245-6.

Feinberg, A. P. (1999). Imprinting of a genomic domain of 11p15 and loss of imprinting in cancer: an introduction. *Cancer Res* 59, 1743s-1746s.

Ferguson-Smith, A., Lin, S. P., Tsai, C. E., Youngson, N. and Tevendale, M. (2003). Genomic imprinting--insights from studies in mice. *Semin Cell Dev Biol* 14, 43-9.

Ferguson-Smith, A. C., Sasaki, H., Cattanach, B. M. and Surani, M. A. (1993). Parental-origin-specific epigenetic modification of the mouse *H19* gene. *Nature* 362, 751-5.

Filion, G. J., Zhenilo, S., Salozhin, S., Yamada, D., Prokhortchouk, E. and Defossez, P. A. (2006). A family of human zinc finger proteins that bind methylated DNA and repress transcription. *Mol Cell Biol* 26, 169-81.

Filippova, G. N. (2008). Genetics and epigenetics of the multifunctional protein CTCF. *Curr Top Dev Biol* 80, 337-60.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806-11.

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. *Nat Genet* 32, 426-31.

Fitzpatrick, G. V., Pugacheva, E. M., Shin, J. Y., Abdullaev, Z., Yang, Y., Khatod, K., Lobanenkov, V. V. and Higgins, M. J. (2007). Allele-specific binding of CTCF to the multipartite imprinting control region KvDMR1. *Mol Cell Biol* 27, 2636-47.

Forne, T., Oswald, J., Dean, W., Saam, J. R., Bailleul, B., Dandolo, L., Tilghman, S. M., Walter, J. and Reik, W. (1997). Loss of the maternal *H19* gene induces changes in *Igf2* methylation in both *cis* and *trans*. *Proc Natl Acad Sci U S A* 94, 10243-8.

Freking, B. A., Murphy, S. K., Wylie, A. A., Rhodes, S. J., Keele, J. W., Leymaster, K. A., Jirtle, R. L. and Smith, T. P. (2002). Identification of the single base change causing the callipyge muscle hypertrophy phenotype, the only known example of polar overdominance in mammals. *Genome Res* 12, 1496-506.

Frevel, M. A., Sowerby, S. J., Petersen, G. B. and Reeve, A. E. (1999). Methylation sequencing analysis refines the region of *H19* epimutation in Wilms tumor. *J Biol Chem* 274, 29331-40.

Fuks, F., Hurd, P. J., Deplus, R. and Kouzarides, T. (2003a). The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res* 31, 2305-12.

Fuks, F., Hurd, P. J., Wolf, D., Nan, X., Bird, A. P. and Kouzarides, T. (2003b). The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *J Biol Chem* 278, 4035-40.

Fuks, F. (2005). DNA methylation and histone modifications: teaming up to silence genes. *Curr Opin Genet Dev* 15, 490-5.

## G

Gabory, A. and Dandolo, L. (2005). [Epigenetics and development: genomic imprinting]. *Med Sci (Paris)* 21, 390-5.

Gabory, A., Ripoche, M. A., Yoshimizu, T. and Dandolo, L. (2006). The *H19* gene: regulation and function of a non-coding RNA. *Cytogenet Genome Res* 113, 188-93.

Georges, M., Charlier, C. and Cockett, N. (2003). The callipyge locus: evidence for the trans interaction of reciprocally imprinted genes. *Trends Genet* 19, 248-52.

Gerard, M., Hernandez, L., Wevrick, R. and Stewart, C. L. (1999). Disruption of the mouse *Necdin* gene results in early post-natal lethality. *Nat Genet* 23, 199-202.

Giannoukakis, N., Deal, C., Paquette, J., Goodyer, C. G. and Polychronakos, C. (1993). Parental genomic imprinting of the human *IGF2* gene. *Nat Genet* 4, 98-101.

Gicquel, C., Rossignol, S., Cabrol, S., Houang, M., Steunou, V., Barbu, V., Danton, F., Thibaud, N., Le Merrer, M., Burglen, L., Bertrand, A. M., Netchine, I. and Le Bouc, Y. (2005). Epimutation of the telomeric imprinting center region on chromosome 11p15 in Silver-Russell syndrome. *Nat Genet* 37, 1003-7.

Giddings, S. J., King, C. D., Harman, K. W., Flood, J. F. and Carnaghi, L. R. (1994). Allele specific inactivation of *insulin 1 and 2*, in the mouse yolk sac, indicates imprinting. *Nat Genet* 6, 310-3.

Girard, A., Sachidanandam, R., Hannon, G. J. and Carmell, M. A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* 442, 199-202.

Godfrey, P., Rahal, J. O., Beamer, W. G., Copeland, N. G., Jenkins, N. A. and Mayo, K. E. (1993). GHRH receptor of little mice contains a missense mutation in the extracellular domain that disrupts receptor function. *Nat Genet* 4, 227-32.

Goldberg, A. D., Allis, C. D. and Bernstein, E. (2007). Epigenetics: a landscape takes shape. *Cell* 128, 635-8.

Goll, M. G., Kirpekar, F., Maggert, K. A., Yoder, J. A., Hsieh, C. L., Zhang, X., Golic, K. G., Jacobsen, S. E. and Bestor, T. H. (2006). Methylation of tRNA<sup>Asp</sup> by the DNA methyltransferase homolog Dnmt2. *Science* 311, 395-8.

Gosden, R., Trasler, J., Lucifero, D. and Faddy, M. (2003). Rare congenital disorders, imprinted genes, and assisted reproductive technology. *Lancet* 361, 1975-7.

Grandjean, V., Smith, J., Schofield, P. N. and Ferguson-Smith, A. C. (2000). Increased IGF-II protein affects p57kip2 expression in vivo and in vitro: implications for Beckwith-Wiedemann syndrome. *Proc Natl Acad Sci U S A* 97, 5279-84.

Grunstein, M. (1997). Histone acetylation in chromatin structure and transcription. *Nature* 389, 349-52.

## H

Hagege, H., Klous, P., Braem, C., Splinter, E., Dekker, J., Cathala, G., de Laat, W. and Forne, T. (2007). Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nat Protoc* 2, 1722-33.

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. *Mech Dev* 117, 15-23.

Han, L., Lee, D. H. and Szabo, P. E. (2008). CTCF is the master organizer of domain-wide allele-specific chromatin at the *H19/Igf2* imprinted region. *Mol Cell Biol* 28, 1124-35.

Hansen, T. V., Hammer, N. A., Nielsen, J., Madsen, M., Dalbaeck, C., Wewer, U. M., Christiansen, J. and Nielsen, F. C. (2004). Dwarfism and impaired gut development in insulin-like growth factor II mRNA-binding protein 1-deficient mice. *Mol Cell Biol* 24, 4448-64.

Hao, Y., Crenshaw, T., Moulton, T., Newcomb, E. and Tycko, B. (1993). Tumour-suppressor activity of *H19* RNA. *Nature* 365, 764-7.

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, 486-9.

Hassan, A. B. and Howell, J. A. (2000). Insulin-like growth factor II supply modifies growth of intestinal adenoma in Apc(Min/+) mice. *Cancer Res* 60, 1070-6.

Hasty, P., Bradley, A., Morris, J. H., Edmondson, D. G., Venuti, J. M., Olson, E. N. and Klein, W. H. (1993). Muscle deficiency and neonatal death in mice with a targeted mutation in the *myogenin* gene. *Nature* 364, 501-6.

Hata, K., Okano, M., Lei, H. and Li, E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* 129, 1983-93.

