

UNIVERSITE LILLE NORD DE FRANCE
Laboratoire INSERM U908, Université de Lille

THESE DE DOCTORAT

Ecole Doctorale Biologie-Santé de Lille (ED446)

Présentée et soutenue publiquement par

Constance VENNIN

Le 04 Mars 2016

**CONTRIBUTION ET ROLES DANS LA
TUMORIGENESE DES ARN NON CODANTS
TRANSCRITS AU LOCUS *H19/IGF2* : *H19* ET *91H***

Jury

Président : Pr Xuefen LE BOURHIS

Rapporteurs : Dr Thierry FORNE
Pr Philippe BECUWE

Examineurs : Dr Bruno LEFEBVRE
Pr Franck TOLEDO

Directeur : Pr Eric ADRIAENSSENS

REMERCIEMENTS

Je souhaite remercier dans un premier temps les directeurs de mes laboratoires d'accueil : le Dr **Yvan DE LAUNOIT** à l'unité CNRS UMR8161 et le Pr **Xuefen LE BOURHIS** à l'unité INSERM U908 de m'avoir accueillie dans leur laboratoire pour la réalisation de ma thèse.

Je remercie les membres du jury, les Dr **Thierry FORNE**, Pr **Philippe BECUWE**, Pr **Franck TOLEDO**, Dr **Bruno LEFEBVRE** d'avoir pris le temps de lire ce manuscrit. Je vous remercie de me consacrer une journée dans votre emploi du temps déjà chargé, pour assister à ma soutenance de thèse. Je remercie également les rapporteurs, le Dr **Thierry FORNE** et le Pr **Philippe BECUWE** pour leurs critiques constructives.

Je tiens à remercier la Fondation ARC pour son soutien financier qu'elle m'a accordée. Cette petite prolongation m'a permis d'approfondir mes travaux de thèses.

Je tiens à remercier Monsieur le Pr **Eric ADRIAENSSENS** de m'avoir encadrée pendant ces 3 années de thèse ainsi que pendant mon master 2. Je te remercie de m'avoir donné l'opportunité de mener à bien cette thèse, de m'avoir fait confiance et d'avoir cru en moi. Je te remercie pour tes précieux conseils et tes idées qui surviennent en pleine nuit ! Eric, tes « *j'ai réfléchi cette nuit, pendant que je ne dormais pas, et je me suis dit...* » vont me manquer ! Merci d'avoir été là dans les moments où je n'y croyais plus. Pour finir, il faut bien que je l'admette : « *le chef a très souvent raison... pas toujours quand même !* »

Un grand merci à **Nathalie SPRUYT** pour son aide technique au cours de ces 4 années. Je la remercie également pour ses précieux conseils, ses « trucs et astuces » qui facilitent la vie de laboratoire ! Merci de m'avoir enseigné toutes les techniques que nous utilisons en routine. Nathalie, tes petits gâteaux, notamment tes macarons, m'ont manqué lors de ma dernière année ! J'espère que tu es bien intégrée dans ta nouvelle équipe.

Je remercie également les Dr **Roland BOURETTE** et **Jérôme VICOONE** pour les conseils qu'ils ont pu me donner. Roland, merci de m'avoir enseigné les bases permettant l'analyse des cellules souches et de m'avoir fait confiance pour les révisions de l'article. Je te remercie également pour ton soutien !

Je remercie **Alexandra** pour tous les bons moments passés au laboratoire ! Je te remercie pour tous tes conseils. Nos échanges autour d'un café m'ont bien manqué pendant ma dernière année, heureusement que la salle de sport était toujours là. Merci d'avoir été là et de m'avoir soutenue dans toutes les péripéties de la thèse !

Je remercie **Claire**, ma copine de thèse ! Dans la même galère, au même moment. Je te remercie pour ton soutien perpétuel au cours de ces trois années. Je te remercie d'avoir partagé avec moi mes doutes et mes inquiétudes. Merci de m'avoir soutenue à tout moment ! Je te souhaite le meilleur ! « *Oui, Claire, on va y arriver !! Nous devons croire en nous !* »

Merci à **Hélène**, d'avoir partagé mon bureau pendant un peu plus d'un an. Nos nombreuses discussions nous ont permis de publier mon premier article ! Merci d'avoir été là, de m'avoir écoutée, réconfortée quand rien ne fonctionnait, d'abord dans notre bureau, puis au détour d'un tri ou d'une analyse au cytomètre !

Je souhaite beaucoup de courage à **Lu** pour la réalisation de sa thèse. J'ai été heureuse de partager mon bureau avec toi ces quelques mois ! J'étais heureuse de te retrouver pour nos réunions hebdomadaires. Heureuse d'avoir fait ta connaissance.

Roland, Jérôme, Alexandra, Hélène, Claire et Lu, je suis fascinée de voir à quel point les nouvelles vont vites !

Je remercie également l'équipe de Chimiste de l'IBL pour les bons moments passés autour des pauses déjeuners.

Je remercie également ma nouvelle équipe, l'unité INSERM U908, que j'ai intégrée il y a un peu plus d'un an. Merci de votre accueil, de votre disponibilité permanente, de votre bonne humeur. Un grand merci à Madame le Pr Xuefen LE BOURHIS de m'avoir accueillie pour quelques mois, d'avoir veillé à ma bonne installation ainsi qu'à la réussite de cette thèse.

Merci à **Sylvain**, mon colocataire de bureau, j'ai appris beaucoup de choses en peu de temps ! Un grand merci à **Chann, Pierre-Olivier, Robert-Alain, Valérie Chopin, Pamela et Nathalie** pour leur enthousiasme, leurs conseils et d'avoir contribué à la réalisation de cette fin de thèse dans de bonnes conditions. Je remercie également **Valérie Cendré et Isabelle** pour leur aide au quotidien.

Je remercie également mes nouveaux copains de thèses : **Mathieu** (non non, je ne t'ai pas oublié !), **Barbara, Justine, Romain** et sans oublier **Jérémy** ! Merci de m'avoir accueillie, vos sourires, votre bonne humeur, votre enthousiasme et votre disponibilité ont facilité mon adaptation ! Merci d'avoir couru à mon secours dès que je cherchais désespérément quelque chose ! Barbara, Justine et Romain, je vous souhaite bon courage pour la thèse. Mathieu, bon courage pour la reprise des études ! **Bernadette**, bon courage pour la thèse !

Enfin, je remercie tous les stagiaires que j'ai pu encadrer et qui ont contribué au bon déroulement de cette thèse : **Florian, Andréa, Amélie et Mylène**.

Je souhaite bon courage à mon successeur qui entre en M2R : **Jordan** et qui, je l'espère, réalisera une thèse sur ce sujet.

Un grand merci à mes amis, **Aurélie, Jérôme, Lucie, Geoffrey, Coralie et Laurent** de m'avoir soutenue dès le début. Merci d'avoir cru en moi et de m'avoir réconfortée dans toutes les situations ! Aurélie, je pense encore souvent à nos fous rires que l'on avait lorsque nous étions sur les bancs de la fac, juste avec un regard ! Merci d'être vous et d'être là dans toutes les situations ! Et, pour répondre à votre question : « *Ca y est, je m'arrête là ! Oui j'ai fini ! Et oui, 3 ans, cela passe très vite !* ».

Merci au Dr **Thierry DUGIMONT**, sans vous, je ne serai pas là aujourd'hui. Lors de vos enseignements, vous m'avez donné goût à l'étude des ARN. Merci d'avoir pris de mes nouvelles pendant ces trois années.

Enfin, je ne peux terminer sans remercier ma **famille** et surtout mes **parents ! MAMAN, PAPA**, merci de m'avoir fait confiance, de m'avoir soutenue à tout moment malgré les épreuves ! Et nous savons qu'il y en a eu. Je ne vous le dirais jamais assez, MERCI !

RESUME

Le locus *H19/IGF2* est soumis à l’empreinte génomique. A partir de l’allèle maternel, deux ARN non codants sont exprimés : *H19* et *91H*. Ces ARN sont fréquemment surexprimés dans les cancers du sein mais leurs fonctions et/ou modes d’action lors de la tumorigenèse ne sont pas encore définis.

L’ARN *H19* est le précurseur de microARN (miR-675) dont aucune cible n’a été décrite dans la glande mammaire. Dans un premier temps, nous avons donc recherché et validé plusieurs cibles de ces microARN permettant d’expliquer les propriétés oncogéniques d’*H19*. En effet, le miR-675 régule l’expression des gènes *c-Cbl*, *Cbl-b* et *FADD* dans les cellules cancéreuses mammaires. Ces nouvelles régulations permettent l’augmentation de la prolifération cellulaire, la migration/invasion mais aussi la résistance aux apoptogènes. Dans cette étude, de manière inattendue et surprenante, nous avons également découvert un nouveau mécanisme de recrutement des microARN. Nous montrons que les protéines associées aux ARNm peuvent favoriser le recrutement du complexe RISC, mais surtout moduler l’action de celui-ci.

De plus, dans certains cancers, l’ARN *H19* peut s’associer aux protéines, notamment la protéine P53 dans les cancers gastriques, pour moduler leur fonction. Dans le cancer du sein, cette interaction empêche l’action de la protéine P53 fonctionnelle et favorise sa dégradation. Un nouveau mécanisme de régulation de P53 a donc été identifié. Ce mécanisme peut être responsable de certaines résistances tumorales aux thérapies.

Par ailleurs, nous avons mis en évidence le rôle du gène *H19* dans la formation et le maintien des cellules souches. Cette nouvelle fonction peut permettre à l’ARN *H19* de promouvoir l’initiation ou la récurrence tumorale. Néanmoins, le mécanisme associé à l’ARN *H19* dans cette population n’est pas encore établi.

Pour finir, les fonctions du long ARN non codant *91H* dans les cellules cancéreuses mammaires ont été étudiées. Cet ARN possède des propriétés oncogéniques et permet l’expression des gènes *H19* et *IGF2* en modulant la conformation de la chromatine au locus.

Pour conclure, j’ai identifié plusieurs modes d’action de l’ARN *H19* lui permettant de promouvoir l’apparition de la tumeur, sa progression ou encore ses résistances aux thérapies. J’ai également défini le rôle de l’ARN *91H* au locus *H19/IGF2*, ainsi que sa fonction lors de la tumorigenèse.

Mots clefs : *H19*, *91H*, long ARN non codant, microARN, cancer du sein, cellules souches.

ABSTRACT

The *H19/IGF2* locus is submitted to the genomic imprinting. From the maternal allele, two long non-coding RNA are transcribed: *H19* and *91H*. These RNA are overexpressed in breast cancer. However their functions and mechanism, in tumorigenesis, remains undeciphered.

The *H19* RNA is a precursor of microRNA (miR-675). Few targets have been identified but neither in breast cancer cells lines. Firstly, I have identified and checked three targets of miR-675 involved in oncogenic properties of *H19* RNA. Indeed, the miR-675 regulate expression of *c-Cbl*, *Cbl-b* and *FADD* mRNA. These new regulation promotes cell proliferation, cell migration/invasion and resistance of cell apoptosis. Moreover, surprisingly, in these studies, I have highlight new mechanism of microRNA recruitment and function. Indeed, I have established that proteins associated to mRNA regulate microRNA recruitment and function.

In cancers, the *H19* could be associated with protein, for example the P53 protein in gastric cancer, in order to regulate their function. In breast cancer cells, the *H19/P53* association prevents the P53 function and promotes its degradation. We have shown a new mechanism of P53 regulation. This mechanism could be involved in drug resistance in cancer.

Otherwise, I have shown that *H19* is involved in stem cell formation and maintenance. This new *H19* function could be involved in tumor initiation or in tumor recurrence. However, the *H19* mechanism in this cell population remains unclear.

To finish, I have also determined the function of the *H19* antisense long non-coding RNA, *91H*. In breast cancer cells, I have demonstrated that *91H* acts as an oncogene and promotes *H19* and *IGF2* expression by modulating chromatin conformation at the locus.

In conclusion, I have identified several *H19* mechanism involved in tumor formation, progression or resistance to treatment. I have also decipher *91H* function in tumorigenesis and at the *H19/IGF2* locus.

Key words: *H19*, *91H*, long non-coding RNA, microRNA, breast cancer, stem cells.

SOMMAIRE

INTRODUCTION	1
La glande mammaire	3
I. La glande mammaire.....	3
1. Anatomie et histologie	3
2. Développement de la glande mammaire.....	5
II. Le cancer du sein	8
1. Epidémiologie.....	8
a. Quelques chiffres	8
b. Facteurs de risque.....	9
2. Le développement du cancer	11
3. La tumorigenèse mammaire	12
a. Altérations génétiques	12
b. Origine du cancer du sein.....	15
c. Influence du microenvironnement tumoral	19
4. Classification des cancers du sein	22
a. La classification histologique	22
b. La classification TNM	24
c. La classification moléculaire	25
III. Les thérapies dans le cancer du sein	27
1. Les tumeurs hormono-dépendantes.....	27
2. Les tumeurs HER2+	28
3. Les tumeurs triples négatives	29
4. Les inhibiteurs non spécifiques	30
Les longs ARN non codants	33
I. Les longs ARN non codants : classification	33
1. Découverte des longs ARN non codants.....	33
2. Classification des longs ARN non codants	35
II. Modes d'action des longs ARN non codants	37
1. Modes d'action nucléaire.....	37
a. Formation de sous-domaines nucléaires	37
b. Stabilité chromosomique	39
c. Régulation de l'expression des gènes	41
d. Formation de triplex ARN-ADN•ADN	44
2. Modes d'action cytosolique.....	44
a. Précurseur de petits ARN.....	44

b. Eponge à ARN	45
c. Interaction lncRNA/protéine	48
III. Les ARN non codants dans la glande mammaire et sa tumorigenèse	49
1. ARN non codant et développement de la glande mammaire	49
2. ARN non codant et cancer du sein	51
a. Les ARN non codants suppresseurs de tumeur	51
b. Les ARN non codants oncogènes	51
Les microARN	55
I. Découverte des microARN	55
1. Découverte des microARN	55
2. Nomenclature	55
II. Les voies de synthèse des microARN	57
1. La voie canonique	57
a. Formation du pré-miRNA.....	57
b. Export du pré-miRNA dans le cytosol	61
c. Maturation du pré-miRNA dans le cytosol.....	61
d. Maturation du duplex de microARN	62
2. Les voies non canoniques.....	65
a. La voie non canonique Drosha-indépendante	65
b. La voie non canonique Dicer-indépendante.....	65
III. Rôles des microARN	67
1. Mode d'action et régulation des microARN.....	67
a. Régulation négative par les microARN	67
b. Régulation positive par les microARN	68
c. Autres modes de régulation.....	69
2. Les microARN dans le cancer	69
a. Suppresseur de tumeur	69
b. Oncogène.....	70
c. Oncogène ou suppresseur de tumeur ?.....	70
IV. Les microARN : une cible thérapeutique ?	71
Le locus <i>H19/IGF2</i>.....	73
I. L'empreinte génomique parentale	73
1. Historique de l'empreinte génomique.....	73
2. Mise en place et maintien de l'empreinte génomique.....	75
3. Quelques exemples de locus soumis à l'empreinte génomique.....	79
II. Le locus <i>H19/IGF2</i>	81
1. L'empreinte génomique au locus <i>H19/IGF2</i>	81

2. Organisation tridimensionnelle de la chromatine	83
3. Modifications épigénétiques au locus <i>H19/IGF2</i>	86
4. Pathologies associées au locus <i>H19/IGF2</i>	87
5. Autres transcrits au locus <i>H19/IGF2</i>	87
III. Fonctions cellulaires associées au locus <i>H19/IGF2</i>	90
1. Le gène <i>IGF2</i>	90
2. Le gène <i>H19</i>	92
a. Historique	92
b. Le gène <i>H19</i> dans l’empreinte génomique	93
c. Régulation de l’expression d’ <i>H19</i>	94
d. <i>H19</i> dans la tumorigenèse	97
e. Mode d’action de l’ARN <i>H19</i>	100
OBJECTIFS	105
RESULTATS	109
I. Identification de cibles du miR-675 dans les cellules cancéreuses mammaires. 111	
1. Les ARNm c-Cbl et Cbl-b : deux nouvelles cibles du miR-675	111
2. Identification d’une nouvelle cible (l’ARNm FADD) et d’un nouveau mode de régulation des microARN	151
II. Régulation de la protéine P53 par <i>H19</i>	165
III. <i>H19</i> et les cellules souches	177
IV. L’ARN <i>91H</i> régule le locus <i>H19/IGF2</i> et est impliqué dans la tumorigenèse	219
DISCUSSION-PERSPECTIVES	249
BIBLIOGRAPHIE	269
ANNEXE	317
Annexe 1	319
Annexe 2	349

LISTE DES FIGURES

Figure 1. Structure de la glande mammaire.	4
Figure 2. Evolution de la glande mammaire au rythme des grossesses de la souris.	6
Figure 3. Cancérisation de la glande mammaire.	10
Figure 4. Hypothèse de l'évolution tumorale dans le cancer du sein.	16
Figure 5. Mécanismes d'apparition des cancers selon Visvader.	18
Figure 6. Interaction de la cellule tumorale avec son microenvironnement.	20
Figure 7. Classification des longs ARN non codants (lncRNA).	34
Figure 8. Représentation des principaux modes d'action des longs ARN non codants.	36
Figure 9. Stabilité chromosomique.	38
Figure 10. Mécanisme d'inactivation du chromosome X.	40
Figure 11. Mécanisme d'action des ARN non codants <i>HOTAIR</i> et <i>HOTTIP</i> .	42
Figure 12. Modulation de l'effet des microARN par les longs ARN non codants.	46
Figure 13. Implication des ARN non codants dans la glande mammaire normale et cancéreuse.	50
Figure 14. Synthèse des microARN.	56
Figure 15. Schéma représentant le pri-miRNA.	58
Figure 16. Schéma représentant le pré-miRNA.	60
Figure 17. Autres voies de synthèse des microARN: voie non canonique.	64
Figure 18. Mode d'action des microARN.	66
Figure 19. Mise en évidence de l'importance des deux génomes : maternel et paternel.	74
Figure 20. Exemples de locus soumis à l'empreinte génomique parentale.	78
Figure 21. Le locus <i>H19/IGF2</i> .	82

Figure 22. Conformation tridimensionnelle de la chromatine au locus <i>H19/IGF2</i> chez l'Homme.	84
Figure 23. Nouveaux transcrits au locus <i>H19/IGF2</i> chez l'Homme.	88
Figure 24. Les différents modes d'action de l'ARN non codant <i>H19</i> .	102
Figure 25. <i>H19/miR675</i> favorise l'expression de <i>FADD</i> dans les cellules cancéreuses mammaires.	159
Figure 26. Stabilité des ARNm <i>FADD</i> et <i>c-Myc</i> .	160
Figure 27. Le miR-675 régule directement <i>FADD</i> en se fixant sur l'UTR3'.	161
Figure 28. Détection par PCR des ARNm associés à la protéine AGO2 ou FXR1.	161
Figure 29. FXR1 recrute le miR-675 et permet son action positive.	162
Figure 30. <i>H19</i> permet la survie et la prolifération cellulaire.	163
Figure 31. <i>H19</i> fixe P53 et induit sa dégradation.	173
Figure 32. <i>H19</i> favorise la dégradation de P53.	174
Figure 33. <i>H19</i> empêche les fonctions de P53.	175
Figure 34. <i>H19</i> retient P53 dans le cytoplasme.	176
Figure 35. Nouveau mécanisme d'action des microARN.	254
Figure 36. L'ARN <i>H19</i> et le cycle cellulaire.	258
Figure 37. Rôle des ARN non codants <i>H19</i> et <i>91H</i> dans une cellule cancéreuse mammaire.	266

LISTE DES TABLES

Tableau 1. Evolution de l'incidence et de la mortalité du cancer chez la Femme en France depuis 1980.	8
Tableau 2. Quelques exemples de facteurs, autres que les altérations génétiques, modifiant le risque de développer un cancer du sein.	9
Tableau 3. Caractéristiques moléculaires des différents sous-types de cancer du sein.	26
Tableau 4. Tableau regroupant les différentes thérapies utilisées pour traiter le cancer du sein en fonction du sous-type tumoral.	31
Tableau 5. Liste non exhaustive des travaux montrant le rôle oncogène ou suppresseur de tumeur de l'ARN <i>H19</i> .	98
Tableau 6. Liste des cibles des miR-675-5p et miR-675-3p déjà identifiées dans les tissus.	101
Tableau 7. Séquence des amorces et siRNA utilisés dans l'étude.	164

INTRODUCTION

La glande mammaire

I. La glande mammaire

1. Anatomie et histologie

Le sein est une glande exocrine assurant la sécrétion du lait chez la Femme. Le sein est composé d'une glande mammaire entourée d'un tissu graisseux reposant sur les muscles pectoraux grâce à des crêtes fibro-glandulaires (crêtes de Duret), des ligaments suspenseurs et des ligaments appelés ligaments de Cooper (fig 1). La peau entourant le mamelon est mince et pigmentée : c'est l'aréole. Elle contient des glandes sébacées, appelées glandes de Morgani, qui s'hypertrophient à la grossesse et prennent alors le nom de tubercules de Montgomery. La sécrétion de ces glandes participe probablement à la protection du mamelon et de l'aréole lors de l'allaitement. La structure de la glande mammaire est en perpétuelle évolution chez la Femme. En effet, ses aspects morphologiques et fonctionnels varient au cours des différentes périodes de la vie génitale de la Femme.