Hawkins, P. G. and Morris, K. V. (2008). RNA and transcriptional modulation of gene expression. *Cell Cycle* 7, 602-7.

Heard, E., Clerc, P. and Avner, P. (1997). X-chromosome inactivation in mammals. *Annu Rev Genet* 31, 571-610.

Heard, E. and Disteché, C. M. (2006). Dosage compensation in mammals: fine-tuning the expression of the X chromosome. *Genes Dev* 20, 1848-67.

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 Dev* 22, 1607-16.

Holmgren, C., Kanduri, C., Dell, G., Ward, A., Mukhopadhyay, R., Kanduri, M., Lobanenko, V. and Ohlsson, R. (2001). CpG methylation regulates the *Igf2/H19* insulator. *Curr Biol* 11, 1128-30.

## I

Ishihara, K., Hatano, N., Furuumi, H., Kato, R., Iwaki, T., Miura, K., Jinno, Y. and Sasaki, H. (2000). Comparative genomic sequencing identifies novel tissue-specific enhancers and sequence elements for methylation-sensitive factors implicated in *Igf2/H19* imprinting. *Genome Res* 10, 664-71.

## J

Jelinic, P., Stehle, J. C. and Shaw, P. (2006). The testis-specific factor CTCFL cooperates with the protein methyltransferase PRMT7 in *H19* imprinting control region methylation. *PLoS Biol* 4, e355.

Jenuwein, T. and Allis, C. D. (2001). Translating the histone code. *Science* 293, 1074-80.

Jones, B. K., LeVorse, J. M. and Tilghman, S. M. (1998). *Igf2* imprinting does not require its own DNA methylation or H19 RNA. *Genes Dev* 12, 2200-7.

Jones, B. K., LeVorse, J. and Tilghman, S. M. (2001). Deletion of a nuclease-sensitive region between the *Igf2* and *H19* genes leads to *Igf2* misregulation and increased adiposity. *Hum Mol Genet* 10, 807-14.

Jones, N. C., Tyner, K. J., Nibarger, L., Stanley, H. M., Cornelison, D. D., Fedorov, Y. V. and Olwin, B. B. (2005). The p38alpha/beta MAPK functions as a molecular switch to activate the quiescent satellite cell. *J Cell Biol* 169, 105-16.

Jones, P. A. and Takai, D. (2001). The role of DNA methylation in mammalian epigenetics. *Science* 293, 1068-70.

Joshi, R. L., Lamothe, B., Cordonnier, N., Mesbah, K., Monthieux, E., Jami, J. and Bucchini, D. (1996). Targeted disruption of the *insulin receptor* gene in the mouse results in neonatal lethality. *EMBO J* 15, 1542-7.

Juan, V., Crain, C. and Wilson, C. (2000). Evidence for evolutionarily conserved secondary structure in the *H19* tumor suppressor RNA. *Nucleic Acids Res* 28, 1221-7.

**K**

Kaffer, C. R., Srivastava, M., Park, K. Y., Ives, E., Hsieh, S., Batlle, J., Grinberg, A., Huang, S. P. and Pfeifer, K. (2000). A transcriptional insulator at the imprinted *H19/Igf2* locus. *Genes Dev* 14, 1908-19.

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. *Curr Biol* 10, 853-6.

Kassar-Duchossoy, L., Gayraud-Morel, B., Gomes, D., Rocancourt, D., Buckingham, M., Shinin, V. and Tajbakhsh, S. (2004). *Mrf4* determines skeletal muscle identity in *Myf5:Myod* double-mutant mice. *Nature* 431, 466-71.

Kato, Y. and Sasaki, H. (2005). Imprinting and looping: epigenetic marks control interactions between regulatory elements. *Bioessays* 27, 1-4.

Kawahara, M., Wu, Q., Ferguson-Smith, A. C. and Kono, T. (2007a). Appropriate expression of imprinted genes on mouse chromosome 12 extends development of bi-maternal embryos to term. *FEBS Lett* 581, 5178-84.

Kawahara, M., Wu, Q., Takahashi, N., Morita, S., Yamada, K., Ito, M., Ferguson-Smith, A. C. and Kono, T. (2007b). High-frequency generation of viable mice from engineered bi-maternal embryos. *Nat Biotechnol* 25, 1045-50.

Kawaji, H. and Hayashizaki, Y. (2008). Exploration of small RNAs. *PLoS Genet* 4, e22.

Khvorova, A., Reynolds, A. and Jayasena, S. D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell* 115, 209-16.

Killian, J. K., Nolan, C. M., Stewart, N., Munday, B. L., Andersen, N. A., Nicol, S. and Jirtle, R. L. (2001). Monotreme *IGF2* expression and ancestral origin of genomic imprinting. *J Exp Zool* 291, 205-12.

Klose, R. J. and Bird, A. P. (2006). Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci* 31, 89-97.

Kohlmaier, A., Savarese, F., Lachner, M., Martens, J., Jenuwein, T. and Wutz, A. (2004). A chromosomal memory triggered by *Xist* regulates histone methylation in X inactivation. *PLoS Biol* 2, E171.

Kono, T., Obata, Y., Yoshimizu, T., Nakahara, T. and Carroll, J. (1996). Epigenetic modifications during oocyte growth correlates with extended parthenogenetic development in the mouse. *Nat Genet* 13, 91-4.

Kono, T., Sotomaru, Y., Katsuzawa, Y. and Dandolo, L. (2002). Mouse parthenogenetic embryos with monoallelic *H19* expression can develop to day 17.5 of gestation. *Dev Biol* 243, 294-300.

Kono, T., Obata, Y., Wu, Q., Niwa, K., Ono, Y., Yamamoto, Y., Park, E. S., Seo, J. S. and Ogawa, H. (2004). Birth of parthenogenetic mice that can develop to adulthood. *Nature* 428, 860-4.

Kono, T. (2006). Genomic imprinting is a barrier to parthenogenesis in mammals. *Cytogenet Genome Res* 113, 31-5.

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693-705.

Kuang, S., Kuroda, K., Le Grand, F. and Rudnicki, M. A. (2007). Asymmetric self-renewal and commitment of satellite stem cells in muscle. *Cell* 129, 999-1010.

Kuo, C. H., Uetsuki, T., Kim, C. H., Tanaka, H., Li, B. S., Taira, E., Higuchi, H., Okamoto, H., Yoshikawa, K. and Miki, N. (1995). Determination of a neclin cis-acting element required for neuron specific expression by using zebra fish. *Biochem Biophys Res Commun* 211, 438-46.

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Totoki, Y., Toyoda, A., Ikawa, M., Asada, N., Kojima, K., Yamaguchi, Y., Ijiri, T. W., Hata, K., Li, E., Matsuda, Y., Kimura, T., Okabe, M., Sakaki, Y., Sasaki, H. and Nakano, T. (2008). DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes Dev* 22, 908-17.

Kurukuti, S., Tiwari, V. K., Tavoosidana, G., Pugacheva, E., Murrell, A., Zhao, Z., Lobanenkov, V., Reik, W. and Ohlsson, R. (2006). CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to *Igf2*. *Proc Natl Acad Sci U S A* 103, 10684-9.

## L

Lansdorp, P. M. (2007). Immortal strands? Give me a break. *Cell* 129, 1244-7.

Le Grand, F. and Rudnicki, M. A. (2007). Skeletal muscle satellite cells and adult myogenesis. *Curr Opin Cell Biol* 19, 628-33.

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, 1807-17.

Lee, J. E., Pintar, J. and Efstratiadis, A. (1990). Pattern of the *insulin-like growth factor II* gene expression during early mouse embryogenesis. *Development* 110, 151-9.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S. and Kim, V. N. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, 415-9.

Lee, Y., Kim, M., Han, J., Yeom, K. H., Lee, S., Baek, S. H. and Kim, V. N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* 23, 4051-60.

Lefebvre, L., Viville, S., Barton, S. C., Ishino, F., Keverne, E. B. and Surani, M. A. (1998). Abnormal maternal behaviour and growth retardation associated with loss of the imprinted gene *Mest*. *Nat Genet* 20, 163-9.

Lehnertz, B., Ueda, Y., Derijck, A. A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T. and Peters, A. H. (2003). Suv39h-mediated histone H3 lysine 9

- methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* 13, 1192-200.
- 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, 3195-205.
- Leighton, P. A., Ingram, R. S., Eggenschwiler, J., Efstratiadis, A. and Tilghman, S. M. (1995a). Disruption of imprinting caused by deletion of the *H19* gene region in mice. *Nature* 375, 34-9.
- Leighton, P. A., Saam, J. R., Ingram, R. S., Stewart, C. L. and Tilghman, S. M. (1995b). An enhancer deletion affects both *H19* and *Igf2* expression. *Genes Dev* 9, 2079-89.
- 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, 865-73.
- 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. *Nat Genet* 36, 1291-5.
- Lewis, J. D., Meehan, R. R., Henzel, W. J., Maurer-Fogy, I., Jeppesen, P., Klein, F. and Bird, A. (1992). Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* 69, 905-14.
- Li, B., Carey, M. and Workman, J. L. (2007). The role of chromatin during transcription. *Cell* 128, 707-19.
- Li, J. Y., Lees-Murdock, D. J., Xu, G. L. and Walsh, C. P. (2004). Timing of establishment of paternal methylation imprints in the mouse. *Genomics* 84, 952-60.
- Li, L., Keverne, E. B., Aparicio, S. A., Ishino, F., Barton, S. C. and Surani, M. A. (1999). Regulation of maternal behavior and offspring growth by paternally expressed *Peg3*. *Science* 284, 330-3.
- Li, S., Czubryt, M. P., McAnally, J., Bassel-Duby, R., Richardson, J. A., Wiebel, F. F., Nordheim, A. and Olson, E. N. (2005). Requirement for serum response factor for skeletal muscle growth and maturation revealed by tissue-specific gene deletion in mice. *Proc Natl Acad Sci U S A* 102, 1082-7.
- Li, Y. M., Franklin, G., Cui, H. M., Svensson, K., He, X. B., Adam, G., Ohlsson, R. and Pfeifer, S. (1998). The *H19* transcript is associated with polysomes and may regulate *IGF2* expression in trans. *J Biol Chem* 273, 28247-52.
- Lin, S. C., Lin, C. R., Gukovsky, I., Lusk, A. J., Sawchenko, P. E. and Rosenfeld, M. G. (1993). Molecular basis of the little mouse phenotype and implications for cell type-specific growth. *Nature* 364, 208-13.
- Liu, J. P., Baker, J., Perkins, A. S., Robertson, E. J. and Efstratiadis, A. (1993). Mice carrying null mutations of the genes encoding *insulin-like growth factor I (Igf-1)* and *type I IGF receptor (Igf1r)*. *Cell* 75, 59-72.

Lopes, S., Lewis, A., Hajkova, P., Dean, W., Oswald, J., Forne, T., Murrell, A., Constancia, M., Bartolomei, M., Walter, J. and Reik, W. (2003). Epigenetic modifications in an imprinting cluster are controlled by a hierarchy of DMRs suggesting long-range chromatin interactions. *Hum Mol Genet* 12, 295-305.

Loukinov, D. I., Pugacheva, E., Vatolin, S., Pack, S. D., Moon, H., Chernukhin, I., Mannan, P., Larsson, E., Kanduri, C., Vostrov, A. A., Cui, H., Niemitz, E. L., Rasko, J. E., Docquier, F. M., Kistler, M., Breen, J. J., Zhuang, Z., Quitschke, W. W., Renkawitz, R., Klenova, E. M., Feinberg, A. P., Ohlsson, R., Morse, H. C., 3rd and Lobanenkov, V. V. (2002). BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma. *Proc Natl Acad Sci U S A* 99, 6806-11.

Louvi, A., Accili, D. and Efstratiadis, A. (1997). Growth-promoting interaction of IGF-II with the insulin receptor during mouse embryonic development. *Dev Biol* 189, 33-48.

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, 530-8.

Ludwig, T., Eggenschwiler, J., Fisher, P., D'Ercole, A. J., Davenport, M. L. and Efstratiadis, A. (1996). Mouse mutants lacking the type 2 IGF receptor (IGF2R) are rescued from perinatal lethality in *Igf2* and *Igf1r* null backgrounds. *Dev Biol* 177, 517-35.

Luedi, P. P., Hartemink, A. J. and Jirtle, R. L. (2005). Genome-wide prediction of imprinted murine genes. *Genome Res* 15, 875-84.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. and Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251-60.

Lui, J. C., Finkielstain, G. P., Barnes, K. M. and Baron, J. (2008). An imprinted gene network that controls mammalian somatic growth is down-regulated during postnatal growth deceleration in multiple organs. *Am J Physiol Regul Integr Comp Physiol* 295, R189-96.

Lustig-Yariv, O., Schulze, E., Komitowski, D., Erdmann, V., Schneider, T., de Groot, N. and Hochberg, A. (1997). The expression of the imprinted genes *H19* and *IGF-2* in choriocarcinoma cell lines. Is *H19* a tumor suppressor gene? *Oncogene* 15, 169-77.

## M

Maher, E. R. and Reik, W. (2000). Beckwith-Wiedemann syndrome: imprinting in clusters revisited. *J Clin Invest* 105, 247-52.

Mancini-Dinardo, D., Steele, S. J., Levorse, J. M., Ingram, R. S. and Tilghman, S. M. (2006). Elongation of the *Kcnqlot1* transcript is required for genomic imprinting of neighboring genes. *Genes Dev* 20, 1268-82.

Mann, J. R., Gadi, I., Harbison, M. L., Abbondanzo, S. J. and Stewart, C. L. (1990). Androgenetic mouse embryonic stem cells are pluripotent and cause skeletal defects in chimeras: implications for genetic imprinting. *Cell* 62, 251-60.

Mann, J. R. (2001). Imprinting in the germ line. *Stem Cells* 19, 287-94.