La glande mammaire est une glande exocrine de type tubulo-alvéolaire, constituée de 15 à 25 lobes irréguliers. Les lobes sont séparés par une couche de tissu conjonctif dense et sont enfouis dans du tissu graisseux. Chaque lobe est drainé par un canal unique, le canal galactophore, qui possède un orifice à la surface du mamelon (fig 1). Le canal galactophore se ramifie progressivement depuis le revêtement cutané jusqu'en profondeur du sein pour se terminer en canaux terminaux (canaux lobulaires), aboutissant chacun à un lobule. Les lobules sont séparés entre-eux par du tissu conjonctif inter-lobulaire moyennement dense et sont composés de multiples alvéoles ou acini reliés au canal lobulaire grâce à un canalicule.

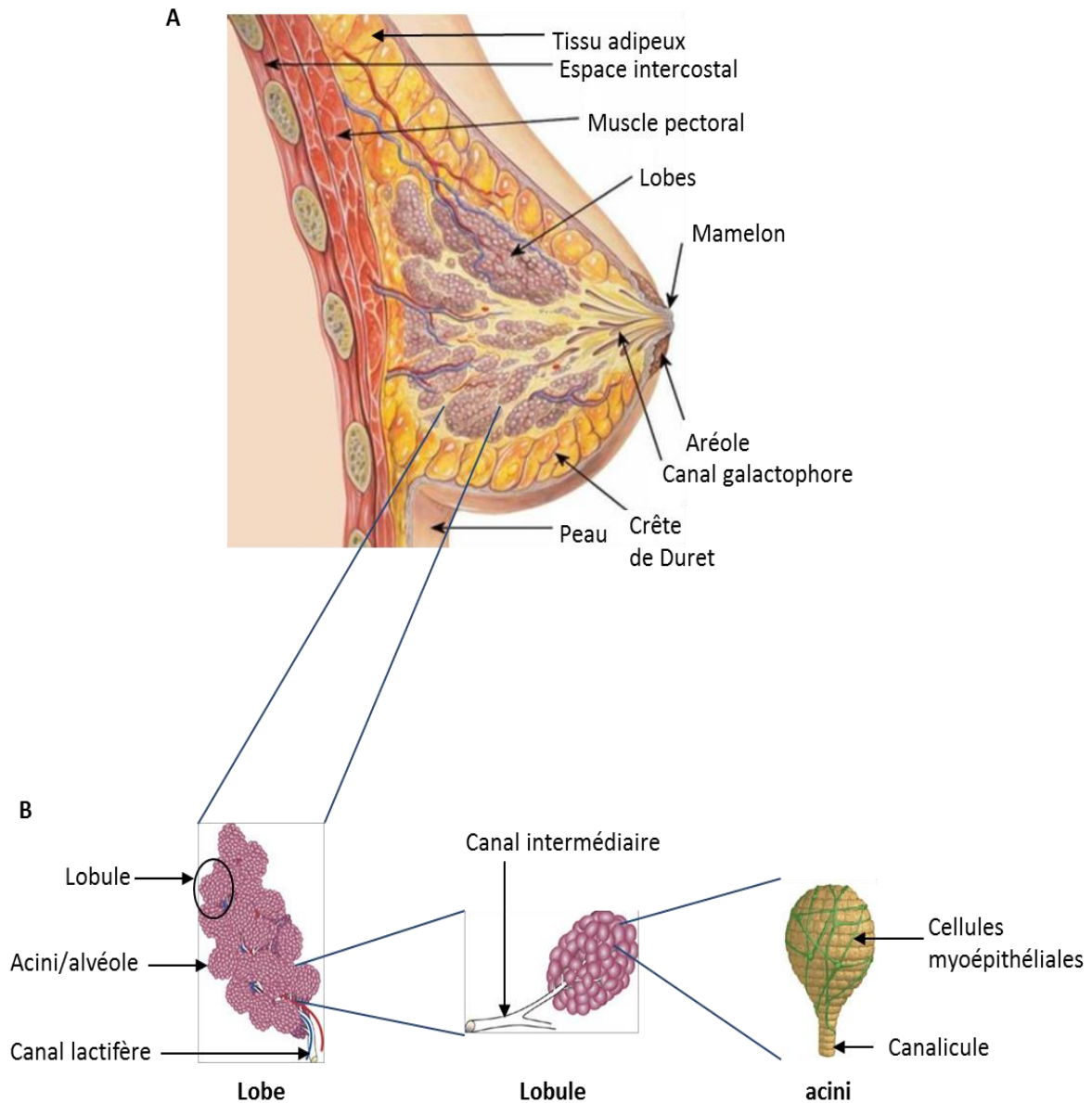


Figure 1. Structure de la glande mammaire. **A.** Représentation schématique d'une coupe sagittale de la glande mammaire chez la Femme. La glande mammaire est principalement constituée de tissu adipeux et de 15 à 20 lobes irréguliers. **B.** Les lobes constituant la glande mammaire sont reliés entre-eux par le canal lactifère qui se termine en canal galactophore. Les lobes se composent de lobules eux-mêmes constitués d'acini ou alvéoles entourés de cellules myoépithéliales. Lors de la succion, pendant la lactation, les cellules myoépithéliales se contractent. Le lait, sécrété par les acini, est éjecté dans le canalicule puis dans les canaux intermédiaires et lactifères pour finir dans le canal galactophore. Ce dernier possède un orifice au niveau du mamelon ; ce qui permet de nourrir le nourrisson. (Adapté de Patrick J Lynch et santeallaitementmaternel.com)

Chaque canal terminal relié à son lobule forme une unité terminale ducto-lobulaire (UTDL), très sensible aux variations hormonales et constituant l'unité fonctionnelle de la glande mammaire. Les UTDL assurent la production de lait lors de la grossesse et de la lactation. A ce jour, il est admis que la plupart des lésions mammaires mastosiques et carcinomateuses se développent à partir de l'UTDL.

La glande mammaire est constituée de plusieurs types cellulaires distincts. Bien que le système canalaire représente la structure fonctionnelle de la glande mammaire, le système fibro-adipeux entourant les canaux constitue la majorité du sein. Le système canalaire est constitué de deux types cellulaires : une couche interne, composée de cellules épithéliales sécrétrices de lait et une couche externe, composée de cellules myoépithéliales fusiformes contractiles. En réponse au stimulus de succion lors de l'allaitement, les cellules myoépithéliales se contractent, ce qui induit l'éjection du lait et son acheminement à travers les canaux alvéolaires, terminaux et galactophores. Ces deux couches cellulaires sont délimitées par une membrane basale et baignent dans un tissu conjonctif lâche richement vascularisé : le stroma ou tissu palléal. Ce stroma est constitué d'une matrice extracellulaire riche en fibres de collagène, en fibroblastes, en adipocytes et en vaisseaux sanguins et lymphatiques nécessaires au développement de la glande mammaire.

2. Développement de la glande mammaire

La glande mammaire évolue constamment au cours de la vie de la Femme, en fonction des cycles menstruels et des grossesses. Bien que le développement de la glande mammaire débute pendant la vie fœtale, l'essentiel de la croissance se produit après la puberté et ne se termine qu'une fois la première grossesse menée à terme.

Au cours de la 4^{ème} semaine du développement embryonnaire, une ébauche du sein apparaît sous forme de crête mammaire (ou ligne lactéale) correspondant à un épaississement de l'ectoderme qui s'étend de la future aisselle jusqu'à la future aine. Le long de cette crête, des épaississements symétriques, nommés bourgeons mammaires primitifs, apparaissent. A la 6^{ème} semaine, la crête mammaire disparaît, seuls les bourgeons mammaires subsistent ; c'est la fin de la phase embryonnaire. A partir du 5^{ème} mois, les bourgeons mammaires émettent 15 à 25 prolongements cylindriques dans le mésoderme qui vont se ramifier et se dilater à leur extrémité. Ensuite, la lumière formant les futurs canaux galactophores va apparaître. A la naissance, la structure de la glande mammaire est donc rudimentaire et identique entre les hommes et les femmes (fig 2). Chez la femme, la croissance de la glande continuera au moment de la puberté et des futures grossesses ; tandis que chez l'homme, sa croissance est modeste et elle restera à ce stade toute la vie.

Introduction

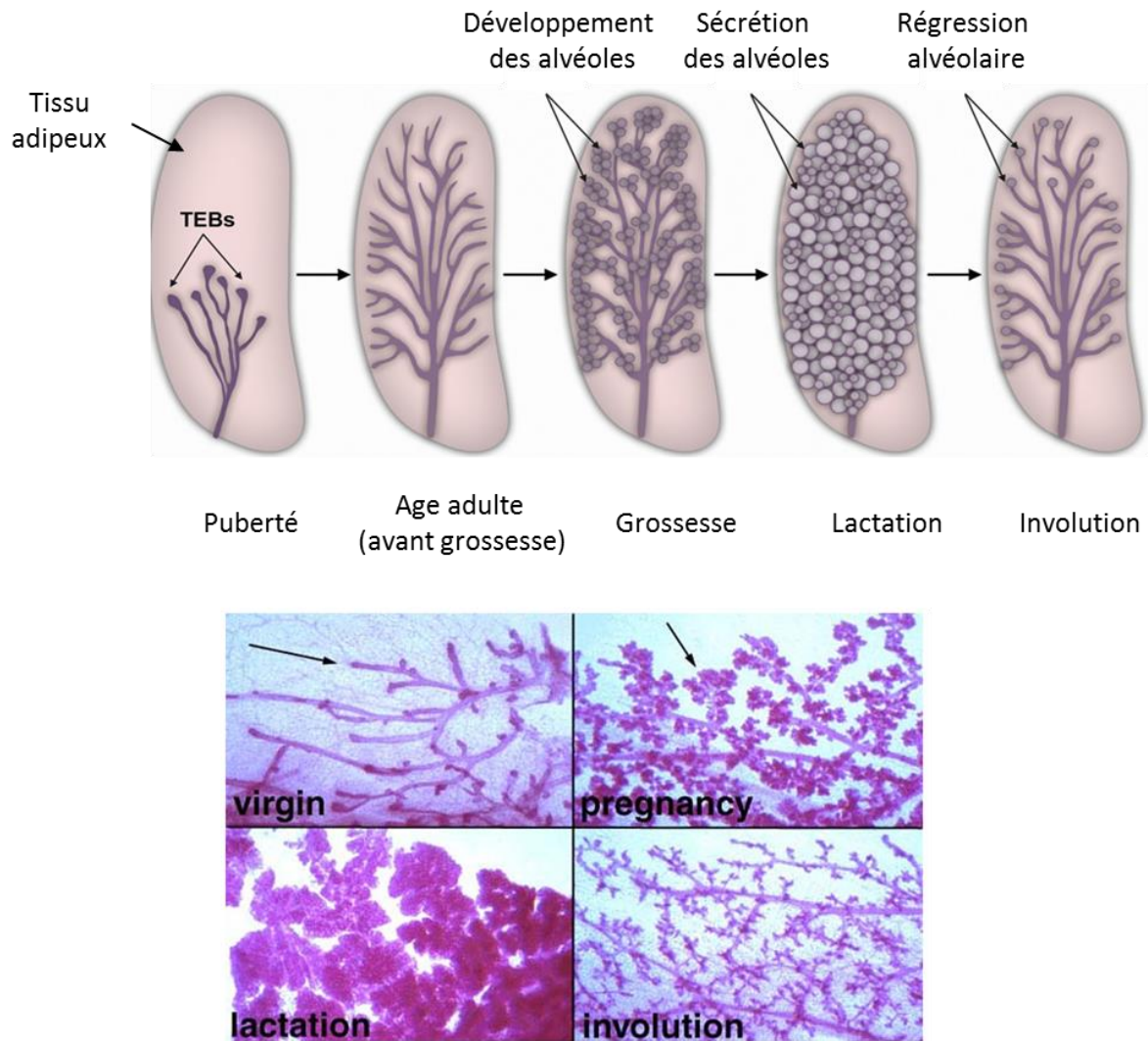


Figure 2. Evolution de la glande mammaire au rythme des grossesses de la souris. A la puberté, le tissu adipeux s'accumule et la croissance des canaux est observée. Pendant la grossesse, la croissance des bourgeons terminaux (TEBs) ou bourgeons alvéolaires se produit, pour prendre la place du tissu adipeux. Au cours de la lactation, les alvéoles se différencient permettant ainsi la production de lait. Après la période de sevrage, l'involution permet le retour de la glande mammaire à une structure similaire à celle observée avant la grossesse. (adapté de Shore & Rosen, 2014; mammary.nih.gov)

Pendant l'enfance, le développement de la glande mammaire est symétrique et lent : le futur coussin adipeux mammaire se forme. Dans ce dernier, quelques canaux épithéliaux apparaissent, ils sont coiffés à leur extrémité par des bourgeons terminaux. A la puberté, sous l'influence de l'hormone de croissance et des hormones sexuelles (œstrogènes, progestérone), la croissance de la glande mammaire s'accélère, notamment par une expansion rapide du réseau canalaire et du stroma périphérique. A ce stade, l'augmentation du volume du sein est principalement due aux dépôts lipidiques. Une pigmentation ainsi qu'une saillie de l'aréole et du mamelon se produisent également à la puberté. Ce n'est qu'au cours de la gestation que les alvéoles se développent activement pour prendre la place du tissu adipeux (fig 2). En effet, lors de la grossesse, les concentrations plasmatiques en œstrogènes, progestérone et hormones lactogènes placentaires sont élevées. La progestérone et les œstrogènes favorisent la sécrétion de facteurs de croissance comme le TGF α (*Transforming Growth Factor*), le MDGF-1 (*Mammary Derived Growth Factor-1*), l'IGF (*Insulin Growth Factor*) et l'EGF (*Epidermal Growth Factor*) par les cellules épithéliales. Ces facteurs vont, à leur tour, réguler de façon paracrine la prolifération et la différenciation de l'épithélium mammaire. Cette phase de prolifération et différenciation se produit principalement au cours des vingt premières semaines de la grossesse. Au cours du dernier trimestre, une accumulation de granules de sécrétion dans les cellules épithéliales mammaires est observée. A la fin de la grossesse, la glande mammaire est entièrement différenciée, les alvéoles occupent tout l'espace du coussin adipeux et elles deviennent fonctionnelles (fig 2).

Pendant la grossesse, la sécrétion du lait est inhibée par la progestérone d'origine placentaire qui inhibe la prolactine. Après l'accouchement, le placenta étant éliminé, on observe une modification de l'équilibre progestérone/prolactine, ce qui se traduit par une production de lait. Lors des tétées, la succion du mamelon va permettre l'entretien de la sécrétion de prolactine par l'adénohypophyse, mais également la sécrétion d'ocytocine. Cette hormone est responsable de la contraction des cellules myoépithéliales nécessaire à l'éjection du lait.

Après une période de sevrage, la glande mammaire involue, elle retourne à un état post-pubertaire, quiescent jusqu'à une nouvelle grossesse. La sécrétion de lait s'arrête et les cellules épithéliales et myoépithéliales entrent en apoptose.

A la ménopause, suite à la chute des œstrogènes et de la progestérone, les acini disparaissent, les cellules épithéliales et myoépithéliales s'atrophient et la membrane basale s'épaissit. Les fibres élastiques et de collagène du tissu conjonctif sont également altérées. Le sein sera alors principalement constitué de tissu fibro-adipeux.

II. Le cancer du sein

1. Epidémiologie

a. Quelques chiffres

Le cancer du sein est le premier cancer chez la Femme dans le monde. En effet, selon l'Organisation Mondiale de la Santé (OMS), 1,67 millions de nouveaux cas de cancer du sein ont été diagnostiqués dans le monde en 2012. Avec 521 000 décès en 2012, il est également la première cause de mortalité par cancer chez la Femme dans le monde (OMS, 2015). Dans les pays en voie de développement, l'incidence du cancer du sein est moins élevée par rapport aux pays développés. Cependant, lorsqu'il est diagnostiqué dans ces pays, le cancer est à un stade plus avancé, il est donc associé à un taux de mortalité plus élevé.

En France, chez la femme, le cancer du sein est le premier cancer (48 000 nouveaux cas en 2012 selon l'Inca) et il est également la première cause de mortalité par cancer (11 900 décès en 2012 selon l'Inca). Entre les années 1980 et 2005, le taux d'incidence du cancer du sein a augmenté fortement. En effet, en 25 ans, l'incidence du cancer du sein a doublé en France : 21 387 nouveaux cas étaient diagnostiqués en 1980 et 49 087 en 2005 (Institut National de Veille Sanitaire, invs, rapport de 2013). Cette forte hausse peut être expliquée par la mise en place des mammographies de dépistage du cancer chez la Femme de plus de 50 ans. Au cours de cette période, le taux de mortalité a également augmenté : 8 596 décès en 1980 et 11 886 décès en 2005. Cependant, au cours des 10 dernières années, l'incidence du cancer a tendance à diminuer légèrement, puisque, en 2012, 48 000 nouveaux cas étaient diagnostiqués (tableau 1). Mais le taux de mortalité reste stable, ceci peut être expliqué par le diagnostic précoce de la maladie ainsi qu'une bonne prise en charge thérapeutique.

Le cancer du sein possède l'un des meilleurs taux de survie, puisque 5 ans après le diagnostic, le taux de survie est de 89%, tous sous-types confondus (données Inca). Ce taux de survie est variable en fonction de la localisation de la tumeur : 98,3% pour une tumeur de stade local, 83,5% pour le stade régional avec envahissement ganglionnaire et 23,3% pour le stade métastatique (Inca). Ainsi la tumeur primaire, locale, est facilement soignée ; les métastases, qui surviennent plus tard, sont les principales causes de décès.

année	1980	2005	2012
incidence	21 387	49 087	48 000
mortalité	8 596	11 886	11 900

Tableau 1. Evolution de l'incidence et de la mortalité du cancer chez la Femme en France depuis 1980. (Selon l'Inca)

b. Facteurs de risque

Le cancer du sein est une maladie multifactorielle qui touche aussi les Hommes. En effet, chaque année, environ 1% des cancers du sein sont déclarés chez l'Homme.

Des altérations génétiques ou épigénétiques peuvent induire ce cancer, notamment des mutations sur les gènes *BRCA1/2* (*Breast CAncer 1/2*) (gènes de prédisposition au cancer du sein). De plus, outre l'âge et le sexe des individus, les contraceptifs hormonaux, les grossesses, l'allaitement, les traitements associés à la ménopause ainsi que la densité mammaire modifient le risque de développer un cancer du sein. Par ailleurs, le mode de vie des individus : la consommation d'alcool, le tabagisme, les habitudes alimentaires (surpoids), les rayons ionisants ou encore l'exercice physique modifient également le risque de développer un cancer du sein (Nkondjock & Ghadirian, 2005 ; Dieterich et al., 2014) (tableau 2).

facteur de risque	diminution du risque	augmentation du risque	mode d'action du risque
obésité (à la ménopause) et prise de poids (à partir de 18 ans)		*	augmentation taux adiponectine, leptine et oestrogènes
tabagisme		*	carcinogène, augmentation de 10% du risque de cancer du sein
alcool		*	augmentation taux oestrogènes
thérapie hormonale à la ménopause		*	augmentation taux progestérone
rayons ionisants		*	dommages à l'ADN
contraceptifs hormonaux		*	augmentation taux oestrogènes
densité mammaire		*	augmentation d'IGF avant la ménopause augmentation prolactine après la ménopause
grossesse précoce et multiparité	*		différenciation rapide de la glande mammaire
allaitement	*		différenciation des alvéoles réduction de la production d'oestrogènes
régime alimentaire	*		consommation de fruits et légumes permet diminution de l'obésité
activité physique	*		maintien de l'équilibre énergétique réduction de la production d'oestrogènes

Tableau 2. Quelques exemples de facteurs, autres que les altérations génétiques, modifiant le risque de développer un cancer du sein.

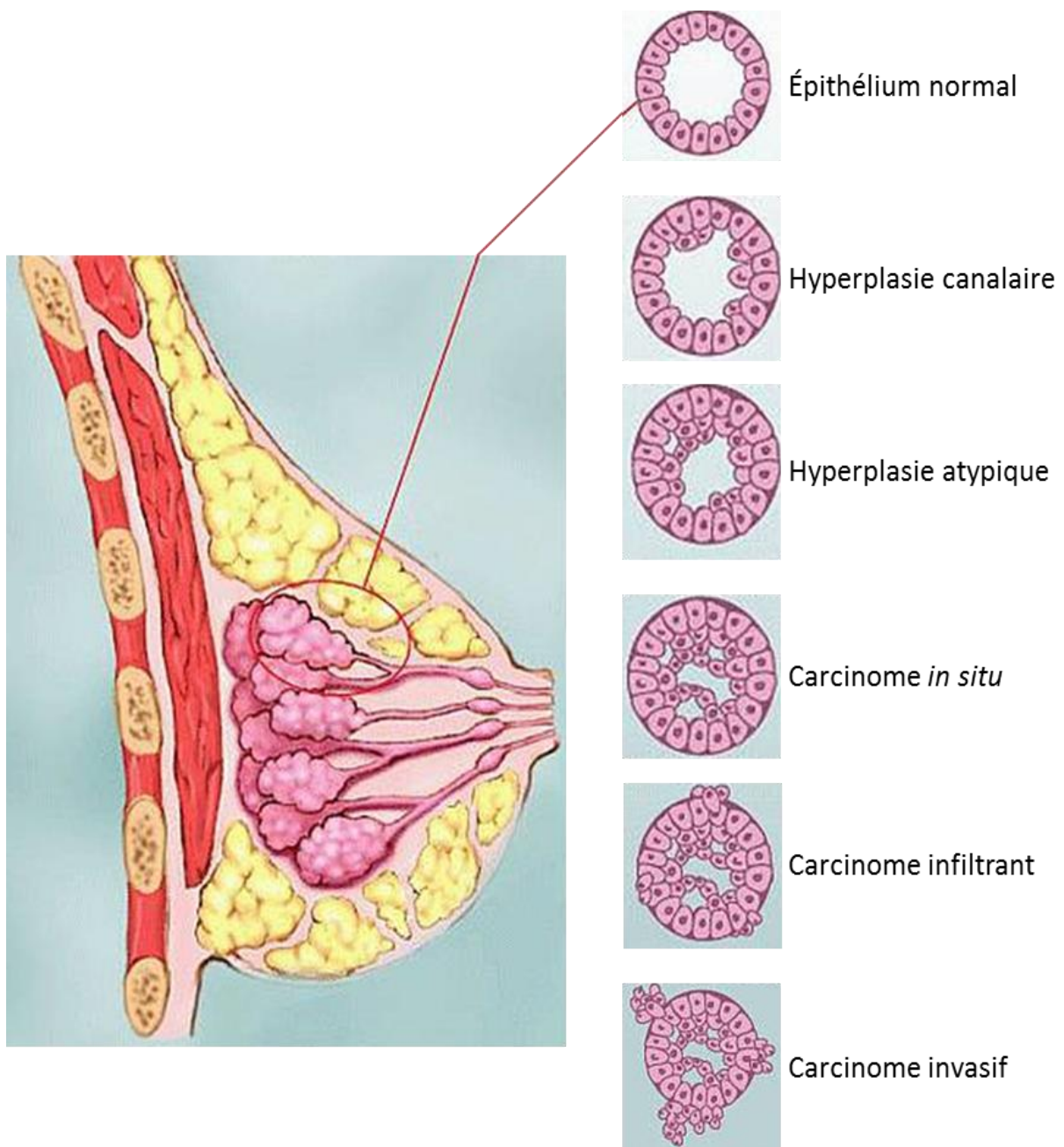


Figure 3. Cancérisation de la glande mammaire. Les cellules épithéliales des alvéoles ou des canaux prolifèrent lentement par mutation spontanée : c'est le stade de l'hyperplasie. Les cellules normales se transforment en cellules cancéreuses progressivement. De nouvelles modifications épigénétiques ou génétiques augmentent la prolifération de ces dernières : c'est l'hyperplasie atypique. Ensuite, lorsque l'amas cellulaire obstrue le canal : la transformation des cellules en cellules cancéreuses est terminée, le cancer de type carcinome *in situ* est déclaré. Les cellules tumorales vont ensuite franchir la lame basale : carcinome infiltrant et disséminer dans l'organisme : carcinome invasif. (adapté de www.centredesmaladiesdusein.ca)

2. Le développement du cancer

Plusieurs étapes clés sont nécessaires à la transformation d'une cellule normale en une cellule cancéreuse. Ces différentes étapes sont détaillées dans le paragraphe suivant (voir P. 12). Dans ce paragraphe, nous étudierons les différents stades tumoraux observés dans les cancers du sein.

Plusieurs stades peuvent être caractérisés et analysés (fig 3). Le premier stade est « l'hyperplasie » : elle correspond à une prolifération accrue des cellules épithéliales canalaire normales luminales ou basales ; elle entraîne ainsi un épaississement de l'épithélium canalaire et une obstruction partielle du canal galactophore. Ce stade hyperplasique n'augmente pas le risque de cancer du sein chez les femmes de moins de 55 ans.

Nous pouvons également distinguer un stade « hyperplasique dit atypique ». Cette hyperplasie est caractérisée par une prolifération augmentée des cellules luminales uniquement. Elle multiplie le risque de cancer du sein par un facteur 4 à 5.

Ensuite, les cellules acquièrent des anomalies génétiques et architecturales qui conduisent le plus souvent à la formation de lésions pré-cancéreuses ou dysplasie. Dans ces cellules, l'accumulation de modifications génétiques et épigénétiques permet la formation d'un « carcinome *in situ* » ; les cellules restent localisées au niveau de la lésion, la membrane basale est intacte et aucune invasion dans le tissu n'est observée.

Au sein de la tumeur ainsi formée, de nouvelles altérations génétiques peuvent modifier les capacités invasives et métastatiques des cellules tumorales. Lorsque les cellules de la tumeur primaire sont capables d'altérer l'intégrité de la lame basale et d'envahir le tissu conjonctif environnant de la tumeur primaire, le « carcinome est dit infiltrant ». Le franchissement de la lame basale est une étape décisive dans la progression tumorale ; elle signifie que les cellules tumorales peuvent envahir le stroma par perte d'adhérence, sécrétion de protéase, et peuvent mettre en place de l'angiogenèse. Enfin, la dernière étape reflétant l'agressivité tumorale est la capacité des cellules tumorales à former des métastases dans l'organisme entier : c'est « le carcinome invasif ». Cette étape est caractérisée par la circulation des cellules tumorales dans les réseaux lymphatiques et/ou sanguins pour coloniser différents organes notamment le cerveau pour les cellules cancéreuses mammaires et former une tumeur secondaire.

3. La tumorigenèse mammaire

La tumorigenèse mammaire est un processus multifactoriel faisant appel à une succession d'altérations génétiques permettant la transformation d'une cellule normale en une cellule cancéreuse. Selon Hanahan et Weinberg, six modifications physiologiques cellulaires, distinctes et complémentaires sont nécessaires pour permettre à la cellule transformée de prendre le dessus sur les cellules saines du tissu (Hanahan & Weinberg, 2000). Ces modifications sont :

- Indépendance vis-à-vis des signaux de croissance
- Perte de sensibilité aux signaux anti-prolifératifs
- Résistance à l'apoptose
- Potentiel réplicatif illimité
- Capacité à envahir les tissus et à métastaser
- Néo-angiogenèse

Plus récemment, deux autres caractéristiques pouvant intervenir dans le processus tumoral ont été proposées : la reprogrammation du métabolisme énergétique pour supporter la croissance accrue des cellules cancéreuses et l'échappement à la destruction par le système immunitaire (Hanahan & Weinberg, 2011). Ces modifications physiologiques sont directement liées à l'activation des oncogènes et/ou la suppression des gènes suppresseurs de tumeur qui sont des résultantes d'altérations génétiques et épigénétiques sporadiques, ou plus rarement héréditaires (5-10% des cancers du sein) (Leredours & Lidereau, 2002 ; Mavaddat et al., 2010 ; Apostolou & Fostira, 2013 ; Mahdi et al., 2013).

a. Altérations génétiques

Activation d'oncogène

On considère comme oncogène tous les gènes dont la régulation positive participe au processus tumoral. Dans le cancer du sein, cette activation peut être due à une amplification génique, des mutations ponctuelles ou encore des translocations chromosomiques (Osborne et al., 2004). La plupart des oncogènes est impliquée dans la transduction du signal de prolifération (FGF, PDGF, EGF, Erb-B2, Ras, Raf, MEK, Myc, Jun, ...), dans la régulation du cycle cellulaire (Cycline D1, Cdk4, Cdc25), dans la régulation du pouvoir métastatique (métalloprotéases, protéines d'adhésions cellulaires) ou encore dans l'angiogenèse (VEGF).

Dans le cancer du sein, les amplifications les plus fréquentes concernent les oncogènes CCND1 (cycline D1), c-myc et ERBB2. Néanmoins, l'expression de ces différents oncogènes peut également être régulée positivement par des facteurs de transcription par exemple.

- **CCND1**, ce gène codant pour la cycline D1 est localisé sur le chromosome 11 en position p13. Cette protéine est un régulateur majeur du cycle cellulaire ; elle intervient notamment dans les transitions G1/S et G2/M du cycle. Elle possède également des activités de facteurs de transcription. Dans 50% des carcinomes mammaires, la cycline D1 est surexprimée alors que son amplification génique n'est retrouvée que dans 15% des cas (Arnold & Papanikolaou, 2005). Cette différence s'explique par le fait que l'expression de la cycline D1 peut être régulée positivement par d'autres facteurs comme les œstrogènes ou encore la protéine P53 *via* son effet sur P21 (Roy & Thompson, 2006). De plus, une étude a mis en évidence que la surexpression de cette protéine est corrélée à une résistance au Tamoxifen lorsque celui-ci est administré dans des cas de cancers hormono-dépendants (Stendahl et al., 2004).
- **c-myc**, ce gène codant pour la protéine nucléaire Myc est localisé sur le chromosome 8 en position q24. Il s'agit d'un facteur de transcription régulant l'expression de nombreux gènes humains impliqués dans divers processus cellulaires : prolifération, croissance, différenciation et apoptose (Chen & Olopade, 2008). Comme pour la cycline D1, Myc est surexprimé dans 40% des tumeurs mammaires alors que son amplification génique n'est retrouvée que dans 15% des cas. Cette différence peut être expliquée par une augmentation de la transcription de l'ARNm et/ou une stabilisation de l'ARNm et de la protéine dans la cellule. Par ailleurs, plusieurs études ont mis en évidence une corrélation entre l'expression de Myc, les tumeurs de haut grade et un mauvais pronostic (Xu et al., 2010b).
- **ERBB2**, ce gène codant pour la protéine HER2 ou Erb-B2 est localisé sur le chromosome 17 en position q21-22. Cette protéine HER2 est un récepteur à activité tyrosine kinase orphelin. Il fait parti d'une famille de récepteurs regroupant 4 membres : HER1 ou EGFR, HER2, HER3 et HER4. Le récepteur HER2 s'hétérodimérise avec les récepteurs HER3 ou HER4 pour induire sa signalisation. Dans 25% des cancers du sein, Erb-B2 est surexprimé et son gène amplifié ; ce qui conduit à son homodimérisation et/ou son hétérodimérisation. Par l'intermédiaire de ses voies de signalisation pro-tumorale, Erb-B2 est impliqué dans divers processus cellulaires : prolifération, angiogenèse, interaction cellulaire, induction de métastase, résistance à l'apoptose (Moasser, 2007). Les tumeurs dites HER+ sont aujourd'hui un sous-type de cancer du sein, elles sont associées à un phénotype invasif et à un mauvais pronostic. En effet, ces tumeurs sont résistantes à l'hormono-thérapie et à certaines chimiothérapies. Afin de pallier ces résistances, les tumeurs HER2+ font l'objet de thérapies ciblées utilisant des anticorps monoclonaux anti-HER2 ou anti-HER2/HER3. Ces différentes thérapies sont détaillées dans le paragraphe P29.

Inhibition de gènes suppresseurs de tumeur

On considère comme gène suppresseur de tumeur tous les gènes dont la régulation négative favorise le développement tumoral. Certains de ces gènes ont un rôle dans la régulation de la prolifération cellulaire, ce sont les « *gatekeepers* ». Nous pouvons citer Rb qui régule négativement la progression du cycle cellulaire ou encore PTEN qui est un inhibiteur de la voie PI3K. D'autres gènes sont impliqués dans la réparation de l'ADN lors de cassures simples ou doubles brins ; ils sont ainsi impliqués dans la stabilité génétique : ce sont les « *caretakers* ». Nous pouvons citer P53 et BRCA qui, lors de dommages à l'ADN, permettent un arrêt de cellules en cycle et la mise en place de complexes de réparation de l'ADN. Selon le modèle de Knudson, leur participation à la tumorigenèse nécessite une altération des deux allèles du gène en question (Knudson, 1971). Lors de cancer héréditaires, pour BRCA1 par exemple, l'un des deux allèles est muté sur la lignée germinale et l'autre allèle est muté de façon sporadique dans l'ADN tumoral (Knudson, 2001). Ces gènes constituent donc les gènes de prédisposition.

Dans le cancer du sein, les gènes suppresseurs de tumeur les plus fréquents sont : BRCA1/2, TP53 et Rb.

- **BRCA1/2** sont des gènes de prédisposition aux cancers du sein, de l'ovaire, du pancréas et de la prostate. Ces gènes sont localisés respectivement sur les chromosomes 17q22 et 13q12-13. Les protéines BRCA, grâce à leur interaction avec les protéines Rad51, P53 et l'ADN polymérase, sont impliquées dans la réparation de l'ADN, la régulation du cycle cellulaire et la régulation de la transcription (Venkitaraman, 2002 ; Foulkes & Shuen, 2013). Dans les cancers familiaux, les gènes BRCA1/2 sont mutés dans 50% des cas. Les femmes porteuses d'une mutation BRCA1 ont un risque de développer un cancer du sein avant 70 ans de 65%. Ce risque est de 39% pour les femmes porteuses d'une mutation BRCA2 (données Inca). Les mutations BRCA sont rarement des mutations ponctuelles ; il s'agit généralement de délétion ou de duplication de grand domaine (de 14 à 17 kb). Elles sont généralement retrouvées dans des tumeurs de haut grade et de mauvais pronostic. Par ailleurs, les mutations BRCA2 sont plus spécifiques des cancers du sein chez l'homme puisqu'elles y sont retrouvées dans 10% des cas (Thompson & Easton, 2001 ; Perou et al., 2000).
- **TP53** est localisé sur le chromosome 17p13.1. Il code pour une protéine impliquée dans la réparation de l'ADN. Lors de dommages à l'ADN, P53, en activant P21, induit un arrêt temporaire du cycle cellulaire durant lequel a lieu la réparation. Si les dommages sont trop importants, P53, en activant la protéine Bax, peut orienter la cellule vers un processus d'apoptose, éliminant ainsi les cellules potentiellement

malignes de l'organisme (Lacroix et al., 2006). Ainsi, P53 est un gardien du génome. Dans 25% des cancers du sein, P53 est muté et est associé à un mauvais pronostic (Yang et al., 2013).

- **Rb** est localisé sur le chromosome 13q14.1. Cette protéine intervient dans le cycle cellulaire notamment dans la régulation négative de la transition G1/S du cycle cellulaire. Lorsque Rb est phosphorylé, il régule négativement le facteur de transcription E2F1 et les CDK (Di Fiore et al., 2013). Ainsi, une perte de p-Rb provoque une perte des checkpoints du cycle cellulaire et donc une prolifération anarchique des cellules ; ce qui contribue à l'agressivité de la tumeur. Dans 25 à 30% des cancers du sein, une perte d'hétérozygotie du gène Rb est retrouvée (Bosco & Knudsen, 2007).

b. Origine du cancer du sein

Les tumeurs, y compris les tumeurs mammaires, sont très hétérogènes en termes de différenciation cellulaire, transformation et altérations génétiques. Cette hétérogénéité est, à l'heure actuelle, expliquée par deux modèles de progression tumorale : le modèle stochastique et le modèle hiérarchique.

Dans le **modèle stochastique**, également nommé modèle de l'évolution clonale, toutes les cellules du tissu, quelque soit leur degré de différenciation, acquièrent une série de mutations sporadiques selon Hanahan et Weinberg (fig 4). Ces altérations génétiques permettent alors à un clone cellulaire un avantage prolifératif ou de survie par rapport aux autres cellules différenciées du tissu. La prolifération aberrante de ce clone permettra ainsi la formation de la masse tumorale. Cependant, aujourd'hui, cette hypothèse semble obsolète puisque, selon cette théorie, peu ou pas de hiérarchie cellulaire serait observée au sein de la tumeur. De plus, il est difficile de comprendre comment une cellule différenciée, qui se divise donc très peu, peut acquérir une série de mutations. Enfin, si un clone cellulaire, par l'intermédiaire d'un set de mutations, est capable de former une tumeur, comment expliquer qu'il faille injecter un grand nombre de cellules en souris immunodéficientes pour obtenir une tumeur par xénogreffe.