- Margueron, R., Trojer, P. and Reinberg, D. (2005). The key to development: interpreting the histone code? *Curr Opin Genet Dev* 15, 163-76.
- Matouk, I. J., DeGroot, N., Mezan, S., Ayesh, S., Abu-lail, R., Hochberg, A. and Galun, E. (2007). The *H19* non-coding RNA is essential for human tumor growth. *PLoS ONE* 2, e845.
- Mayer, W., Niveleau, A., Walter, J., Fundele, R. and Haaf, T. (2000). Demethylation of the zygotic paternal genome. *Nature* 403, 501-2.
- McGrath, J. and Solter, D. (1984). Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* 37, 179-83.
- McKinsey, T. A., Zhang, C. L., Lu, J. and Olson, E. N. (2000). Signal-dependent nuclear export of a histone deacetylase regulates muscle differentiation. *Nature* 408, 106-11.
- Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G. and Tuschl, T. (2004). Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol Cell* 15, 185-97.
- Meister, G. and Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. *Nature* 431, 343-9.
- Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T. K., Koche, R. P., Lee, W., Mendenhall, E., O'Donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Wernig, M., Jaenisch, R., Nusbaum, C., Lander, E. S. and Bernstein, B. E. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448, 553-60.
- Milligan, L., Antoine, E., Bisbal, C., Weber, M., Brunel, C., Forne, T. and Cathala, G. (2000). H19 gene expression is up-regulated exclusively by stabilization of the RNA during muscle cell differentiation. *Oncogene* 19, 5810-6.
- Mineno, J., Okamoto, S., Ando, T., Sato, M., Chono, H., Izu, H., Takayama, M., Asada, K., Mirochnitchenko, O., Inouye, M. and Kato, I. (2006). The expression profile of microRNAs in mouse embryos. *Nucleic Acids Res* 34, 1765-71.
- Miniou, P., Jeanpierre, M., Blanquet, V., Sibella, V., Bonneau, D., Herbelin, C., Fischer, A., Niveleau, A. and Viegas-Pequignot, E. (1994). Abnormal methylation pattern in constitutive and facultative (X inactive chromosome) heterochromatin of ICF patients. *Hum Mol Genet* 3, 2093-102.
- Miranda, T. B. and Jones, P. A. (2007). DNA methylation: the nuts and bolts of repression. *J Cell Physiol* 213, 384-90.
- Monk, D., Wagschal, A., Arnaud, P., Muller, P., Parker-Katirae, L., Bourchis, D., Scherer, S. W., Feil, R., Stanier, P. and Moore, G. E. (2008). Comparative analysis of human chromosome 7q21 and mouse proximal chromosome 6 reveals a placental-specific imprinted gene, *TFPI2/Tfpi2*, which requires G9a and Eed for allelic-silencing. *Genome Res*.
- Monk, M., Boubelik, M. and Lehnert, S. (1987). Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* 99, 371-82.

- Montarras, D., Morgan, J., Collins, C., Relaix, F., Zaffran, S., Cumano, A., Partridge, T. and Buckingham, M. (2005). Direct isolation of satellite cells for skeletal muscle regeneration. *Science* 309, 2064-7.
- Moon, Y. S., Smas, C. M., Lee, K., Villena, J. A., Kim, K. H., Yun, E. J. and Sul, H. S. (2002). Mice lacking paternally expressed *Pref-1/Dlk1* display growth retardation and accelerated adiposity. *Mol Cell Biol* 22, 5585-92.
- Moore, G. E., Abu-Amero, S. N., Bell, G., Wakeling, E. L., Kingsnorth, A., Stanier, P., Jauniaux, E. and Bennett, S. T. (2001). Evidence that *insulin* is imprinted in the human yolk sac. *Diabetes* 50, 199-203.
- Moore, T. and Haig, D. (1991). Genomic imprinting in mammalian development: a parental tug-of-war. *Trends Genet* 7, 45-9.
- Moore, T., Constancia, M., Zubair, M., Bailleul, B., Feil, R., Sasaki, H. and Reik, W. (1997). Multiple imprinted sense and antisense transcripts, differential methylation and tandem repeats in a putative imprinting control region upstream of mouse *Igf2*. *Proc Natl Acad Sci U S A* 94, 12509-14.
- Morison, I. M., Eccles, M. R. and Reeve, A. E. (2000). Imprinting of insulin-like growth factor 2 is modulated during hematopoiesis. *Blood* 96, 3023-8.
- Moulton, T., Crenshaw, T., Hao, Y., Moosikasuwon, J., Lin, N., Dembitzer, F., Hensle, T., Weiss, L., McMorro, L., Loew, T. and et al. (1994). Epigenetic lesions at the *H19* locus in Wilms' tumour patients. *Nat Genet* 7, 440-7.
- Murrell, A., Heeson, S., Bowden, L., Constancia, M., Dean, W., Kelsey, G. and Reik, W. (2001). An intragenic methylated region in the imprinted *Igf2* gene augments transcription. *EMBO Rep* 2, 1101-6.
- Murrell, A., Heeson, S. and Reik, W. (2004). Interaction between differentially methylated regions partitions the imprinted genes *Igf2* and *H19* into parent-specific chromatin loops. *Nat Genet* 36, 889-93.

## N

- Nabeshima, Y., Hanaoka, K., Hayasaka, M., Esumi, E., Li, S. and Nonaka, I. (1993). Myogenin gene disruption results in perinatal lethality because of severe muscle defect. *Nature* 364, 532-5.
- Nakamura, T., Arai, Y., Umehara, H., Masuhara, M., Kimura, T., Taniguchi, H., Sekimoto, T., Ikawa, M., Yoneda, Y., Okabe, M., Tanaka, S., Shiota, K. and Nakano, T. (2007). PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat Cell Biol* 9, 64-71.
- Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N. and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 393, 386-9.
- Netchine, I., Rossignol, S., Dufourg, M. N., Azzi, S., Rousseau, A., Perin, L., Houang, M., Steunou, V., Esteve, B., Thibaud, N., Demay, M. C., Danton, F., Petriczko, E., Bertrand, A.

M., Heinrichs, C., Carel, J. C., Loeuille, G. A., Pinto, G., Jacquemont, M. L., Gicquel, C., Cabrol, S. and Le Bouc, Y. (2007). 11p15 imprinting center region 1 loss of methylation is a common and specific cause of typical Russell-Silver syndrome: clinical scoring system and epigenetic-phenotypic correlations. *J Clin Endocrinol Metab* 92, 3148-54.

Ng, K., Pullirsch, D., Leeb, M. and Wutz, A. (2007). *Xist* and the order of silencing. *EMBO Rep* 8, 34-9.

Nicholls, R. D. and Knepper, J. L. (2001). Genome organization, function, and imprinting in Prader-Willi and Angelman syndromes. *Annu Rev Genomics Hum Genet* 2, 153-75.

Nielsen, F. C. and Christiansen, J. (1992). Endonucleolysis in the turnover of *insulin-like growth factor II* mRNA. *J Biol Chem* 267, 19404-11.

Nielsen, J., Christiansen, J., Lykke-Andersen, J., Johnsen, A. H., Wewer, U. M. and Nielsen, F. C. (1999). A family of *insulin-like growth factor II* mRNA-binding proteins represses translation in late development. *Mol Cell Biol* 19, 1262-70.

## O

Obata, Y., Kaneko-Ishino, T., Koide, T., Takai, Y., Ueda, T., Domeki, I., Shiroishi, T., Ishino, F. and Kono, T. (1998). Disruption of primary imprinting during oocyte growth leads to the modified expression of imprinted genes during embryogenesis. *Development* 125, 1553-60.

Obata, Y., Kono, T. and Hatada, I. (2002). Gene silencing: maturation of mouse fetal germ cells in vitro. *Nature* 418, 497.

Okano, M., Xie, S. and Li, E. (1998). Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucleic Acids Res* 26, 2536-40.

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, 247-57.

Ou, J. N., Torrisani, J., Unterberger, A., Provencal, N., Shikimi, K., Karimi, M., Ekstrom, T. J. and Szyf, M. (2007). Histone deacetylase inhibitor Trichostatin A induces global and gene-specific DNA demethylation in human cancer cell lines. *Biochem Pharmacol* 73, 1297-307.

## P

Pachnis, V., Belayew, A. and Tilghman, S. M. (1984). Locus unlinked to alpha-fetoprotein under the control of the murine raf and Rif genes. *Proc Natl Acad Sci U S A* 81, 5523-7.

Pachnis, V., Brannan, C. I. and Tilghman, S. M. (1988). The structure and expression of a novel gene activated in early mouse embryogenesis. *EMBO J* 7, 673-81.

Pant, V., Kurukuti, S., Pugacheva, E., Shamsuddin, S., Mariano, P., Renkawitz, R., Klenova, E., Lobanenkova, V. and Ohlsson, R. (2004). Mutation of a single CTCF target site within the *H19* imprinting control region leads to loss of *Igf2* imprinting and complex patterns of de novo methylation upon maternal inheritance. *Mol Cell Biol* 24, 3497-504.

Pauler, F. M., Koerner, M. V. and Barlow, D. P. (2007). Silencing by imprinted noncoding RNAs: is transcription the answer? *Trends Genet* 23, 284-92.

Paulsen, M., Takada, S., Youngson, N. A., Benchaib, M., Charlier, C., Segers, K., Georges, M. and Ferguson-Smith, A. C. (2001). Comparative sequence analysis of the imprinted Dlk1-Gtl2 locus in three mammalian species reveals highly conserved genomic elements and refines comparison with the Igf2-H19 region. *Genome Res* 11, 2085-94.

Pfaffle, R. W., DiMattia, G. E., Parks, J. S., Brown, M. R., Wit, J. M., Jansen, M., Van der Nat, H., Van den Brande, J. L., Rosenfeld, M. G. and Ingraham, H. A. (1992). Mutation of the POU-specific domain of Pit-1 and hypopituitarism without pituitary hypoplasia. *Science* 257, 1118-21.

Pfeifer, K. and Tilghman, S. M. (1994). Allele-specific gene expression in mammals: the curious case of the imprinted RNAs. *Genes Dev* 8, 1867-74.

Poirier, F., Chan, C. T., Timmons, P. M., Robertson, E. J., Evans, M. J. and Rigby, P. W. (1991). The murine *H19* gene is activated during embryonic stem cell differentiation in vitro and at the time of implantation in the developing embryo. *Development* 113, 1105-14.

Prawitt, D., Enklaar, T., Gartner-Rupprecht, B., Spangenberg, C., Oswald, M., Lausch, E., Schmidtke, P., Reutzler, D., Fees, S., Lucito, R., Korzon, M., Brozek, I., Limon, J., Housman, D. E., Pelletier, J. and Zabel, B. (2005). Microdeletion of target sites for insulator protein CTCF in a chromosome 11p15 imprinting center in Beckwith-Wiedemann syndrome and Wilms' tumor. *Proc Natl Acad Sci U S A* 102, 4085-90.

Preece, M. A., Price, S. M., Davies, V., Clough, L., Stanier, P., Trembath, R. C. and Moore, G. E. (1997). Maternal uniparental disomy 7 in Silver-Russell syndrome. *J Med Genet* 34, 6-9.

Prokhortchouk, A., Hendrich, B., Jorgensen, H., Ruzov, A., Wilm, M., Georgiev, G., Bird, A. and Prokhortchouk, E. (2001). The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. *Genes Dev* 15, 1613-8.

Purbowasito, W., Suda, C., Yokomine, T., Zubair, M., Sado, T., Tsutsui, K. and Sasaki, H. (2004). Large-scale identification and mapping of nuclear matrix-attachment regions in the distal imprinted domain of mouse chromosome 7. *DNA Res* 11, 391-407.

## Q

Qiu, X., Vu, T. H., Lu, Q., Ling, J. Q., Li, T., Hou, A., Wang, S. K., Chen, H. L., Hu, J. F. and Hoffman, A. R. (2008). A Complex DNA Looping Configuration Associated with the Silencing of the Maternal *Igf2* Allele. *Mol Endocrinol*.

## R

Radovick, S., Nations, M., Du, Y., Berg, L. A., Weintraub, B. D. and Wondisford, F. E. (1992). A mutation in the POU-homeodomain of Pit-1 responsible for combined pituitary hormone deficiency. *Science* 257, 1115-8.

Rana, T. M. (2007). Illuminating the silence: understanding the structure and function of small RNAs. *Nat Rev Mol Cell Biol* 8, 23-36.

Randhawa, R. and Cohen, P. (2005). The role of the insulin-like growth factor system in prenatal growth. *Mol Genet Metab* 86, 84-90.

- Reese, K. J., Lin, S., Verona, R. I., Schultz, R. M. and Bartolomei, M. S. (2007). Maintenance of paternal methylation and repression of the imprinted *H19* gene requires MBD3. *PLoS Genet* 3, e137.
- Reik, W., Dean, W. and Walter, J. (2001). Epigenetic reprogramming in mammalian development. *Science* 293, 1089-93.
- Reik, W., Santos, F., Mitsuya, K., Morgan, H. and Dean, W. (2003). Epigenetic asymmetry in the mammalian zygote and early embryo: relationship to lineage commitment? *Philos Trans R Soc Lond B Biol Sci* 358, 1403-9; discussion 1409.
- Relaix, F., Rocancourt, D., Mansouri, A. and Buckingham, M. (2004). Divergent functions of murine Pax3 and Pax7 in limb muscle development. *Genes Dev* 18, 1088-105.
- Relaix, F., Rocancourt, D., Mansouri, A. and Buckingham, M. (2005). A Pax3/Pax7-dependent population of skeletal muscle progenitor cells. *Nature* 435, 948-53.
- Relaix, F., Montarras, D., Zaffran, S., Gayraud-Morel, B., Rocancourt, D., Tajbakhsh, S., Mansouri, A., Cumano, A. and Buckingham, M. (2006). Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells. *J Cell Biol* 172, 91-102.
- Riedemann, J. and Macaulay, V. M. (2006). IGF1R signalling and its inhibition. *Endocr Relat Cancer* 13 Suppl 1, S33-43.
- Ringrose, L. and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu Rev Genet* 38, 413-43.
- Ripoche, M. A., Kress, C., Poirier, F. and Dandolo, L. (1997). Deletion of the *H19* transcription unit reveals the existence of a putative imprinting control element. *Genes Dev* 11, 1596-604.
- Rosenblatt, J. D., Lunt, A. I., Parry, D. J. and Partridge, T. A. (1995). Culturing satellite cells from living single muscle fiber explants. *In Vitro Cell Dev Biol Anim* 31, 773-9.
- Rougeulle, C., Chaumeil, J., Sarma, K., Allis, C. D., Reinberg, D., Avner, P. and Heard, E. (2004). Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Mol Cell Biol* 24, 5475-84.
- Rountree, M. R., Bachman, K. E. and Baylin, S. B. (2000). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nat Genet* 25, 269-77.
- Royo, H., Bortolin, M. L., Seitz, H. and Cavaille, J. (2006). Small non-coding RNAs and genomic imprinting. *Cytogenet Genome Res* 113, 99-108.
- Royo, H. and Cavaille, J. (2008). Non-coding RNAs in imprinted gene clusters. *Biol Cell* 100, 149-66.
- Rudnicki, M. A., Schnegelsberg, P. N., Stead, R. H., Braun, T., Arnold, H. H. and Jaenisch, R. (1993). MyoD or Myf-5 is required for the formation of skeletal muscle. *Cell* 75, 1351-9.
- Rump, P., Zeegers, M. P. and van Essen, A. J. (2005). Tumor risk in Beckwith-Wiedemann syndrome: A review and meta-analysis. *Am J Med Genet A* 136, 95-104.

Runge, S., Nielsen, F. C., Nielsen, J., Lykke-Andersen, J., Wewer, U. M. and Christiansen, J. (2000). *H19* RNA binds four molecules of insulin-like growth factor II mRNA-binding protein. *J Biol Chem* 275, 29562-9.