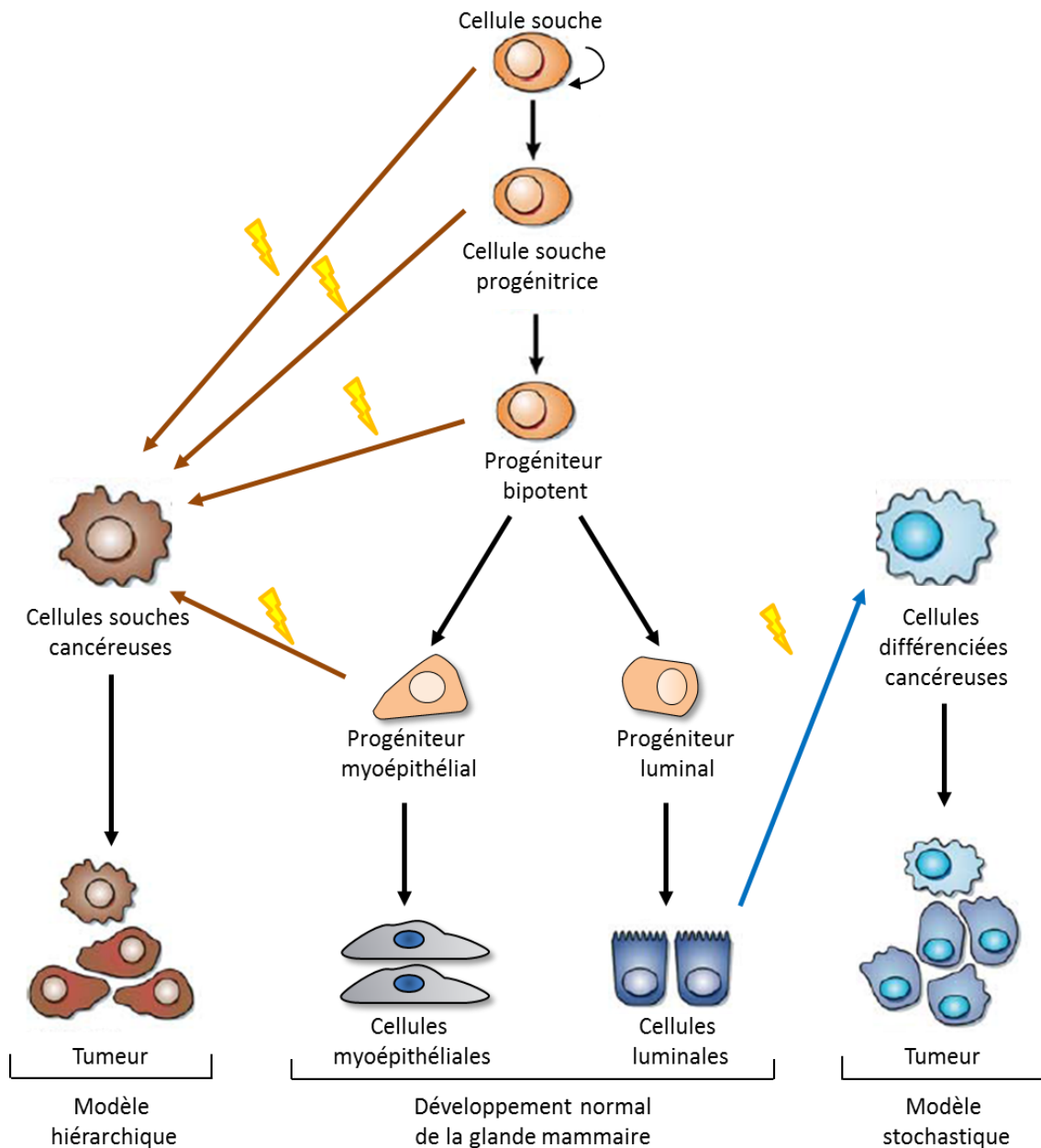


Figure 4. Hypothèse de l'évolution tumorale dans le cancer du sein. En condition normale, les cellules souches de la glande mammaire donnent naissance aux cellules progénitrices myoépithéliales et lumorales. Ces dernières forment respectivement les cellules myoépithéliales et lumorales de la glande mammaire. **Le modèle stochastique** de l'évolution tumorale : les cellules différenciées subissent des altérations géniques ou épigénétiques, elles se transforment alors en cellules cancéreuses, ce qui aboutira à la formation d'une tumeur. L'hétérogénéité tumorale est définie par le(s) type(s) cellulaire(s) à l'origine de la tumeur. **Le modèle hiérarchique** de l'évolution tumorale : les cellules souches et les cellules progénitrices subissent des altérations géniques, elles deviennent cancéreuses et permettent la formation de tumeur. L'hétérogénéité tumorale est apportée par les capacités multipotentes des cellules souches. (Adapté de Shipitisin & Polyak, 2008).

Au cours des dernières années, plusieurs données expérimentales ont mis en évidence l'existence d'une sous-population au sein de la tumeur. En 1997, l'équipe de Dick a mis en évidence que, dans les leucémies, cette sous-population est capable de former une tumeur identique à la tumeur d'origine après xénogreffe en souris immunodéficientes (Bonnet & Dick, 1997). Ils introduisent alors la notion de cellules souches cancéreuses (CSC). Les cellules souches possèdent une capacité d'auto-renouvellement illimité et sont multipotentes. Elles donnent naissance à des cellules progénitrices précoces puis tardives. Ces cellules progénitrices sont ensuite capables de donner naissance à l'ensemble des types cellulaires du tissu considéré (fig 4). Les CSC ont ensuite été isolées dans de nombreux cancers, notamment dans le cancer du sein en 2003 (Al-Hajj et al., 2003). L'isolement et la caractérisation des CSC font intervenir des combinaisons de marqueur variable en fonction du cancer étudié. Par exemple, dans le cancer du sein, les CSC sont caractérisées par une forte expression du marqueur de surface CD44 et une faible expression, voire une absence, du marqueur CD24 (phénotype CD44⁺/CD24^{-low}). L'existence de ces cellules a permis l'élaboration d'un nouveau modèle tumoral : **le modèle hiérarchique** (fig 4). Selon cette théorie, les cellules souches adultes de l'épithélium acquièrent des altérations génétiques qui leur permettent, en plus de leur capacité d'auto-renouvellement illimité, de proliférer de façon continue et de manière indépendante vis-à-vis de leur « niche environnementale ». Les cellules souches étant multipotentes, elles sont également responsables de l'hétérogénéité cellulaire observée au sein d'une tumeur. L'existence des CSC expliquerait également le gradient de différenciation cellulaire observé dans les tumeurs (Reya et al., 2001 ; Clevers, 2011).

Bien que ces deux théories soient différentes, elles reposent toutes deux sur les mêmes principes : une cellule, par altération génétique, acquiert des capacités de prolifération augmentées par rapport aux autres cellules du tissu. Ces mutations leur confèrent un avantage sélectif de prolifération ; elles permettent la formation d'une tumeur hétérogène. Cependant, ces deux théories possèdent leur limite puisque selon le modèle stochastique, il est difficile d'expliquer l'hétérogénéité cellulaire observée dans les tumeurs. Selon le modèle hiérarchique, il est difficile d'expliquer la différenciation des cellules souches car elles perdraient leur potentiel d'auto-renouvellement illimité.

Pour répondre à ces questions, récemment, deux nouveaux concepts ont été proposés. Ces concepts tiennent compte des deux théories déjà existantes. Selon ces concepts, une cellule initiatrice de cancer (CIC) acquiert les premières mutations sporadiques à l'origine du cancer. Il est important de noter que dans ces théories, la cellule initiatrice n'est pas nécessairement une cellule souche. Le premier concept, émis par Visvader, est : une CIC, par mutations génétiques ou épigénétiques, forme tous les sous-types tumoraux (fig 5A) (Visvader, 2011).

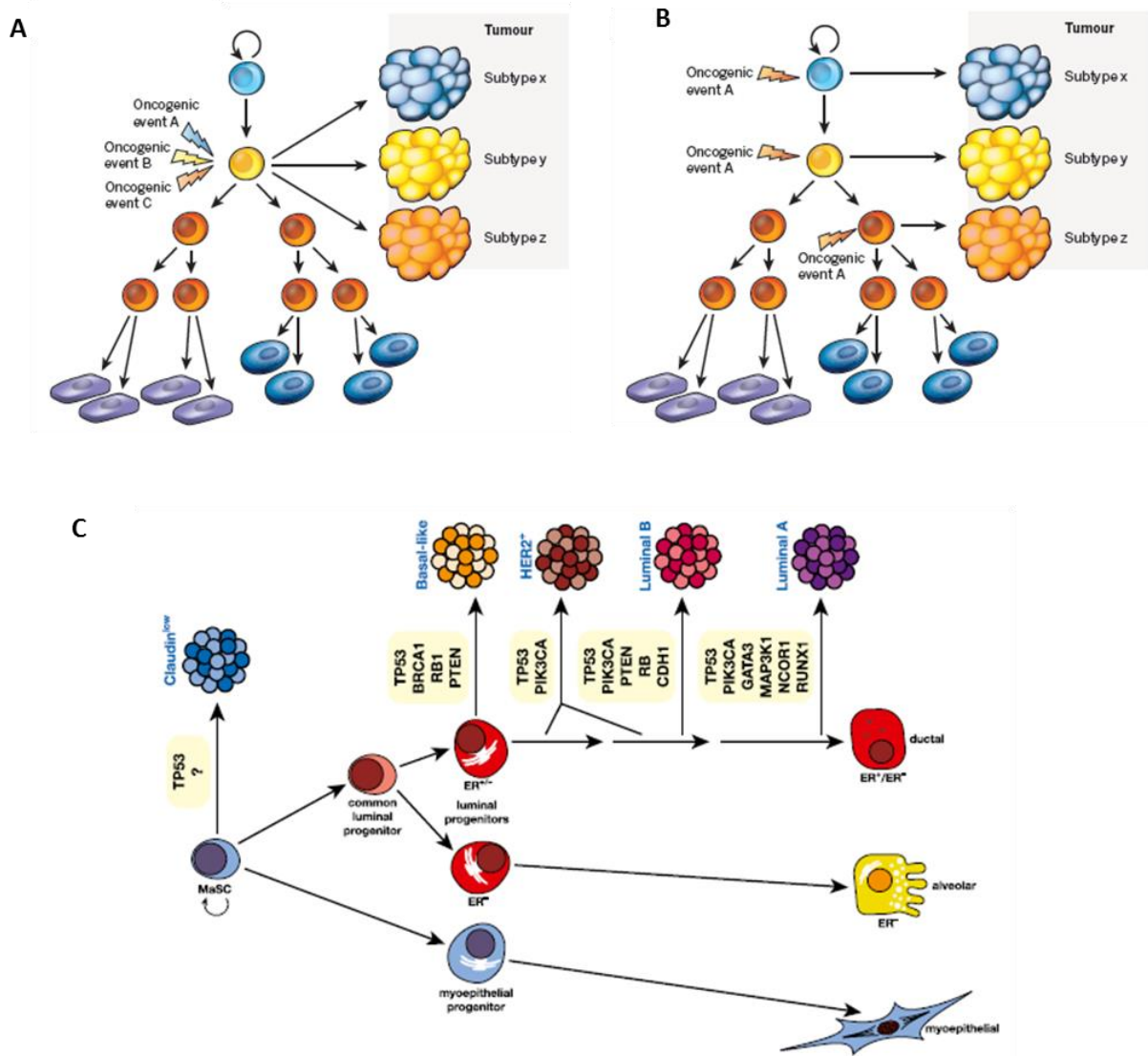


Figure 5. Mécanismes d'apparition des cancers selon Visvader. (A) Une cellule initiatrice de cancer (en jaune) subit des altérations génétiques conduisant à l'apparition des différents sous-types de cancer. (B) Un même évènement oncogénique se produit sur des cellules à différents stades de différenciation ; ce qui permet l'apparition des sous-types tumoraux. (C) Mécanisme d'apparition du cancer du sein selon Visvader. Plusieurs évènements oncogéniques apparaissent à différents stades de différenciation cellulaire conduisant à l'apparition des principaux sous-types tumoraux existant dans le cancer du sein. (Visvader, 2011; Visvader & Stingl, 2014).

Le second concept, émis par ce même auteur, est : un même évènement oncogénique peut survenir à différents stades de différenciation cellulaire et par conséquent former les différents sous-types tumoraux (fig 5B). Selon Hanahan et Weinberg, 6 altérations géniques sont nécessaires pour induire la transformation d'une cellule normale en une cellule cancéreuse. L'ordre d'apparition de ces mutations peut également influencer sur le sous-type tumoral.

Récemment, cet auteur a appliqué ces concepts de l'apparition tumorale dans le cancer du sein (Visvader & Stingl, 2014). Selon lui, les cellules initiatrices de cancer du sein sont différentes en fonction du sous-type tumoral (fig 5C). En effet, en comparant les profils d'expressions géniques entre les différentes tumeurs et les différents types cellulaires de la glande mammaire, il est apparu que les tumeurs de type Claudin-low dériveraient des cellules souches de la glande mammaire et que les tumeurs basal-like, HER2+, lumineales A et B dériveraient des cellules progénitrices lumineales. Ces trois derniers sous-types seraient ensuite différenciés par de nouvelles altérations génétiques. Dans ce mécanisme d'oncogenèse, les deux concepts précédemment émis sont utilisés.

L'hétérogénéité tumorale mise en avant dans l'ensemble de ces théories est responsable des différents échecs thérapeutiques. Afin de palier à ces échecs, il est donc primordial d'identifier toutes les sources de cette hétérogénéité.

c. Influence du microenvironnement tumoral

La survenue du cancer résulte donc d'une succession d'altérations génétiques et épigénétiques conduisant à la transformation d'une cellule saine en une cellule tumorale. Cependant dans les cancers, les mutations sur les cellules épithéliales ne conduisent pas à elles seules à l'apparition de la tumeur. Aujourd'hui, il est admis que le microenvironnement tumoral subit également des altérations qui sont nécessaires au développement de la tumeur. En effet, selon Hanahan et Coussens, 7 des 8 altérations cellulaires essentielles à la transformation cellulaire sont influencées par le microenvironnement tumoral (Hanahan & Coussens, 2012).

Le microenvironnement des cellules épithéliales mammaires est constitué d'une matrice extracellulaire et d'une multitude de cellules non épithéliales : les adipocytes (en grande majorité), les macrophages, les fibroblastes, les lymphocytes et les cellules endothéliales. L'ensemble de ces cellules permet de maintenir l'homéostasie tissulaire de la glande mammaire.

Les cellules tumorales mammaires modifient ce microenvironnement pour leur permettre d'échapper au système immunitaire. Ces modifications du microenvironnement facilitent également la prolifération, la survie, l'invasion, la transition épithélio-mésenchymateuse (TEM), l'angiogenèse mais aussi le remodelage de la matrice extracellulaire (fig 6).

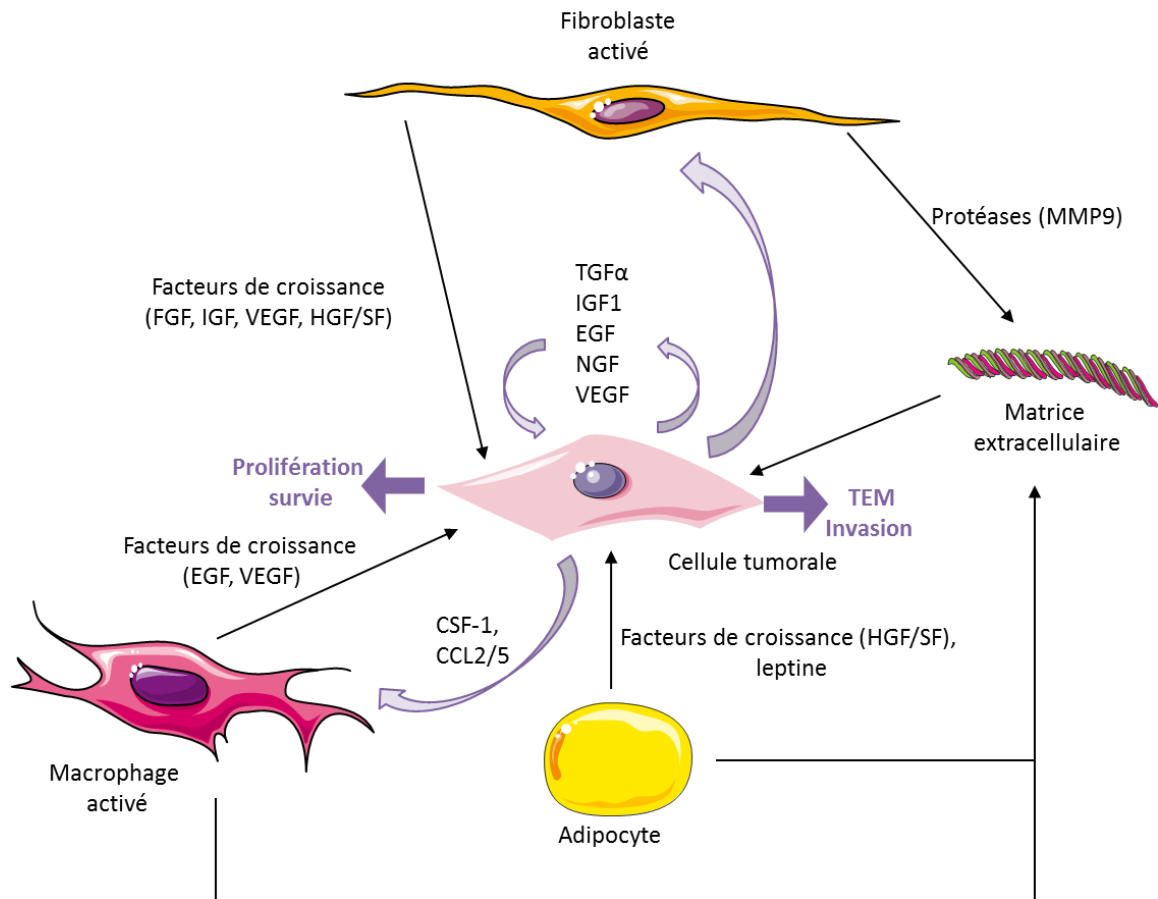


Figure 6. Interaction de la cellule tumorale avec son microenvironnement. La cellule tumorale, par sécrétion de plusieurs composés, active les fibroblastes ainsi que les macrophages. Ces derniers ainsi que les adipocytes, en sécrétant des facteurs de croissance, des protéases et des composants de la matrice extracellulaire, favorisent la prolifération de la cellule tumorale, son invasion et l'angiogenèse.

Les cellules tumorales mammaires dialoguent avec l'ensemble des cellules du microenvironnement et activent leurs effets pro-tumoraux (Mueller & Fusenig, 2004 ; Hu & Poyak, 2008 ; Polyak & Kalluri, 2010 ; Mao et al., 2013). En effet, elles sont capables de recruter les fibroblastes de ce microenvironnement pour les convertir en fibroblastes associés au cancer ou CAF (*cancer associated fibroblasts*). Une fois activés, les CAF sécrètent un grand nombre de composés, notamment des composants de la matrice extracellulaire (fibronectine, collagène, ...), des cytokines, des protéases mais aussi des facteurs de croissance (FGF, IGF, VEGF, HGF/SF) qui favorisent le développement du cancer (Elenbaas & Weinberg, 2001 ; Kalluri & Zeisberg, 2006). En effet, les cytokines augmentent la prolifération cellulaire en activant les boucles autocrine et paracrine (Orimo et al., 2005) ; les protéases, en remodelant la matrice extracellulaire, sont impliquées dans l'invasion tumorale et la transition épithélio-mésenchymateuse (Lochter et al., 1997 ; Kleer et al., 2008) ; et enfin, les facteurs de croissance permettent la prolifération cellulaire, la survie, l'angiogenèse et l'invasion tumorale (fig 6).

Les cellules cancéreuses mammaires, en sécrétant des chimioattractants tels que le CSF-1 (*Colony-Stimulating Factor-1*) ou les CCL2/5 (*Chemokine C-C motif ligand 2/5*), sont également capables de recruter des macrophages associés aux tumeurs (TAM, *tumor-associated macrophages*) (Soria & Ben-Baruch, 2008 ; Laoui et al., 2011 ; Obeid et al., 2013). Ces derniers sécrètent des facteurs de croissance notamment l'EGF qui sera capable de stimuler les cellules épithéliales cancéreuses mammaires et donc d'augmenter leur prolifération, migration et invasion (Wyckoff et al., 2004 ; Goswami et al., 2005). Les TAM sont également capables de sécréter du VEGF qui favorisera l'angiogenèse de la tumeur (Carmeliet, 2003) (fig 6).

De plus, les adipocytes, présents en majorité dans le sein, sont capables de sécréter des facteurs de croissance (HGF/SF), des adipokines (leptine) ou encore des protéines de la matrice extracellulaire (collagène) qui favorisent la croissance et la migration/invasion des cellules tumorales (Iyengar et al., 2005 ; Vona-Davis & Rose, 2007 ; Walter et al., 2009 ; Dirat et al., 2011) (fig 6). En effet l'obésité, en plus d'augmenter le risque de cancer du sein, est associée à un mauvais pronostic et à un phénotype tumoral agressif (Rose & Vona-Davis, 2014).

Le microenvironnement, outre ces effets pro-tumoraux, est également capable de sécréter des facteurs anti-tumoraux. En effet, notre équipe a mis en évidence que les cellules épithéliales normales mammaires sécrètent de l'IGFBP-3 (*Insulin-like Growth Factor Binding Protein 3*) et de la maspine afin d'induire l'apoptose des cellules cancéreuses de ce tissu (Toillon et al., 2007). Ces cellules sécrètent également des facteurs inhibiteurs de croissance tels que le

TGF- β 1, le TNF- α , le MDGI ou encore des interleukines ou des protéoglycanes (Bourhis et al., 2000).

Outre le microenvironnement, notre laboratoire a mis en évidence que les cellules cancéreuses mammaires sécrètent des facteurs de croissance, de façon autocrine : les neurotrophines. Ces facteurs de croissance sont le NGF (*Nerve Growth Factor*), le BDNF (*Brain-Derived Neurotrophic Factor*) et la NT4/5 (*Neurotrophin-4/5*). Ils sont capables d'induire la prolifération, la survie et la migration/invasion de ces cellules cancéreuses mammaires (Descamps et al., 2001 ; Dolle et al., 2003 ; Vanhecke et al., 2011 ; Demont et al., 2012). Par ailleurs, nous avons mis en évidence que ces facteurs peuvent faire l'objet de cible thérapeutique. En effet, des inhibiteurs de NGF, BDNF ou de NT4/5 réduisent la formation de tumeur mammaire dans des modèles de xénogreffes (Adriaenssens et al., 2008 ; Vanhecke et al., 2011 ; Aubert et al., 2015).

Ainsi, cibler les facteurs de croissance et l'activation de leurs voies de signalisation semble avoir un intérêt thérapeutique privilégié dans le traitement du cancer du sein.

4. Classification des cancers du sein

A l'heure actuelle, il est impossible de définir un stade précoce et un stade tardif de la tumeur, puisque cela supposerait une évolution régulière et linéaire de la formation tumorale. Cependant, il est important de classer les tumeurs afin de créer des groupes homogènes sur le plan pronostic mais aussi thérapeutique. Aujourd'hui, il existe plusieurs types de classification des tumeurs mammaires. Dans ce paragraphe, trois classifications différentes seront abordées : la première tient compte des caractéristiques macroscopiques, histologiques et pronostiques des tumeurs ; la seconde tient compte de la taille de la tumeur (T), de l'invasion ganglionnaire (N) et de l'invasion métastatique (M) des autres tissus : il s'agit de la classification dite « TNM ». Enfin la dernière classification tient compte des altérations génétiques observées dans la tumeur.

a. La classification histologique

En fonction de leur diagnostic et de leur évolution, les tumeurs mammaires peuvent être bénignes ou malignes. Ces deux groupes de tumeurs sont définis en fonction des caractéristiques macroscopiques, histologiques et pronostiques des tumeurs.

Une tumeur est dite « **bénigne** » lorsqu'elle est bien délimitée et encapsulée dans le tissu. Le plus souvent, il s'agit d'une masse tissulaire qui est non cancéreuse. Les cellules constituant cette masse ont une croissance lente et restent localisées à proximité de leur lieu de naissance dans le tissu. Ainsi, ces tumeurs ne forment aucune métastase et après exérèse complète,

aucune récurrence n'est observée. Ces tumeurs ont un diagnostic très favorable et sont fréquemment diagnostiquées avant l'âge de 30 ans chez la femme.

A l'inverse du groupe précédent, les tumeurs **malignes** sont généralement non délimitées et non encapsulées. Les cellules présentes dans ces tumeurs ont une croissance rapide et sont capables d'infiltrer et de détruire les tissus adjacents à leur lieu de naissance. Ces tumeurs peuvent former des métastases et induire une récurrence après excision complète de la tumeur primaire. La majorité des tumeurs malignes (95% des cas) sont des adénocarcinomes : des tumeurs ayant pour origine les cellules épithéliales glandulaires. Lorsque les cellules cancéreuses sont issues des cellules des canaux, nous sommes en présence d'un cancer canalaire ; lorsque les cellules cancéreuses proviennent des cellules lobulaires, nous parlons de cancer lobulaire. En fonction de leur envahissement dans le tissu environnant, les adénocarcinomes sont classés en plusieurs groupes :

➤ Les carcinomes *in situ* (non infiltrants) :

Il existe deux types de carcinomes *in situ* : les carcinomes canauxiers *in situ* (CCIS) qui proviennent des cellules de la lumière du canal galactophore ; les carcinomes lobulaires *in situ* (CLIS) qui proviennent de la prolifération des acini des lobules. Ces carcinomes n'envahissent pas le tissu conjonctif, la membrane basale est intacte ; ils sont donc généralement de bon pronostic. Cependant, ces carcinomes sont rares (2-3%) puisqu'ils ne sont jamais diagnostiqués à ce stade.

➤ Les carcinomes invasifs (infiltrants) :

Ces carcinomes sont caractérisés par une membrane basale altérée et un envahissement tumoral dans le tissu conjonctif environnant. Comme les carcinomes *in situ*, deux types de carcinomes invasifs existent : les carcinomes invasifs canauxiers (75% des cas de cancer du sein) et les carcinomes invasifs lobulaires (5 à 15% des cas de cancer du sein). Ces carcinomes invasifs représentent la majorité (98% des cas) des cancers du sein puisque le cancer du sein est généralement diagnostiqué à ce stade. Des formes rares de ces carcinomes invasifs ont été diagnostiquées : les carcinomes tubuleux, mucineux, médullaires, papillaires ou encore les carcinomes apocrines, neuroendocrines et adénoïdes kystiques.

➤ Les cancers du sein métastatiques :

Cette forme de cancer du sein provient généralement des cellules néoplasiques des carcinomes invasifs. Ces cellules sont capables de disséminer à partir de la tumeur primaire dans le réseau lymphatique préférentiellement mais aussi dans les vaisseaux sanguins. Les premiers relais ganglionnaires (ganglions axillaires, les ganglions des chaînes mammaires internes et sub-claviculaires) sont alors colonisés et peuvent être à l'origine d'une

Introduction

dissémination générale du cancer. Cet envahissement ganglionnaire est le reflet du potentiel métastatique des cellules de la tumeur primaire. La dissémination des cellules cancéreuses mammaires n'est pas aléatoire dans l'organisme ; un tropisme pour les os, les poumons, le cerveau et le foie a été observé.

- Les autres formes de cancer du sein (rare) :

Le cancer inflammatoire : Les cellules cancéreuses infiltrent le réseau lymphatique dans la peau du sein, ce qui engendre une embolie lymphatique ; il s'agit donc d'un cancer très agressif et de mauvais pronostic. Il est dit « inflammatoire » puisqu'il est souvent associé à une augmentation du volume du sein, et des changements au niveau de la peau du sein (rougeur, chaleur, œdème).

La maladie de Paget du mamelon : il s'agit d'un cancer du sein qui est caractérisé par une éruption cutanée de la peau du mamelon et de l'aréole. Le plus souvent, ce signe clinique est associé à un adénocarcinome sous-jacent qui prend naissance dans les canaux du sein et se répand jusqu'à l'épiderme du mamelon et de l'aréole.

Dans de très rares cas, d'autres formes de cancer du sein peuvent être diagnostiquées : des tumeurs phyllodes, des sarcomes ou encore des lymphomes du sein. Les tumeurs phyllodes sont, le plus souvent, des tumeurs dites « bénignes » ; elles sont traitées par une chirurgie. Enfin, les sarcomes et lymphomes du sein sont traités de la même manière que les autres sarcomes et lymphomes non Hodgkinien, par chimiothérapie.

b. La classification TNM

Cette classification repose principalement sur des observations anatomiques de la tumeur. En effet, dans un premier temps, le volume tumoral est mesuré et un score T compris entre 0 et 4 lui est attribué. Il est à noter que plus la valeur est élevée, plus le volume de la tumeur et son extension à la cage thoracique est important. Ensuite, l'invasion ganglionnaire par les cellules tumorales est analysée et également scorée avec des valeurs N comprises entre 0 et 3. La cotation N dépend du territoire ganglionnaire envahi par les cellules tumorales, plus ou moins proche du site de la tumeur, des dimensions des adénopathies, de leur nombre et de leur éventuelle fixation aux tissus voisins. Enfin, l'invasion métastatique M est analysée et notée M0 ou M1. La valeur M0 est associée à l'absence de métastase et la valeur M1 à la présence de métastase, quelque soit leur nombre, leur volume et leur localisation. Une fois l'ensemble des valeurs TNM diagnostiquées, nous pouvons déterminer le stade de la progression tumorale. Ce stade est noté de 0 à IV ; où 0 correspond à un cancer *in situ* de petite taille, sans invasion ganglionnaire ni de métastase et IV correspond à un cancer avancé avec envahissement ganglionnaire et la présence de métastases multi-sites T4N3M1.

c. La classification moléculaire

Pendant de nombreuses années, la classification des différents sous-types de cancer du sein reposait sur l'expression des récepteurs hormonaux, notamment des récepteurs aux œstrogènes, à la progestérone et du récepteur HER2. Mais, les différents sous-groupes de patients ainsi générés ne répondaient pas de la même manière aux traitements. En 2000, Perou et collaborateurs ont étudié les profils d'expression des gènes de 65 tumeurs mammaires obtenues à partir de 42 individus par microarray (Perrou et al., 2000). Sur les 8102 gènes présents sur la puce, ils n'en ont retenu que 496. L'expression de ces gènes était significativement différente entre les tumeurs, mais avec de faibles variations entre deux tumeurs issues d'un même patient. L'analyse de ces gènes a permis de mettre en évidence deux groupes : un groupe dit œstrogènes récepteur-négatif (ER-) et un groupe dit œstrogènes récepteur-positif (ER+). Une analyse plus précise de ces groupes montre que les tumeurs ER+ expriment les récepteurs aux facteurs hormonaux (œstrogènes et progestérone), les cytokératines spécifiques des cellules luminales ainsi que d'autres marqueurs de ce type cellulaire : elles sont donc nommées tumeurs **luminales**. De plus, des détections des récepteurs aux œstrogènes et progestérone, d'HER2 et du marqueur de prolifération Ki-67 par immunohistochimies ont permis de mettre en évidence 2 sous-groupes de tumeurs luminales : **luminal-A** et **luminal-B**.

Dans la branche des tumeurs ER-, 3 sous-groupes ont été identifiés : les **HER2+**, **basal-like** et **normal-like**. Les différences entre ces groupes de tumeurs reposent sur des modifications d'expression du récepteur HER2 et sur l'expression de cytokératine spécifique des différents types cellulaires. Les caractéristiques de ces sous-groupes sont détaillées dans le tableau 3.

Récemment, deux autres types de tumeurs mammaires ont été mis en évidence : les tumeurs dites « molecular apocrine » qui montrent une activation du récepteur aux androgènes, et les tumeurs dites « claudin-low » qui sont caractérisées par une faible expression des marqueurs de jonction cellulaire (CDH1, Cadherin-1, CLDN, claudin).

sous-type	ER/PR/ HER	fréquence dans les cancers	autres marqueurs	gènes caractéristiques	grade histologi que	pronostic
luminal-A	ER+ PR+ HER2-	50-60%	CK8/18+ ER1+ Ki67-	LIV1, FOXA1, XBP1, GATA3, BLC2, erBb3, erbB4	grade I ou II	+
luminal-B	ER+/- PR+/- HER2- /+	15-20%	Ki67+	v-MYB, GGH, LAPTMB4, NSEP1, CCNE1	grade II ou III	+/-
HER2 positive	ER- PR- HER2+	15-20%	Ki67+ P53 mutée	EGFR, P-cadherine	grade III	-
basal-like	ER- PR- HER2-	8-37%	CK5/14/17+ laminine+ P53 mutée Ki67+	P-cadherine, calvéoline, EGFR,	grade III	- -
normal-like	ER-/+ PR? HER2-	5-10%	CK5- EGFR-		grade I	+/-
molecular apocrine	ER- PR- HER2+ /-		CK5+	AR, FAS, ERBB2, XBP1	grade II ou III	+/-
claudin low	ER- PR- HER2-	12-14%	CK5, CDH1 bas/- CLND1 bas/-	CD44, SNAI3	grade III	-

Tableau 3. Caractéristiques moléculaires des différents sous-types de cancer du sein. Pronostic : - - très défavorable ; - défavorable ; +/- intermédiaire ; + bon

III. Les thérapies dans le cancer du sein

Le cancer du sein est une maladie multifactorielle avec une hétérogénéité inter- et intra-tumorale très diversifiée. A l'heure actuelle, plusieurs stratégies thérapeutiques pour traiter ce cancer sont employées. Dans tous les types de cancer, la chirurgie est requise. Lorsque la tumeur est petite, une mastectomie partielle est réalisée. Dans des cas volume tumoral important (T3 ou T4), le chirurgien recourt à une mastectomie totale du sein.

La chirurgie est complétée par une radiothérapie, chimiothérapie ou hormono-thérapie, lorsque celle-ci est possible. Ces traitements sont souvent administrés de manière adjuvante, c'est-à-dire après la chirurgie ; ils permettent ainsi d'éliminer les cellules cancéreuses proliférantes qui seraient encore présentes dans l'organisme. Dans certains cas, ils seront administrés de façon néoadjuvante (avant la chirurgie). Cette option est proposée aux personnes ayant une tumeur de volume important (T3 ou T4) pour permettre une réduction du volume tumoral et ainsi éviter la mastectomie complète. L'ensemble des traitements sont décrits en fonction des trois types majeurs de cancer du sein : les ER⁺/PR⁺, les HER2⁺ et les triples négatifs (ER⁻/PR⁻/HER2⁻) et regroupés dans le tableau 4 (voir P. 31).

1. Les tumeurs hormono-dépendantes

Ces tumeurs représentent 60 à 65% des cancers du sein, elles sont souvent traitées par chimiothérapies et/ou hormono-thérapies. La chimiothérapie repose sur l'utilisation de molécules anthracyclines ou de taxanes. Ces deux composés possèdent des modes d'action différents mais permettent l'apoptose des cellules proliférantes. Les anthracyclines modifient la structure de l'ADN, ce qui empêche sa réplication et par conséquent entraîne la mort cellulaire. Les principales molécules anthracycliques utilisées pour traiter le cancer du sein sont : la doxorubicine, la gemcitabine ou encore le cisplatine. Elles sont administrées également dans les cancers du sein dit triples négatifs.

Les taxanes (le paclitaxel par exemple) sont des inhibiteurs de la formation des microtubules. En leur présence, toute cellule se divisant est incapable de former de nouveaux microtubules ; les cellules vont donc mourir. Ces différents groupes d'agents chimiothérapeutiques sont relativement efficaces ; cependant, ils ne sont pas spécifiques des cellules cancéreuses.

Les tumeurs hormono-dépendantes peuvent également être traitées par des inhibiteurs de l'activation du récepteur aux œstrogènes : le tamoxifène ; ou encore des inhibiteurs d'aromatases (IA). L'aromatase est un enzyme permettant la conversion de la testostérone en œstradiol. Ce dernier est ensuite converti en œstrogènes sous l'action d'autres enzymes. Les IA empêchent la synthèse de l'œstrogène par l'organisme. En l'absence d'œstrogènes, les cellules œstrogènes-dépendantes ne prolifèrent plus ; elles vont donc mourir. Une étude a comparé les effets du tamoxifène et des IA chez les femmes ménopausées. Il est apparu que

les femmes traitées avec les IA montrent une meilleure réponse au traitement avec une diminution significative du taux de métastases (Howell et al., 2005). Chez ces personnes, les IA sont donc les molécules de premier choix dans ce type de cancer.

2. Les tumeurs HER2+

Ce type tumoral représente 20 à 25% des cancers du sein ; il est caractérisé par une surexpression du récepteur orphelin HER2. Comme décrit précédemment, ce récepteur appartient à la famille des récepteurs HER 1 à 4. Les principales stratégies thérapeutiques mises en œuvre pour traiter ces tumeurs utilisent des anticorps monoclonaux anti-HER. La molécule la plus utilisée est le trastuzumab aussi connu sous le nom d'Herceptine. Il s'agit d'un anticorps monoclonal spécifique du domaine juxta-membranaire du récepteur HER2. Cet anticorps, en se fixant sur ce récepteur, empêche ainsi son activation. Il permet également l'activation des protéines régulatrices du cycle cellulaire p21 et p27. Ces protéines vont alors inhiber la prolifération cellulaire en induisant un arrêt du cycle cellulaire (Miller et al., 2014).

Des résistances à ce traitement ont rapidement été constatées, notamment dans les tumeurs ayant une forme tronquée de HER2. Pour pallier ces phénomènes, d'autres anticorps anti-HER sont apparus. Ces anticorps monoclonaux reconnaissent différents domaines sur les récepteurs HER ; ils possèdent donc des modes d'actions complémentaires. La première molécule apparue, le lapatinib (Tykerb), est un inhibiteur spécifique du domaine intracellulaire du récepteur anti-HER2. Cet inhibiteur est également capable de se fixer sur le récepteur HER1 (EGFR) et d'empêcher son activation. Il possède donc deux cibles potentielles sur les cellules tumorales.

Le pertuzumab (Perjeta) se fixe sur le domaine de dimérisation de HER2 avec le récepteur HER3. Une étude préclinique a montré un effet synergique de 50% du trastuzumab avec le pertuzumab (Scheuer et al., 2009 ; Baselga, 2010).

Un anticorps anti-HER3 (MM-121) ainsi qu'un anticorps anti-HER2/HER3 (MM-111) sont actuellement en essai clinique de phase I/II. La première molécule empêche l'activation du récepteur HER3 tandis que la seconde empêche l'hétérodimérisation du récepteur HER2 avec le récepteur HER3.

Enfin, la molécule T-DM1, composée du trastuzumab couplé à un dérivé de maytansine, montre des effets prometteurs dans le cadre de tumeur HER2+. La maytansine est une molécule inhibitrice de la formation des microtubules. Ce composé T-DM1 se fixe sur le récepteur HER2, il est alors internalisé dans la cellule. Ensuite, sa dégradation par le protéasome permet la libération de la maytansine qui va alors empêcher la formation des microtubules et par conséquent induire la mort cellulaire. Des essais cliniques de phase I ont

montré que le T-DM1 possède les mêmes effets thérapeutiques que la combinaison du trastuzumab avec le docetaxel (un agent chimiothérapeutique) (Higgins & Baselga, 2011). Cette étude montre également que le T-DM1 ne possède pas les effets secondaires usuellement observés lors d'un traitement chimiothérapeutique puisque la maytansine est active uniquement dans les cellules cancéreuses HER2+.