Ruther, U., Woodroffe, C., Fattori, E. and Ciliberto, G. (1993). Inducible formation of liver tumors in transgenic mice. *Oncogene* 8, 87-93.

## S

Sakatani, T., Kaneda, A., Iacobuzio-Donahue, C. A., Carter, M. G., de Boom Witzel, S., Okano, H., Ko, M. S., Ohlsson, R., Longo, D. L. and Feinberg, A. P. (2005). Loss of imprinting of *Igf2* alters intestinal maturation and tumorigenesis in mice. *Science* 307, 1976-8.

Sambasivan, R. and Tajbakhsh, S. (2007). Skeletal muscle stem cell birth and properties. *Semin Cell Dev Biol* 18, 870-82.

Sasaki, H., Ferguson-Smith, A. C., Shum, A. S., Barton, S. C. and Surani, M. A. (1995). Temporal and spatial regulation of *H19* imprinting in normal and uniparental mouse embryos. *Development* 121, 4195-202.

Sasaki, T., Shiohama, A., Minoshima, S. and Shimizu, N. (2003). Identification of eight members of the Argonaute family in the human genome small star, filled. *Genomics* 82, 323-30.

Schmidt, J. V., Leverage, J. M. and Tilghman, S. M. (1999). Enhancer competition between *H19* and *Igf2* does not mediate their imprinting. *Proc Natl Acad Sci U S A* 96, 9733-8.

Schoeftner, S., Sengupta, A. K., Kubicek, S., Mechtler, K., Spahn, L., Koseki, H., Jenuwein, T. and Wutz, A. (2006). Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *EMBO J* 25, 3110-22.

Schoenherr, C. J., Leverage, J. M. and Tilghman, S. M. (2003). CTCF maintains differential methylation at the *Igf2/H19* locus. *Nat Genet* 33, 66-9.

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B. and Cavalli, G. (2007). Genome regulation by polycomb and trithorax proteins. *Cell* 128, 735-45.

Seitz, H., Royo, H., Bortolin, M. L., Lin, S. P., Ferguson-Smith, A. C. and Cavaille, J. (2004). A large imprinted microRNA gene cluster at the mouse *Dlk1-Gtl2* domain. *Genome Res* 14, 1741-8.

Sekita, Y., Wagatsuma, H., Nakamura, K., Ono, R., Kagami, M., Wakisaka, N., Hino, T., Suzuki-Migishima, R., Kohda, T., Ogura, A., Ogata, T., Yokoyama, M., Kaneko-Ishino, T. and Ishino, F. (2008). Role of retrotransposon-derived imprinted gene, *Rtl1*, in the fetomaternal interface of mouse placenta. *Nat Genet* 40, 243-8.

SenGupta, D. J., Zhang, B., Kraemer, B., Pochart, P., Fields, S. and Wickens, M. (1996). A three-hybrid system to detect RNA-protein interactions in vivo. *Proc Natl Acad Sci U S A* 93, 8496-501.

Shin, J. Y., Fitzpatrick, G. V. and Higgins, M. J. (2008). Two distinct mechanisms of silencing by the KvDMR1 imprinting control region. *EMBO J* 27, 168-78.

- Shinin, V., Gayraud-Morel, B., Gomes, D. and Tajbakhsh, S. (2006). Asymmetric division and cosegregation of template DNA strands in adult muscle satellite cells. *Nat Cell Biol* 8, 677-87.
- Shogren-Knaak, M., Ishii, H., Sun, J. M., Pazin, M. J., Davie, J. R. and Peterson, C. L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* 311, 844-7.
- Sinkkonen, L., Hugenschmidt, T., Berninger, P., Gaidatzis, D., Mohn, F., Artus-Revel, C. G., Zavolan, M., Svoboda, P. and Filipowicz, W. (2008). MicroRNAs control de novo DNA methylation through regulation of transcriptional repressors in mouse embryonic stem cells. *Nat Struct Mol Biol* 15, 259-67.
- Sleutels, F., Zwart, R. and Barlow, D. P. (2002). The non-coding *Air* RNA is required for silencing autosomal imprinted genes. *Nature* 415, 810-3.
- Smilnich, N. J., Day, C. D., Fitzpatrick, G. V., Caldwell, G. M., Lossie, A. C., Cooper, P. R., Smallwood, A. C., Joyce, J. A., Schofield, P. N., Reik, W., Nicholls, R. D., Weksberg, R., Driscoll, D. J., Maher, E. R., Shows, T. B. and Higgins, M. J. (1999). A maternally methylated CpG island in *KvLQT1* is associated with an antisense paternal transcript and loss of imprinting in Beckwith-Wiedemann syndrome. *Proc Natl Acad Sci U S A* 96, 8064-9.
- Smith, F. M., Garfield, A. S. and Ward, A. (2006). Regulation of growth and metabolism by imprinted genes. *Cytogenet Genome Res* 113, 279-91.
- Smits, G., Mungall, A. J., Griffiths-Jones, S., Smith, P., Beury, D., Matthews, L., Rogers, J., Pask, A. J., Shaw, G., VandeBerg, J. L., McCarrey, J. R., the SAVOIR consortium, Renfree, M. B., Reik, W. and Dunham, I. (2008). Conservation of the *H19* non-coding RNA and *H19-Igf2* imprinting mechanism in therians. *Nat Genet*.
- Sparago, A., Cerrato, F., Vernucci, M., Ferrero, G. B., Silengo, M. C. and Riccio, A. (2004). Microdeletions in the human *H19* DMR result in loss of *IGF2* imprinting and Beckwith-Wiedemann syndrome. *Nat Genet* 36, 958-60.
- Sparanese, D. and Lee, C. H. (2007). CRD-BP shields c-myc and MDR-1 RNA from endonucleolytic attack by a mammalian endoribonuclease. *Nucleic Acids Res* 35, 1209-21.
- Srivastava, M., Hsieh, S., Grinberg, A., Williams-Simons, L., Huang, S. P. and Pfeifer, K. (2000). *H19* and *Igf2* monoallelic expression is regulated in two distinct ways by a shared *cis* acting regulatory region upstream of *H19*. *Genes Dev* 14, 1186-95.
- Stevens, L. C. (1970). The development of transplantable teratocarcinomas from intratesticular grafts of pre- and postimplantation mouse embryos. *Dev Biol* 21, 364-82.
- Suetake, I., Shinozaki, F., Miyagawa, J., Takeshima, H. and Tajima, S. (2004). DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. *J Biol Chem* 279, 27816-23.
- Sun, F. L., Dean, W. L., Kelsey, G., Allen, N. D. and Reik, W. (1997). Transactivation of *Igf2* in a mouse model of Beckwith-Wiedemann syndrome. *Nature* 389, 809-15.

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, 548-50.

Svensson, K., Walsh, C., Fundele, R. and Ohlsson, R. (1995). *H19* is imprinted in the choroid plexus and leptomeninges of the mouse foetus. *Mech Dev* 51, 31-7.

Szabo, P., Tang, S. H., Rentsendorj, A., Pfeifer, G. P. and Mann, J. R. (2000). Maternal-specific footprints at putative CTCF sites in the *H19* imprinting control region give evidence for insulator function. *Curr Biol* 10, 607-10.

Szabo, P. E. and Mann, J. R. (1995a). Biallelic expression of imprinted genes in the mouse germ line: implications for erasure, establishment, and mechanisms of genomic imprinting. *Genes Dev* 9, 1857-68.