Par ailleurs, pour pallier aux résistances aux anticorps, des inhibiteurs de voies de signalisation, notamment la voie PI3K/Akt/mTOR, sont également utilisés. La voie de signalisation mTOR (*mammalian Targets Of Rapamycin*) joue un rôle central dans la prolifération cellulaire, la survie, la migration ou encore l'angiogenèse. Un inhibiteur naturel de cette voie de signalisation existe : la rapamycine. De plus, une étude a mis en évidence que la rapamycine inhibe la prolifération de 20 lignées cellulaires cancéreuses mammaires et augmente les effets de l'Herceptine (Wang et al., 2007a). Ainsi des dérivés de la rapamycine, plus facilement solubilisable et avec une meilleure stabilité, ont été développés : l'everolimus par exemple. Cette molécule restaure la sensibilité cellulaire aux anti-HER2 (Serra et al., 2008). Cette molécule est actuellement en essai clinique de phase I.

Outre l'activation constitutive des voies de signalisation des récepteurs, un second mécanisme de résistance aux anti-HER2 a été identifié : c'est l'activation du récepteur IGF1R. Afin de contrebalancer ce mécanisme de résistance, un inhibiteur de l'activation d'IGF1R (dalotuzumab) est actuellement en essai clinique de phase II (Merck, clinicaltrials.gov).

3. Les tumeurs triples négatives

Ces tumeurs représentent 10 à 15% des cancers du sein. De par l'absence de récepteurs hormonaux, il est difficile de les cibler spécifiquement. Ces tumeurs sont très agressives et disséminent facilement dans l'organisme. Le principal traitement est l'utilisation d'anthracyclines et de taxanes.

Cependant, ces tumeurs sont souvent associées à un statut *BRCA1/2* muté. Une étude a mis en évidence que l'inhibition de PARP dans les tumeurs où la réparation de l'ADN est impossible, conduit à la mort cellulaire (Iglehart & Silver, 2009). Ainsi, des molécules inhibitrices de PARP (olaparib et iniparib) sont actuellement en cours d'essai clinique. L'olaparib montre des effets bénéfiques lorsqu'il est administré seul à des patientes possédant des mutations pour *BRCA1* et/ou 2. Plusieurs essais cliniques de phase I testent les effets de l'olaparib en combinaison avec des agents chimiothérapeutiques (cisplatine, carboplatine, paclitaxel) (Tomao et al., 2015). L'iniparib est un inhibiteur de PARP ; cependant, une récente étude a montré qu'il est capable d'induire les foyers de réparation γ -H2AX et un arrêt du cycle cellulaire en phase G2/M. La combinaison de l'iniparib avec la gemcitabine ou le cisplatine

augmente la durée de vie sans métastases de façon significative comparée à la chimiothérapie seule (O'Shaughnessy et al., 2011).

Par ailleurs, le récepteur HER1 (EGFR) est fréquemment surexprimé dans les tumeurs triples négatives. Il est également associé à un mauvais pronostic. Un anticorps anti-EGFR est actuellement en essai clinique de phase II : il s'agit du cetuximab (Erbatus). Lorsqu'il est utilisé en combinaison avec un agent chimiothérapeutique (cisplatine), de bons résultats ont été obtenus (Baselga et al., 2010). En effet, l'utilisation de cette combinaison double le taux de réponse comparé aux résultats obtenus avec le cisplatine seul. D'autres combinaisons sont actuellement en cours d'essai clinique : cetuximab+carboplatine ; cetuximab+irinotecan.

D'autres inhibiteurs de l'EGFR (erlotinib et lapatinib) sont actuellement en cours d'essai en combinaison avec des agents chimiothérapeutiques (carboplatine, docetaxel).

Les tumeurs triples négatives étant fortement prolifératives, un essai a combiné l'inhibiteur de la voie de signalisation mTOR (l'everolimus) avec un inhibiteur d'EGFR (lapatinib). Il est apparu que ces deux molécules possèdent un effet synergique dans ce sous-type tumoral. Par ailleurs, sur des souris immunodéficientes xénotransplantées, la tumeur formée possède un retard de croissance lorsque les souris sont traitées avec cette combinaison thérapeutique comparé aux souris contrôles (Liu et al., 2011).

4. Les inhibiteurs non spécifiques

Pour finir, les tumeurs ont besoin de nutriments et d'oxygène pour se développer. Pour cela, elles sont capables de former de nouveaux vaisseaux sanguins : c'est la néo-angiogenèse. Un puissant anti-VEGF est actuellement utilisé pour traiter les cancers du sein quel que soit le sous-type tumoral : c'est le bevacizumab (Avastin). Il s'agit d'un anticorps monoclonal ciblant le récepteur VEGF et empêchant ainsi son activation. Une étude a comparé les effets du paclitaxel seul ou en combinaison avec le bevacizumab. Une meilleure réponse thérapeutique chez les patientes ayant reçu la combinaison a été observée (de 21% à 37% respectivement) (Miller et al., 2007). Cependant, le bevacizumab est associé à d'importants effets secondaires : accidents vasculaires cérébraux, défaut de cicatrisation des plaies, dommages sur des organes (Choueiri et al., 2011). Ainsi les Etats-Unis ont retiré cette molécule dans le cadre du traitement du cancer du sein.

A ce jour, d'autres inhibiteurs multi-kinases sont en cours d'essai clinique. Le sunitinib est une molécule ciblant les trois récepteurs aux VEGF, PDGFR α et β , RET, c-Kit, FLT3 et le récepteur au CSF-1. Un essai clinique de phase II a mis en évidence que le sunitinib administré en monothérapie possède un effet bénéfique de 16% chez les patientes traitées. Aucun bénéfice n'a été observé si cette molécule est administrée en combinaison avec le docetaxel comparé à son administration seule (Tomao et al., 2015).

Un second composé multi-kinase est également en cours d'essai thérapeutique : le sorafenib. Cette molécule cible les récepteurs au VEGF1 et 2, PDGFR α et β , c-Kit, FLT3 mais aussi la kinase Raf. Cette molécule est donc inhibitrice de l'angiogenèse mais aussi inhibitrice de la prolifération cellulaire. Elle a montré des effets bénéfiques lorsqu'elle est administrée en combinaison avec le capecitabine (un agent chimiothérapeutique) (Higgins & Baselga, 2011).

D'autres agents thérapeutiques sont actuellement en cours d'essai clinique. Ces agents sont des inhibiteurs : du cycle cellulaire (anti-Chk (UCN-01), anti-CDK (Dinaciclib)) ; du protéasome (Bortezomib) ; des histones désacétylases (Panobinostat) ; de kinase (Dasatinib) ; de la voie Wnt (PORCN, LGK974) (Tomao et al., 2015). Ces molécules doivent encore montrer leur efficacité en administration seule ou en combinaison et être peu cytotoxiques. Par conséquent, il faut évaluer le rapport bénéfice/risque de l'utilisation de ces drogues pour traiter le cancer du sein.

L'ensemble de ces thérapies est regroupé et classé dans le tableau suivant (tableau 4).

ER+/PR+	HER2+	ER-/PR-/HER2-
tamoxifen	trastuzumab : anti-HER2	olaparib : inhibiteur PARP
inhibiteurs d'aromatase	lapatinib : inhibiteur HER2 inhibiteur HER1	iniparib : inhibiteur PARP
	pertuzumab : anti-HER2	cetuximab : anti-EGFR
	T-DM1 : anti-HER2+taxane	erolotinib : anti-EGFR
	MM-121 : anti-HER3	lapatinib: anti-EGFR
	MM-111: anti-HER2/HER3	
	dalotuzumab : anti-IGF1R	
chimiothérapies : doxorubicine, gemcitabine, cisplatine, paclitaxel, docetaxel, capecitabine		
inhibiteurs : everolimus : inhibiteur de mTOR bevacizumab : inhibiteur VEGF sunitinib : inhibiteur multi-kinase sorafenib : inhibiteur multi-kinase		
Tableau 4. Tableau regroupant les différentes thérapies utilisées pour traiter le cancer du sein en fonction du sous-type tumoral.		

Malgré l'éventail de molécules thérapeutiques déjà sur le marché, le cancer du sein est encore la première cause de mortalité par cancer chez la Femme en France et dans le monde. Afin de réduire cette mortalité, il est essentiel d'identifier tous les acteurs impliqués dans la tumorigenèse de la glande mammaire ainsi que dans les résistances aux traitements.

Les longs ARN non codants

Récemment le projet GENCODE a mis en évidence que 90% du génome est transcrit. Parmi ces transcrits, uniquement 2% sont des ARNm permettant la synthèse de protéines, les autres transcrits sont des ARN non codants mais aussi du bruit de fond transcriptionnel. Les ARN non codants sont divisés en deux groupes : les petits ARN non codants d'une taille inférieure à 200 nts ; et les longs ARN non codants d'une taille supérieure à 200 nts. Les longs ARN non codants (lncRNA) sont impliqués dans tous les processus cellulaires tels que la prolifération, la migration, l'apoptose, ou encore le maintien des cellules souches. Dans ce chapitre, nous étudierons l'ensemble des mécanismes mis en place par les lncRNA pour induire ces différents phénotypes. Les petits ARN non codants feront l'objet du chapitre suivant (voir P. 55).

I. Les longs ARN non codants : classification

1. Découverte des longs ARN non codants

En 1950, les scientifiques de l'époque pensaient que l'Homme (un organisme fortement développé et complexe) possédait un génome plus important qu'un organisme peu développé (Kung et al., 2013). En 1970, des expériences d'hybridation ADN-ARN ont mis en évidence que l'Homme ne possède que 20 000 à 30 000 gènes codants. A l'époque, le reste du génome est appelé « ADN poubelle » ou « *junk DNA* » puisqu'il contient des transposons, des pseudogènes ou des séquences répétées. En 1980, il a été mis en évidence que ces séquences non codantes pouvaient réguler l'intégrité du génome, l'expression des gènes, la synthèse des ARNm matures ou encore servir de réservoir à l'évolution des espèces. En 2005, par séquençage du génome entier, il a été mis en évidence que le taux de séquences non codantes dans un organisme est le reflet de sa complexité. Ces séquences non codantes ont longtemps été considérées comme étant un bruit de fond transcriptionnel sans fonction dans la cellule. Au cours des dernières années, un intérêt grandissant a été porté à l'étude de ces fragments d'ADN.

En 2012, le consortium ENCODE, a mis en évidence 9640 loci permettant l'expression des longs ARN non codants chez l'Homme (Bernstein et al., 2012; Derrien et al., 2012). Le projet GENCODE a, quant à lui, mis en évidence que 40% des longs ARN non codants sont transcrits à partir de locus contenant des gènes codants pour des protéines. Aujourd'hui, selon la base de données LNCipedia 2.0, 32 183 lncRNA ont été annotés. Pour la majorité d'entre eux, leur fonction dans la cellule doit encore être identifiée.

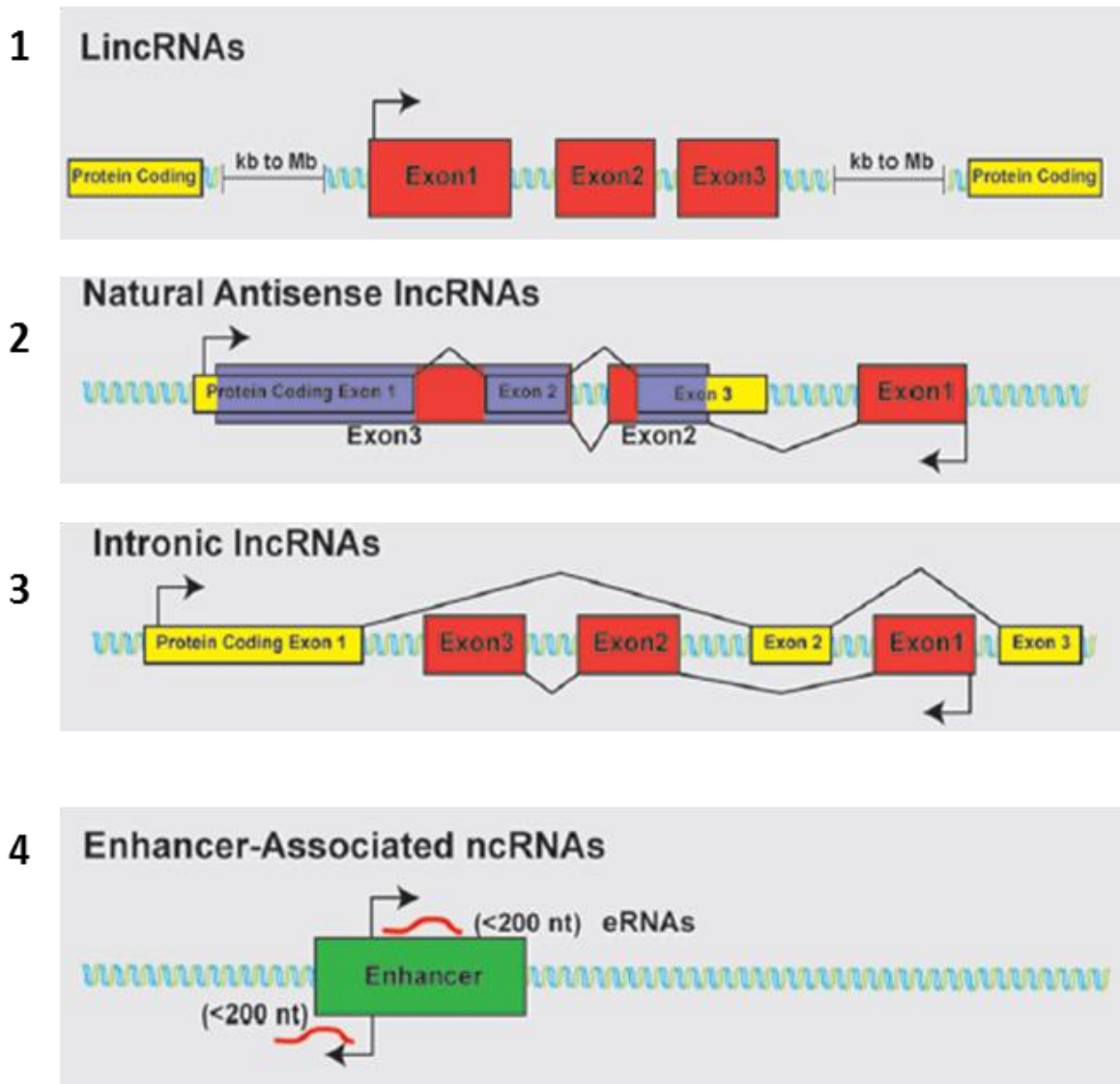


Figure 7. Classification des longs ARN non codants (lncRNA). (1) Les lncRNA sont classés en fonction de leur origine dans le génome. Les longs ARN intergéniques sont issus de fragments distants d'au moins 1 kb d'un gène codant pour une protéine (lincRNA). (2) Ils peuvent être transcrits en orientation antisens aux gènes codants. (3) Lors de l'épissage des ARNm, les introns peuvent former des lncRNA : ce sont les lncRNA introniques. (4) Ils peuvent également être transcrits à partir des enhanceurs (eRNA). (adapté de Clark & Blackshaw, 2014).

2. Classification des longs ARN non codants

Les longs ARN non codants peuvent être classés selon leur origine de transcription dans le génome, dans ce cas, nous comptabiliserons 5 groupes de longs ARN non codants. Certains lincRNA possèdent leur propre promoteur, ils sont transcrits par l'ARN polymérase de type II, polyadénylés et épissés : ce sont les **large-intergenic non coding RNA** (lincRNA) (fig 7-1). Ces lincRNA sont transcrits à partir de séquence comprise entre deux gènes codants mais leur transcription ne déborde pas sur l'expression de ces gènes. Ils sont généralement localisés à plus de 5 kb des gènes codants et ont une taille supérieure à 1000 nts (Kung et al., 2013; Shi et al., 2013). Nous pouvons citer par exemple les lincRNA *Xist* (Brockdorff et al., 1992; Brown et al., 1992), *HOTAIR* (Rinn et al., 2007), *MALAT1* (Ji et al., 2003) ou encore *H19* (Brannan et al., 1990). Les différents modes d'action de ces 4 ARN non codants seront développés dans le paragraphe suivant (voir P. 37).

Ensuite, les lincRNA peuvent être transcrits en orientation **antisens** aux gènes codants (fig 7-2). Ces transcrits antisens peuvent déborder partiellement ou entièrement sur ces gènes codants. Ils sont principalement retrouvés dans des régions soumises à l'empreinte génomique. Les lincRNA les plus connus et étudiés sont transcrits à partir des loci suivants : *Tsix/Xist* (Lee et al., 1999), *Kcnq1/Kcnq1ot1* (Kanduri et al., 2006) et *Igf2r/Air* (Lyle et al., 2000). Ces transcrits y contrôlent la mise en place et le maintien de l'empreinte. Un chapitre sera consacré à l'étude de l'empreinte génomique et le rôle des lincRNA dans la mise en place et le maintien de celle-ci (voir P. 73).

Certains lincRNA sont issus de **pseudogène**. A cause de mutation non sense ou d'un décalage de la phase ouverte de lecture, les pseudogènes sont généralement transcrits en lincRNA. Les lincRNA peuvent acquérir de nouvelles fonctions dans la cellule et réguler de façon post-transcriptionnelle l'expression des gènes par exemple.

Lors de l'épissage des ARNm, les introns éliminés peuvent former des lincRNA : ce sont les **longs ARN non codants introniques**. Ils ont une fonction similaire à l'ARNm à partir duquel ils sont synthétisés. Par exemple, l'ARN non codant *COLDAIR*, transcrit à partir du locus FLC (Flowering repressor locus), est impliqué dans le phénomène de vernalisation de la plante (fig 7-3) (Heo & Sung, 2011).

Enfin, la dernière classe des ARN non codant sont des transcrits peu abondants dans la cellule, qui sont rapidement dégradés par les exosomes et qui possèdent des fonctions peu connues. Ils sont synthétisés à proximité des points d'initiation de la transcription des ARNm mais aussi des ARN antisenses ; ils sont très courts (d'une taille allant de 20 à 2500 nts), coiffés et polyadénylés. Leur transcription est probablement impliquée dans l'ouverture de la chromatine pour favoriser l'expression de l'ARNm auquel ils sont associés : ce sont les **ARN**

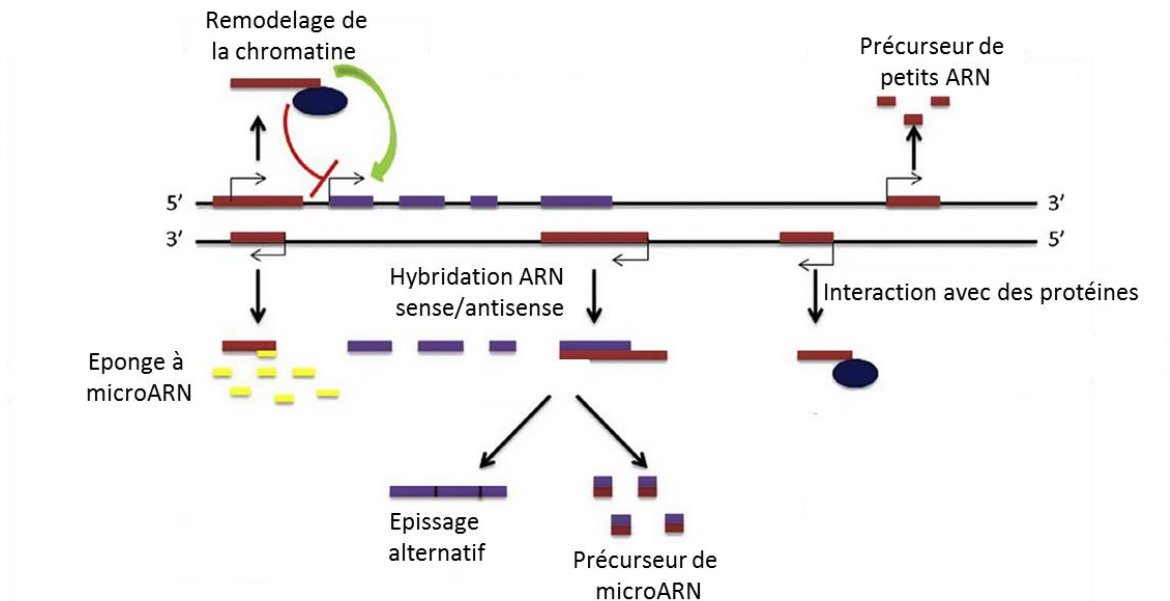


Figure 8. Représentation des principaux modes d'action des longs ARN non codants.