Szabo, P. E. and Mann, J. R. (1995b). Allele-specific expression and total expression levels of imprinted genes during early mouse development: implications for imprinting mechanisms. *Genes Dev* 9, 3097-108.

Szabo, P. E., Hubner, K., Scholer, H. and Mann, J. R. (2002). Allele-specific expression of imprinted genes in mouse migratory primordial germ cells. *Mech Dev* 115, 157-60.

Szabo, P. E., Tang, S. H., Silva, F. J., Tsark, W. M. and Mann, J. R. (2004). Role of CTCF binding sites in the *Igf2/H19* imprinting control region. *Mol Cell Biol* 24, 4791-800.

## T

Tajbakhsh, S., Rocancourt, D., Cossu, G. and Buckingham, M. (1997). Redefining the genetic hierarchies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD. *Cell* 89, 127-38.

Takai, D. and Jones, P. A. (2003). The CpG island searcher: a new WWW resource. *In Silico Biol* 3, 235-40.

Tatsumi, R., Anderson, J. E., Nevoret, C. J., Halevy, O. and Allen, R. E. (1998). HGF/SF is present in normal adult skeletal muscle and is capable of activating satellite cells. *Dev Biol* 194, 114-28.

Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D. and Patel, D. J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* 14, 1025-40.

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 Dev* 12, 3693-702.

Thorvaldsen, J. L., Mann, M. R., Nwoko, O., Duran, K. L. and Bartolomei, M. S. (2002). Analysis of sequence upstream of the endogenous *H19* gene reveals elements both essential and dispensable for imprinting. *Mol Cell Biol* 22, 2450-62.

Thorvaldsen, J. L. and Bartolomei, M. S. (2007). SnapShot: imprinted gene clusters. *Cell* 130, 958.

- Tilghman, S. (2006). The making of a President: an interview with Shirley Tilghman. Interview by Jane Gitschier. *PLoS Genet* 2, e82.
- Tost, J., El Abdalaoui, H. and Gut, I. G. (2006). Serial pyrosequencing for quantitative DNA methylation analysis. *Biotechniques* 40, 721-2, 724, 726.
- Tourriere, H., Gallouzi, I. E., Chebli, K., Capony, J. P., Mouaikel, J., van der Geer, P. and Tazi, J. (2001). RasGAP-associated endoribonuclease G3Bp: selective RNA degradation and phosphorylation-dependent localization. *Mol Cell Biol* 21, 7747-60.
- Tremblay, K. D., Saam, J. R., Ingram, R. S., Tilghman, S. M. and Bartolomei, M. S. (1995). A paternal-specific methylation imprint marks the alleles of the mouse *H19* gene. *Nat Genet* 9, 407-13.
- Tremblay, K. D., Duran, K. L. and Bartolomei, M. S. (1997). A 5' 2-kilobase-pair region of the imprinted mouse *H19* gene exhibits exclusive paternal methylation throughout development. *Mol Cell Biol* 17, 4322-9.
- Tremblay, P., Dietrich, S., Mericskay, M., Schubert, F. R., Li, Z. and Paulin, D. (1998). A crucial role for Pax3 in the development of the hypaxial musculature and the long-range migration of muscle precursors. *Dev Biol* 203, 49-61.

## U

- Ueda, T., Abe, K., Miura, A., Yuzuriha, M., Zubair, M., Noguchi, M., Niwa, K., Kawase, Y., Kono, T., Matsuda, Y., Fujimoto, H., Shibata, H., Hayashizaki, Y. and Sasaki, H. (2000). The paternal methylation imprint of the mouse *H19* locus is acquired in the gonocyte stage during foetal testis development. *Genes Cells* 5, 649-59.
- Umlauf, D., Goto, Y., Cao, R., Cerqueira, F., Wagschal, A., Zhang, Y. and Feil, R. (2004). Imprinting along the *Kcnq1* domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nat Genet* 36, 1296-300.

## V

- Vaissiere, T., Sawan, C. and Herceg, Z. (2008). Epigenetic interplay between histone modifications and DNA methylation in gene silencing. *Mutat Res*.
- van Dijk, E. L., Sussenbach, J. S. and Holthuizen, P. E. (2000). Distinct RNA structural domains cooperate to maintain a specific cleavage site in the 3'-UTR of IGF-II mRNAs. *J Mol Biol* 300, 449-67.
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3.
- Varrault, A., Gueydan, C., Delalbre, A., Bellmann, A., Houssami, S., Aknin, C., Severac, D., Chotard, L., Kahli, M., Le Digarcher, A., Pavlidis, P. and Journot, L. (2006). *Zac1* regulates an imprinted gene network critically involved in the control of embryonic growth. *Dev Cell* 11, 711-22.

Verkerk, A. J., Ariel, I., Dekker, M. C., Schneider, T., van Gurp, R. J., de Groot, N., Gillis, A. J., Oosterhuis, J. W., Hochberg, A. A. and Looijenga, L. H. (1997). Unique expression patterns of *H19* in human testicular cancers of different etiology. *Oncogene* 14, 95-107.

Vernucci, M., Cerrato, F., Besnard, N., Casola, S., Pedone, P. V., Bruni, C. B. and Riccio, A. (2000). The *H19* endodermal enhancer is required for *Igf2* activation and tumor formation in experimental liver carcinogenesis. *Oncogene* 19, 6376-85.

Vernucci, M., Cerrato, F., Pedone, P. V., Dandolo, L., Bruni, C. B. and Riccio, A. (2004). Developmentally regulated functions of the *H19* differentially methylated domain. *Hum Mol Genet* 13, 353-61.

## W

Wagschal, A. and Feil, R. (2006). Genomic imprinting in the placenta. *Cytogenet Genome Res* 113, 90-8.

Wan, L. B. and Bartolomei, M. S. (2008). Chapter 7 Regulation of Imprinting in Clusters: Noncoding RNAs Versus Insulators. *Adv Genet* 61, 207-23.

Wang, G. G., Allis, C. D. and Chi, P. (2007a). Chromatin remodeling and cancer, Part I: Covalent histone modifications. *Trends Mol Med* 13, 363-72.

Wang, G. G., Allis, C. D. and Chi, P. (2007b). Chromatin remodeling and cancer, Part II: ATP-dependent chromatin remodeling. *Trends Mol Med* 13, 373-80.

Wang, Z. Q., Fung, M. R., Barlow, D. P. and Wagner, E. F. (1994). Regulation of embryonic growth and lysosomal targeting by the imprinted *Igf2/Mpr* gene. *Nature* 372, 464-7.

Weber, M., Milligan, L., Delalbre, A., Antoine, E., Brunel, C., Cathala, G. and Forne, T. (2001). Extensive tissue-specific variation of allelic methylation in the *Igf2* gene during mouse fetal development: relation to expression and imprinting. *Mech Dev* 101, 133-41.

Weber, M., Hagege, H., Murrell, A., Brunel, C., Reik, W., Cathala, G. and Forne, T. (2003). Genomic imprinting controls matrix attachment regions in the *Igf2* gene. *Mol Cell Biol* 23, 8953-9.

Weber, M., Davies, J. J., Wittig, D., Oakeley, E. J., Haase, M., Lam, W. L. and Schubeler, D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat Genet* 37, 853-62.

Weber, M., Hellmann, I., Stadler, M. B., Ramos, L., Paabo, S., Rebhan, M. and Schubeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* 39, 457-66.

Weber, M. and Schubeler, D. (2007). Genomic patterns of DNA methylation: targets and function of an epigenetic mark. *Curr Opin Cell Biol* 19, 273-80.