Dans le noyau, les longs ARN non codants peuvent réguler l'expression des gènes par des modifications épigénétiques (*HOTAIR*, *Xist*). Ils peuvent réguler l'épissage des ARNm en formant des sous-domaines nucléaires (*MALAT1*). Dans le cytoplasme, les longs ARN non codants peuvent être précurseurs de petits ARN (siRNA, miRNA) (*MALAT1*, *Xist*). Ils peuvent empêcher la fonction des microARN en piégeant ces derniers (*linc-RoR*). Enfin, ils peuvent fixer des protéines pour en modifier leur localisation cellulaire et/ou leur fonction (*gadd7*, *MALAT1*). (adapté de Shi et al., 2013).

enhancers (eRNA) (fig 7-4). Ils sont parfois considérés comme étant un bruit de fond transcriptionnel généré par la fixation du complexe d'initiation de la transcription sur les enhancers ou à proximité des promoteurs.

Ces différents lncRNA exercent des fonctions cellulaires variées.

II. Modes d'action des longs ARN non codants

En fonction de leur localisation cellulaire, les lncRNA ont différentes fonctions. Ils peuvent être impliqués dans la formation de compartiment cellulaire, dans la régulation de l'expression des gènes ou encore dans la régulation de l'activité des protéines (fig 8).

1. Modes d'action nucléaire

a. Formation de sous-domaines nucléaires

Les lncRNA peuvent réguler la formation de sous-domaines nucléaires appelés *paraspeckles*. Ces structures sont composées de protéines impliquées dans le transport des ARNm, de facteurs d'épissage, d'ARNm en cours de maturation et d'ARN non codants. Plusieurs études ont montré que l'ARN non codant *NEAT1* (*nuclear enriched autosomal transcript 1*) est essentiel à la formation de ces domaines. En effet, dans les cellules souches embryonnaires, où *NEAT1* n'est pas exprimé, les *paraspeckles* ne sont pas formés (Chen & Carmichael, 2009). A l'inverse, dans les cellules en différenciation, une augmentation du nombre et de la taille des *paraspeckles* a été retrouvée associée à une augmentation de l'expression de *NEAT1* (Sunwoo et al., 2009). Par ailleurs, des expériences d'imagerie sur des cellules où l'expression de *NEAT1* est sous l'action d'un promoteur inductible ont montré que la formation des *paraspeckles* se produit uniquement lorsque cet ARN est transcrit (Mao et al., 2011). Un second ARN non codant *MALAT1/NEAT2* (*Metastasis-associated lung adenocarcinoma transcript 1*) est également localisé dans les *paraspeckles* ; cependant il n'est pas essentiel à l'intégrité de ce domaine nucléaire contrairement à *NEAT1* (Bernard et al, 2010 ; Tripathi et al., 2010 ; Nakagawa et al., 2012). Ce dernier est conservé entre les espèces, ce qui lui suggère un rôle important dans la cellule. En effet, il a été mis en évidence que *NEAT2* est capable de se fixer sur des facteurs d'épissage SR, notamment la protéine SF2/ASF, et d'en modifier leur phosphorylation. Cependant, bien que *MALAT1* soit associé à un mauvais pronostic dans les cancers, des souris déficientes en *MALAT1* sont viables et ne possèdent pas de retard de croissance (Eißmann et al., 2012) ; ceci suggère la présence d'autres ARN assurant les fonctions de *MALAT1* dans la cellule.

A

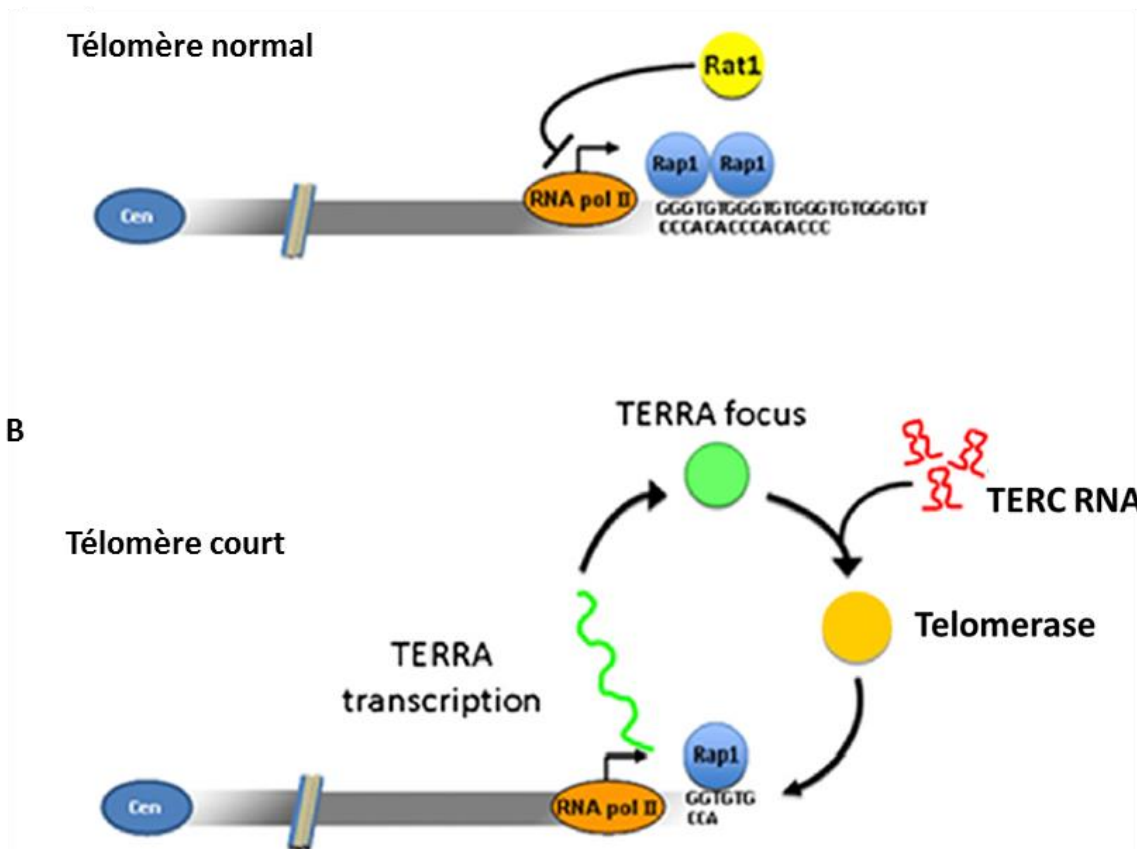


Figure 9. Stabilité chromosomique. (A) Sur un télomère normal, l'extrémité chromosomique, constituée du motif répété UUAGGG, est assez longue pour être reconnue par les protéines du complexe *Shelterin*, notamment Rap1. Ce complexe, en se fixant sur ces extrémités, permet la formation de la boucle T et empêche l'expression de l'ARN non codant *TERRA*. (B) Lorsque les télomères sont courts, les protéines du complexe *Shelterin* ne sont plus associées aux motifs répétés. La boucle T ne se forme plus et l'ARN non codant *TERRA* est exprimé. Cet ARN non codant recrute alors l'ARN *TERC* et la télomérase TERT. L'unité minimale catabolique de la télomérase est ainsi constituée au niveau des télomères courts. Elle permettra l'ajout de motifs UUAGGG en tandem à ces extrémités, ce qui augmentera leur taille et les protégera de l'action d'exonucléases. (adapté de Cusanelli et al., 2013).

b. Stabilité chromosomique

Les longs ARN non codants ont un rôle majeur dans la stabilité chromosomique. L'extrémité des chromosomes est composée d'un motif répété UUAGGG. Ce motif est reconnu par un certain nombre de protéines : TRF1, TRF2, TIN2, RAP1, TPP1 et POT1. Elles forment le complexe « *Shelterin* » qui permettra la formation de la boucle T à chaque extrémité chromosomique (Artandi & DePinho, 2010). Cette boucle protège les chromosomes de l'action des exonucléases. A chaque cycle de réplication, une partie des télomères est éliminée, faute de matrice à répliquer. Lorsque les télomères sont courts, la boucle T ne peut plus se former, les extrémités chromosomiques sont alors reconnues comme des cassures doubles brins susceptibles d'être dégradées par les exonucléases. La télomerase reconnaît ces chromosomes courts, elle y ajoute un motif répété ; ce qui permet d'allonger les extrémités chromosomiques et ainsi de protéger les chromosomes de la dégradation. La télomerase est constituée de la protéine TERT, qui est une *Telomerase reverse transcriptase*, et de l'ARN non codant *TERC* (*Telomerase RNA component*), qui sert de matrice à la télomerase. Une fois associée au télomère, cette enzyme va ajouter ces séquences répétées aux extrémités qui seront reconnues par les protéines du complexe *Shelterin* et la boucle T se formera de nouveau.

Un second ARN non codant est impliqué dans le maintien des télomères : *TERRA* (*Telomeric Repeat containing RNA*). Cet ARN est transcrit par l'ARN polymérase II, il est rarement polyadénylé (7% des cas) et est nucléaire. *TERRA* est transcrit à partir d'une région riche en CpG subtélomérique des chromosomes, sa longueur peut varier de 100 à 9000 bases paires (bp). Cet ARN est transcrit uniquement lorsque les télomères sont courts (fig 9). Lorsqu'il est transcrit, il s'associe au télomère de son chromosome d'origine. Il a souvent été admis que *TERRA* était en compétition avec la télomerase pour se fixer sur les télomères (Artandi & DePinho, 2010; Feuerhahn et al., 2010). Récemment, Cusanelli et collaborateurs ont mis en évidence que *TERRA* était capable de s'associer à la télomerase (TERT) ainsi qu'à l'ARN *TERC* (Cusanelli et al., 2013). De plus, ils ont également observé que l'association *TERRA*/télomère permet une mise en conformation de l'extrémité chromosomique favorable à l'activité de la télomerase. Il semblerait donc que *TERRA* possède un rôle d'ARN « *scaffold* » et que son association avec la télomerase favorise l'activité de cette dernière. Lorsque la télomerase est recrutée au niveau des télomères courts, elle va ajouter plusieurs fois le motif UUAGGG à l'extrémité du chromosome. Elle ajoute en moyenne 200 nts mais elle peut en ajouter jusque 400. Ceci permettra au chromosome de reformer sa boucle T et ainsi de le protéger de la dégradation et donc de le stabiliser. Ce processus est retrouvé dans les cellules en prolifération, non différenciées ; c'est pourquoi, l'ensemble de ces études a été réalisé dans des levures.

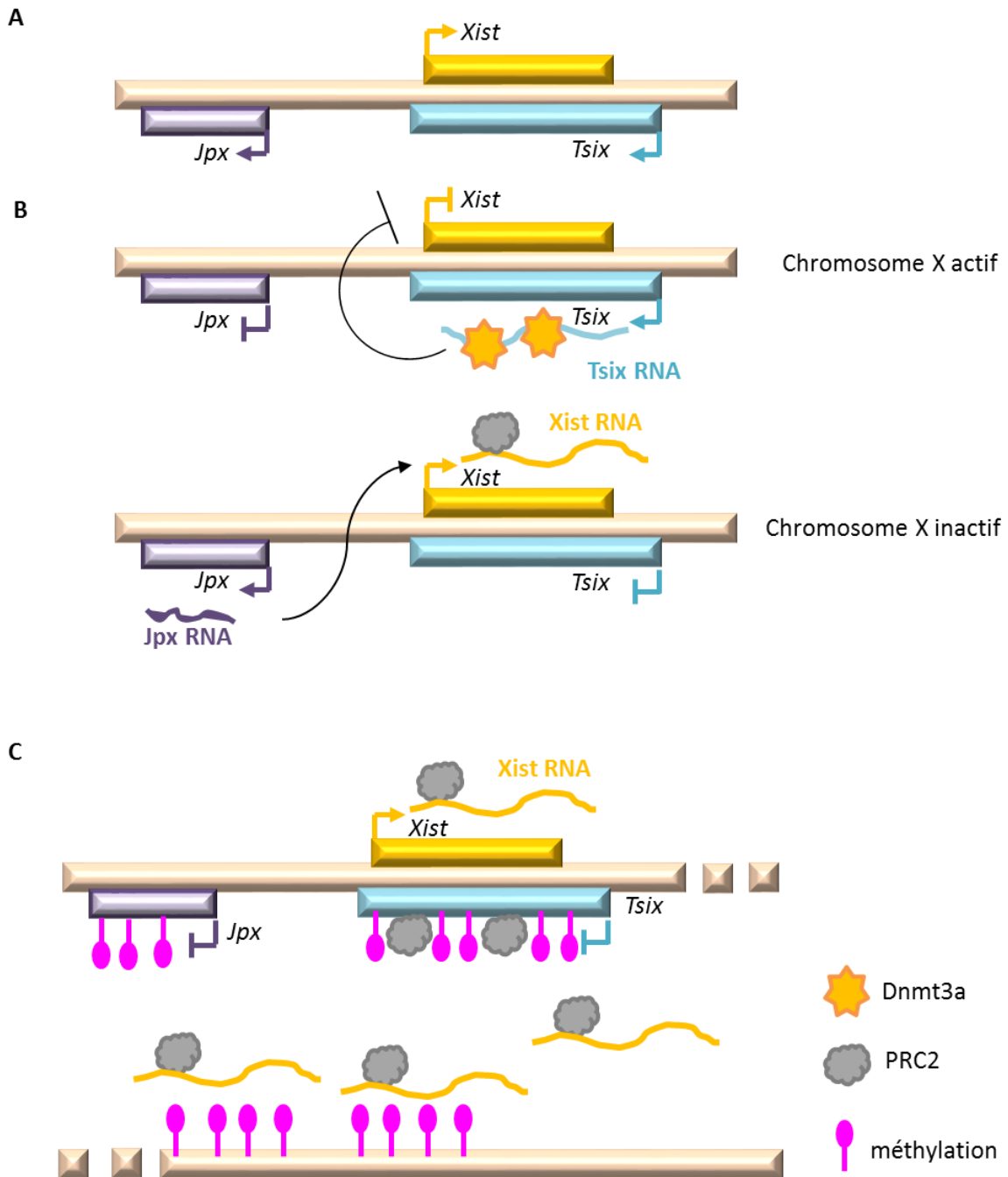


Figure 10. Mécanisme d'inactivation du chromosome X. (A) Composition du centre d'inactivation du chromosome X (Xic). Les gènes du Xic permettent l'expression de plusieurs ARN non codants (*Xist*, *Tsix*, *Jpx*) impliqués dans la régulation de l'inactivation du chromosome X chez les mammifères. (B) A partir du chromosome X actif, le gène *Tsix* s'exprime en un ARN non codant qui recrute Dnmt3a. Cette enzyme est responsable de la méthylation du promoteur de *Xist* sur cet allèle et par conséquent empêche l'expression de ce gène. Ce chromosome restera donc actif dans la cellule. Sur le chromosome inactif, l'expression de l'ARN non codant *Jpx* favorise l'expression de *Xist*. Cet ARN va alors se fixer au complexe PRC2 *via* son motif RepA et induire l'inactivation du chromosome. (C) Lorsque *Xist* est exprimé, il recrute le complexe PRC2 et en s'associant avec son chromosome d'origine, il induit la méthylation du chromosome. Cette méthylation se réalise de proche en proche le long du chromosome X.

c. Régulation de l'expression des gènes

Les lncRNA peuvent réguler l'expression des gènes par des modifications épigénétiques. L'épigénétique est *l'étude des changements héréditaires dans la fonction des gènes, ayant lieu sans altération sur la séquence ADN*. Ils sont également capables de moduler l'activité du complexe d'initiation de la transcription.

L'inactivation du chromosome X

Au cours du développement embryonnaire, les modifications épigénétiques permettent, entres autres, l'inactivation de l'un des deux chromosomes X chez la Femme.

Il s'agit du processus épigénétique le plus décrit et le plus étudié. Elle est permise par l'expression différentielle de trois ARN non codants : *Xist*, *Tsix* et *Jpx*. L'ARN non codant *Xist* est l'un des premiers à avoir été découvert chez les mammifères (Brockdorff et al., 1992 ; Brown et al., 1992). Cet ARN non codant est responsable de l'inactivation de l'un des deux chromosomes X chez la femme. Il a une taille comprise entre 17 et 20kb et son expression est régulée par l'ARN non codant *Tsix* (Lee, 2012). Ces deux ARN non codants (*Xist* et *Tsix*) sont transcrits à partir du centre de régulation : le centre d'inactivation du chromosome X (*Xic*). Lorsque *Tsix* est exprimé, il recrute une ADN méthyltransférase (*Dnmt3a*) qui méthylera l'ADN et empêchera ainsi l'expression de *Xist* ; le chromosome sera donc actif (fig 10). A l'inverse, sur le second chromosome X, l'expression de l'ARN *Jpx* active l'expression de *Xist* *en trans*. Une fois transcrit, *Xist* recrute le complexe polycomb PRC2 par l'intermédiaire de motifs répétés RepA et s'associe à son chromosome d'origine. Sur celui-ci, le complexe PRC2 induira une triméthylation de la lysine 27 de l'histone H3 ; cette marque épigénétique sera responsable de l'inactivation de ce chromosome X (Brockdorff et al., 2013) (fig 10). L'inactivation de l'un des deux chromosomes X chez la Femme est donc un exemple de compensation génique. Ce mécanisme permet ainsi un équilibre de l'expression des gènes présents sur ce chromosome entre les individus.

A partir du chromosome X actif, une transcription associée au motif RepA a été retrouvée. Bien que le rôle de cette transcription ne soit pas encore clairement défini, il semblerait que l'ARN *RepA* fixe les protéines du complexe polycomb et empêche leur action sur ce chromosome.

Ces modifications épigénétiques ont également lieu lors de la mise en place de l'empreinte génomique. Un chapitre sera consacré à cette régulation d'expression génique (voir P. 73).