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., Meachem, S. J., Antonarakis, S. E., de Kretser, D. M., Hedger, M. P., Peterson, P., Carroll, B. J. and Scott, H. S. (2005). Meiotic and epigenetic defects in *Dnmt3L*-knockout mouse spermatogenesis. *Proc Natl Acad Sci U S A* 102, 4068-73.

Weksberg, R., Smith, A. C., Squire, J. and Sadowski, P. (2003). Beckwith-Wiedemann syndrome demonstrates a role for epigenetic control of normal development. *Hum Mol Genet* 12 Spec No 1, R61-8.

Wiles, M. V. (1988). Isolation of differentially expressed human cDNA clones: similarities between mouse and human embryonal carcinoma cell differentiation. *Development* 104, 403-13.

Wilkin, F., Paquette, J., Ledru, E., Hamelin, C., Pollak, M. and Deal, C. L. (2000). *H19* sense and antisense transgenes modify *insulin-like growth factor-II* mRNA levels. *Eur J Biochem* 267, 4020-7.

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, 745-9.

Wutz, A. and Jaenisch, R. (2000). A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol Cell* 5, 695-705.

Wutz, A. and Gribnau, J. (2007). X inactivation Xplained. *Curr Opin Genet Dev* 17, 387-93.

## Y

Yan, Z., Choi, S., Liu, X., Zhang, M., Schageman, J. J., Lee, S. Y., Hart, R., Lin, L., Thurmond, F. A. and Williams, R. S. (2003). Highly coordinated gene regulation in mouse skeletal muscle regeneration. *J Biol Chem* 278, 8826-36.

Yoo-Warren, H., Pachnis, V., Ingram, R. S. and Tilghman, S. M. (1988). Two regulatory domains flank the mouse *H19* gene. *Mol Cell Biol* 8, 4707-15.

Yoon, B., Herman, H., Hu, B., Park, Y. J., Lindroth, A., Bell, A., West, A. G., Chang, Y., Stablewski, A., Piel, J. C., Loukinov, D. I., Lobanenkova, V. V. and Soloway, P. D. (2005). *Rasgrfl* imprinting is regulated by a CTCF-dependent methylation-sensitive enhancer blocker. *Mol Cell Biol* 25, 11184-90.

Yoshimizu, T., Miroglio, A., Ripoche, M. A., Gabory, A., Vernucci, M., Riccio, A., Colnot, S., Godard, C., Terris, B., Jammes, H. and Dandolo, L. (2008). The *H19* locus acts *in vivo* as a tumour suppressor. *Proc Natl Acad Sci U S A* 105, 12417-22.

Yu, S., Yu, D., Lee, E., Eckhaus, M., Lee, R., Corria, Z., Accili, D., Westphal, H. and Weinstein, L. S. (1998). Variable and tissue-specific hormone resistance in heterotrimeric Gs protein alpha-subunit (G $\alpha$ ) knockout mice is due to tissue-specific imprinting of the *galpha* gene. *Proc Natl Acad Sci U S A* 95, 8715-20.

## Z

Zekri, L., Chebli, K., Tourriere, H., Nielsen, F. C., Hansen, T. V., Rami, A. and Tazi, J. (2005). Control of fetal growth and neonatal survival by the RasGAP-associated endoribonuclease G3BP. *Mol Cell Biol* 25, 8703-16.

Zemel, S., Bartolomei, M. S. and Tilghman, S. M. (1992). Physical linkage of two mammalian imprinted genes, *H19* and *insulin-like growth factor 2*. *Nat Genet* 2, 61-5.

Zhang, P., Liegeois, N. J., Wong, C., Finegold, M., Hou, H., Thompson, J. C., Silverman, A., Harper, J. W., DePinho, R. A. and Elledge, S. J. (1997). Altered cell differentiation and proliferation in mice lacking p57KIP2 indicates a role in Beckwith-Wiedemann syndrome. *Nature* 387, 151-8.

Zhang, Y. and Tycko, B. (1992). Monoallelic expression of the human *H19* gene. *Nat Genet* 1, 40-4.

---

## Etude de la fonction de l'ARN non codant *H19* et de la régulation de l'Empreinte Parentale au locus *H19/Igf2* dans des modèles murins transgéniques.

Le gène *H19* a été découvert il y a un vingtaine d'années et sa fonction reste encore inconnue. Ce gène est soumis à l'Empreinte Parentale, son expression est monoallélique maternelle. Par ailleurs le transcrit produit à partir de ce gène est un ARN non codant de 2,3 kb. Il est fortement exprimé au cours de l'embryogenèse dans tous les tissus issus du mésoderme et de l'endoderme et est réprimé après la naissance. Le seul organe où il reste fortement exprimé chez l'adulte est le muscle squelettique. Le gène *H19* est lié au niveau de sa localisation chromosomique et de la régulation transcriptionnelle au gène *Igf2*, qui code un facteur de croissance embryonnaire exprimé de façon monoallélique paternelle. L'invalidation du gène *H19* chez la souris entraîne un phénotype de surcroissance attribué à la perte d'Empreinte du gène *Igf2*.

Nous avons voulu adresser deux questions portant d'une part sur la fonction de l'ARN *H19* au cours de l'embryogenèse et d'autre part sur la régulation de ce locus dans le muscle adulte. L'originalité de notre approche réside dans l'étude *in vivo* de modèles murins transgéniques. Nos résultats montrent, à partir de trois modèles murins de tumorigenèse induite, que le gène *H19* a une fonction de gène suppresseur de tumeur *in vivo*. Par ailleurs nous avons montré à partir de souris exprimant le gène *H19* de façon ectopique que cet ARN est un *trans* régulateur d'un réseau de gènes soumis à l'Empreinte, impliqué dans la régulation de la croissance. Enfin, nous avons observé une expression biallélique des gènes *H19* et *Igf2* dans les cellules satellites, qui sont les cellules souches musculaires, ainsi que dans les cellules souches hématopoïétiques, suggérant qu'un relâchement de l'Empreinte puisse être une caractéristique des cellules souches adultes.

---

## Analysis of the *H19* non-coding RNA function and the *H19/Igf2* locus imprinting regulation in transgenic mouse models.

The *H19* gene was first described twenty years ago and its function remains unknown. This gene is imprinted; it is expressed only from the maternal allele. The product of the gene is a 2.3 kb non-coding RNA abundantly expressed in mesoderm and endoderm derived tissues during embryogenesis. After birth, the expression is repressed in all tissues except skeletal muscle where it remains strongly expressed at adult stage. The *H19* gene is linked genetically to the paternally expressed *Igf2* gene, which encodes an embryonic growth factor, and they share the mechanism of transcriptional regulation. The deletion of the *H19* gene in mice leads to an overgrowth phenotype, which is attributed to a loss of imprinting of the *Igf2* gene.

We addressed the questions of the role of the *H19* RNA during embryogenesis and of the regulation of this imprinted locus in adult muscle tissue. The particularity of our approach is the use of *in vivo* transgenic mouse models. We showed in three independent models of induced tumorigenesis that the *H19* gene acts as a tumour suppressor *in vivo*. We then analysed transgenic animals which ectopically expressed the *H19* gene and found that this RNA is a *trans* regulator of a recently described imprinted gene network, implicated in growth regulation. Finally we observed a biallelic expression of the *H19* and *Igf2* genes in satellites cells, which are muscle stem cells, as well as in haematopoietic stem cells, suggesting that a relaxation of imprinting could be a characteristic of adult stem cells.