Régulation de l'expression des gènes HOX

Les modifications épigénétiques par les lncRNA peuvent aussi avoir lieu *en trans*. Le premier ARN non codant à réguler l'expression des gènes *en trans* identifié est l'ARN non codant *HOTAIR* (*HOX transcript antisense RNA*). Cet ARN est exprimé à partir du locus HOX C situé

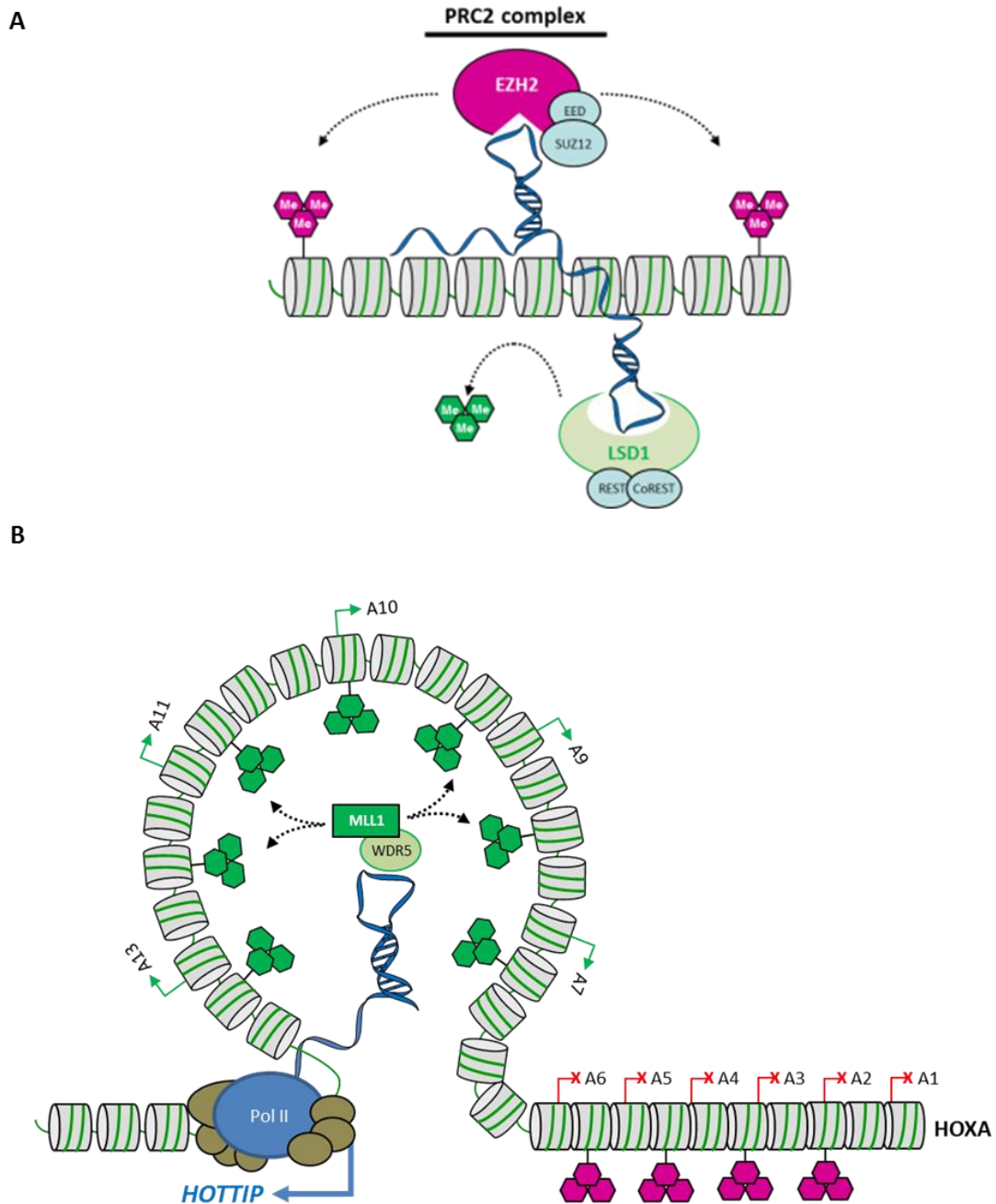


Figure 11. Mécanisme d'action des ARN non codants *HOTAIR* et *HOTTIP*. (A) Le lncRNA *HOTAIR*, exprimé à partir du locus HOX C, en s'associant à la protéine EZH2, permet le recrutement du complexe PRC2 au locus HOX D. *HOTAIR* se fixe également avec le complexe LSD1/CoREST. Ainsi *HOTAIR* permet la triméthylation de la lysine 27 et la déméthylation de la lysine 4 de l'histone H3 ; ce qui réprime l'expression des gènes du locus. (B) Lorsqu'il est exprimé, *HOTTIP* est capable de recruter la protéine MLL1 par l'intermédiaire de WDR5. La MLL1 (une histone méthyltransférase) va méthyler la lysine 4 de l'histone H3 : une marque activatrice de l'expression des gènes. Cette modification permettra l'expression des gènes HOX A. En vert : les marques activatrices de l'expression génique. En rose : les marques inhibitrices de l'activation génique. (Angrand et al., 2015).

sur le chromosome 12. Une fois exprimé, *HOTAIR* fixe la sous-unité catalytique EZH2 du complexe polycomb PRC2. Cette association va alors permettre la triméthylation de la lysine 27 de l'histone H3 sur le locus HOX D, situé sur le chromosome 2. Cette modification épigénétique étant une marque répressive de l'expression, les gènes du locus HOX D ne seront pas exprimés (Brockdorff, 2013 ; Bergmann & Spector 2014) (fig 11A). De plus, il a été mis en évidence que *HOTAIR* fixe la protéine LSD1, une protéine faisant partie du complexe répresseur REST. Cette protéine est une histone déméthylase, permettant la perte de méthylation de la lysine 4 de l'histone H3 (une modification activatrice de l'expression des gènes). Lorsque LSD1 est associée à *HOTAIR*, elle permet alors la perte de la marque H3K4me3 au niveau du locus HOX D ; ce qui contribue également à réprimer l'expression des gènes de ce locus (Tsai et al., 2010). Ainsi, *HOTAIR*, grâce à ses interactions protéiques, est une molécule chaperonne permettant l'inactivation de l'expression des gènes du locus HOX D.

A l'inverse, les lncRNA peuvent également favoriser l'expression des gènes. C'est le cas de l'ARN non codant *HOTTIP* (*HOXA transcript at the distal tip*) qui est transcrit à partir du locus HOX A. Cet ARN est capable de se fixer sur la protéine MLL (Mixed-lineage leukemia) via le motif répété WD (WDR5) (Kanhere & Jenner, 2012 ; Bergamm et al., 2014). Cette protéine est une histone H3 lysine 4 méthyltransférase. Cette modification épigénétique favorise l'expression des gènes. Par ailleurs, des expériences de capture de conformation de la chromatine ont montré que *HOTTIP*, en se fixant sur WRD5/MLL, permet une mise en conformation de la chromatine favorable à l'expression des gènes situés à distance de l'unité de transcription *HOTTIP* (Wang et al., 2011) (fig 11B). Cette conformation de la chromatine permet à la méthyltransférase (MLL1) associée à *HOTTIP* de méthyler l'H3K4. Les gènes présents dans la boucle seront alors exprimés. *HOTTIP* joue donc un rôle d'activateur de la transcription des gènes du locus HOX A.

LncRNA et le complexe d'initiation de la transcription

Dans les cellules, les eRNA ne sont ni épissés ni polyadénylés et sont souvent retrouvés dans les exosomes. Cependant, certains peuvent être épissés et polyadénylés, ce sont les elncRNA. Ces ARN non codants favorisent l'expression des gènes codants situés à proximité de leur loci d'expression. Dans les cellules cancéreuses mammaires (MCF-7), l'expression de plusieurs ARN enhancers (*TFF1*, *FOXC12* ou encore *CA12*) est régulée positivement par le 17- β -oestradiol (Li et al., 2013). Ils régulent l'expression en *cis* des gènes situés plus loin sur le chromosome en favorisant la formation d'une boucle d'interaction entre les enhancers et ces gènes. L'un des gènes cibles de ces eRNA activés par l'oestradiol est le gène codant la cohésine.

Par ailleurs, une autre étude a mis en évidence que l'eRNA transcrit à partir des enhancers du locus *MyoD1* favorise le recrutement de l'ARN polymérase II à ce locus et l'expression des gènes de ce site (Mousavi et al., 2013).

Cependant, dans certains cas, la transcription de l'ARN non codant empêche l'expression du gène codant situé en aval. C'est le cas du gène codant pour la déhydrofolate réductase (*DHFR*). Ce gène possède deux promoteurs : un promoteur dit majoritaire qui permet l'expression du gène et un promoteur minoritaire situé en amont. Dans les cellules en arrêt de cycle, un transcrit court, non codant, exprimé à partir du promoteur minoritaire a été identifié. Les auteurs ont également mis en évidence que cet ARN non codant s'associe au facteur de transcription TFIIB et empêche ainsi l'activation du promoteur majoritaire. Cependant, cet effet est labile puisqu'un ajout de sérum dans le milieu diminue l'expression de l'ARN non codant et favorise l'expression de la DHFR (Martianov et al., 2007).

d. Formation de triplex ARN-ADN•ADN

Les longs ARN non codants peuvent s'associer à la double hélice de l'ADN, ce qui forme un triplex ARN-ADN•ADN. Pour la formation de ce triplex, un appariement de type Hoogsteen est nécessaire. Cet appariement permet l'association d'une uridine avec une adénosine. Il est possible que la formation de ce triplex oligonucléotidique soit permise par la présence de protéines sur l'ARN comme c'est le cas pour le lncRNA *Xist*. En effet, lorsque *Xist* est exprimé, il se fixe sur son chromosome d'origine et induit l'inactivation de ce dernier. Des études ont montré que l'interaction *Xist*/chromosome X est permise par la protéine YY1 qui est un facteur de transcription en doigt de zinc (Jeon & Lee, 2011). Cette protéine, en se fixant sur *Xist* et sur le chromosome X, permet ainsi la formation du triplex et par conséquent l'inactivation du chromosome.

Cependant, dans d'autres cas, il se peut que l'association ARN/protéine empêche la formation du triplex. Par exemple, l'ARN non codant *Jpx* peut se fixer sur la protéine CTCF (CCCTC-binding factor) (Sun et al., 2013). Il semblerait que cette interaction empêche *Jpx* d'activer la transcription de *Xist*.

Le processus de formation mais surtout le rôle de ces triplex sont très controversés. D'autres études doivent être réalisées pour mieux les définir.

2. Modes d'action cytosolique

a. Précurseur de petits ARN

Dans le cytosol, les structures secondaires des longs ARN non codants peuvent être reconnues par des endonucléases spécifiques type Dicer. Ils seront alors clivés par cette enzyme, ce qui générera des petits ARN non codants type siARN, microARN ou encore ARN

de transfert dont les fonctions sont détaillées dans le chapitre suivant (voir P. 55). En effet, les longs ARN non codants peuvent être précurseurs de microARN. C'est le cas de l'ARN non codant *linc-MD1* qui génère deux microARN (miR-206 et miR-133b) dans les cellules musculaires différenciées et lors de dystrophie musculaire (Cesana et al., 2011).

Les longs ARN non codants peuvent également être précurseurs de siARN. Par exemple, les ARN *Xist* et *Tsix*, impliqués dans l'inactivation du chromosome X, sont capables de s'associer en duplex et d'être exportés dans le cytosol. Dans ce compartiment, ce duplex d'ARN est reconnu par l'enzyme Dicer qui va alors générer un siARN fonctionnel dans la cellule. Cet ARN interférant serait nécessaire à l'inactivation du chromosome X chez la femme (Ogawa et al., 2008).

Nous pouvons également citer l'ARN non codant *MALAT1* dont l'extrémité 3' est clivée par les RNase P et RNase Z. Ce clivage génère un ARN de transfert de 61 bp et l'ARN *MALAT1* d'environ 6700 bp qui est localisé dans les *paraspeckles* (Wilusz et al., 2008).

b. Eponge à ARN

Les lncRNA peuvent interagir avec des microARN pour les empêcher de se fixer sur leurs ARNm cibles : ce sont des éponges à microARN (*competitive endogenous RNA*, ceRNA). Les éponges à microARN peuvent se présenter sous différentes formes (fig 12). Nous pouvons citer par exemple, l'ARN non codant *RoR* (*Regulator of Reprogramming*) qui est impliqué dans le maintien des cellules souches, d'où son nom (Loewer et al., 2010). En effet, il a été mis en évidence que le *lincRNA-RoR* était capable de fixer le microARN-145. Ce microARN réprime l'expression des gènes *Sox2*, *Oct3/4* et *Nanog* ; protéine impliquée principalement dans la formation et le maintien des cellules souches (Wang et al., 2013). Ainsi le *lincRNA-RoR*, en fixant le miR-145, favorise l'apparition et le maintien des cellules souches.

Par ailleurs, le lncRNA *LINCMD1*, qui est spécifique des cellules musculaires, est capable de fixer les microARN miR-133 et miR-135 (Geisler & Coller, 2013). Ces deux microARN sont connus pour réprimer l'expression des ARNm *MALM1* (mastermind-like 1) et *MEF2C* (myocyte enhancer factor 2) : deux facteurs de transcription impliqués dans l'expression de gènes spécifiques de la cellule musculaire. Ainsi, de manière indirecte, le *LINCMD1* active l'expression de ces facteurs en se fixant sur les miR-133 et 135.

Récemment, des éponges à microARN plus stables ont été mises en évidence : ce sont les circular RNA (circRNA ou ciRNA). Ces deux formes d'éponges sont issues de l'épissage des exons. Les circRNA peuvent être constitués d'exons et d'introns alors que les ciRNA sont constitués uniquement d'introns (Zhang et al., 2013b). Comme les ceRNA, ces ARN non codants circulaires fixent des microARN et les empêchent d'agir dans la cellule. C'est le cas

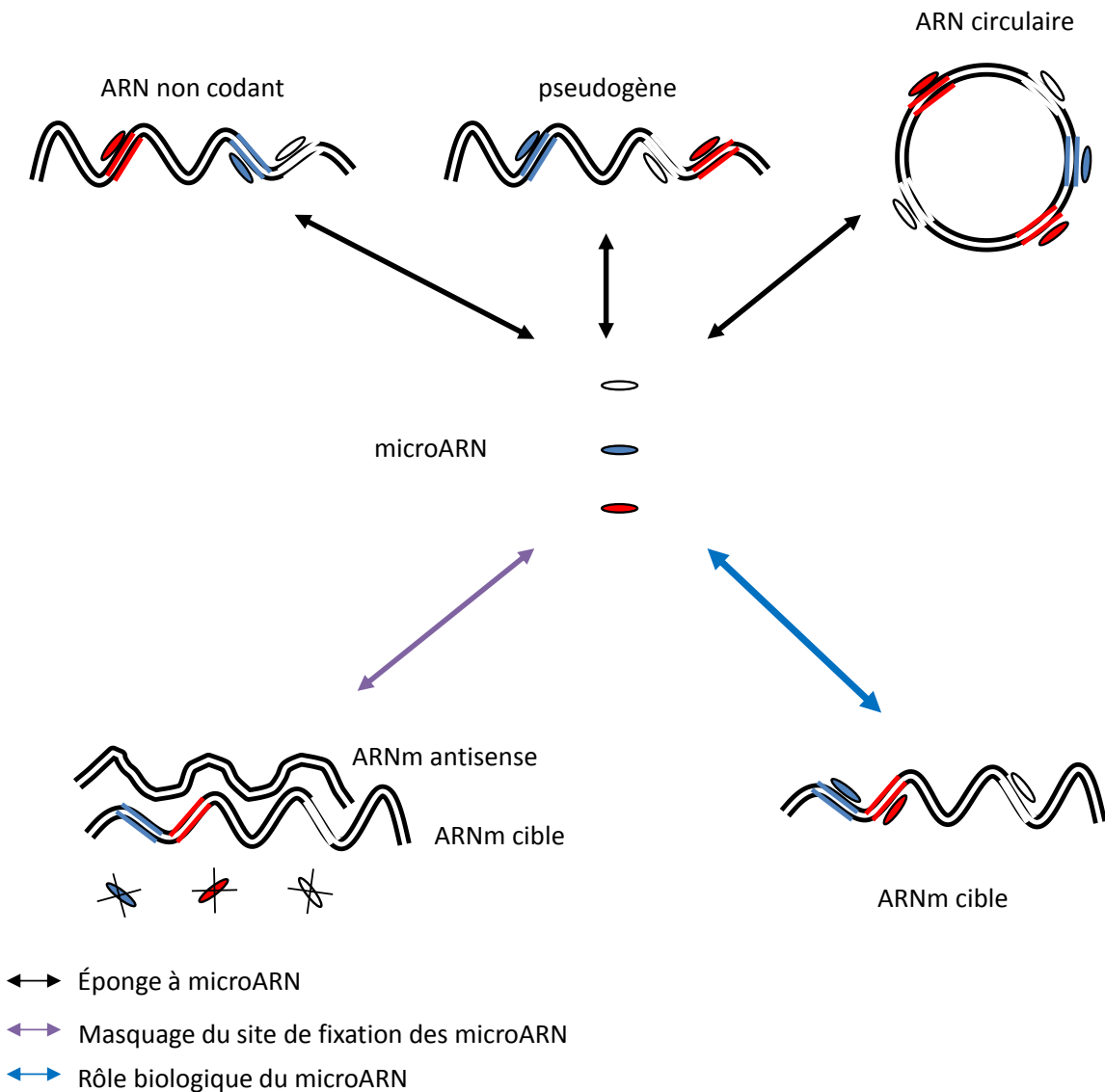


Figure 12. Modulation de l'effet des microARN par les longs ARN non codants. Les éponges à microARN peuvent se présenter sous différentes formes. Les longs ARN non codants, les pseudogènes et les ARN circulaires servent de leurre au microARN. Les ARN circulaires, issus de l'épissage alternatif des ARNm, sont plus stables que les autres types d'éponges connus. Les longs ARN non codants antisens peuvent interagir avec les ARNm transcrits à partir du même locus et par conséquent masquer le site de fixation du microARN. (adapté de Kartha & Subramanian, 2014).

par exemple, du circRNA CDR1as qui agit comme une éponge des miR-7 (Geisler & Coller, 2013) (fig 12).

Cependant, quelques différences entre ces deux types d'éponges sont à noter. La formation des ciRNA est permise par la présence d'un site consensus en 5' du site d'épissage ; tandis que la formation des circRNA est permise par l'absence de queue polyA sur l'ARNm.

De plus, les ciRNA sont plutôt nucléaires et régulent l'expression de leur propre gène alors que les circRNA sont cytosoliques et jouent un rôle d'éponge (Zhang et al., 2013b; Hansen et al., 2013; Memczak et al., 2013). En effet, le ciRNA intronique *ankrd52* régule positivement l'expression du gène auquel il est associé en se fixant sur la machinerie d'élongation de la transcription par l'ARN polymérase II.

L'ARN non codant *PTENP1* est produit à partir d'un pseudogène situé en position 9p13.3. Il possède une forte homologie de séquence avec l'ARNm *PTEN*. Ces deux ARN, *PTEN* et *PTENP1*, sont en compétition pour les mêmes sets de microARN (miR-17, miR-21, miR-214, miR-19 et miR-26) puisque des sites de fixation pour ces derniers sont présents sur les deux ARN (Poliseno et al., 2010) (fig 12). Une corrélation positive entre l'expression de *PTEN* et *PTENP1* est retrouvée dans une série de tumeur. Un faible niveau d'expression du pseudogène permet de restaurer l'expression de l'ARNm *PTEN* dans les tissus. L'ARN non codant *PTENP1*, grâce à sa fonction d'éponge à microARN, possède donc un rôle de suppresseur de tumeur.

Les lncRNA peuvent aussi masquer le site de fixation du microARN en interagissant avec l'ARNm. Ce mode d'action est surtout décrit pour les lncRNA synthétisés en orientation antisens de gènes codants : *BACE-AS* (*Alzheimer-associated β -secretase-1*) par exemple. Cet ARN non codant forme un ARN double brin lorsqu'il s'associe avec l'exon 6 de l'ARNm *BACE*. Lorsque ce duplex est formé, le site de fixation du miR-485-5p sur l'ARNm est masqué, le microARN ne peut plus interagir avec cette cible (fig 12). Par conséquent, une augmentation de protéine BACE est observée dans les cellules. Cette protéine étant une enzyme impliquée dans le clivage de la protéine pre-amyloïde, une synthèse plus élevée de peptide A β 1-42 a été mis en évidence lorsque *BACE-AS* est exprimé. Ce peptide s'agrège en plaques, nommées plaques d'amyloïdes, dans le cerveau. L'apparition de ces plaques est responsable de la maladie d'Alzheimer (Faghihi et al., 2008, 2010).

Ainsi, les lncRNA, en fixant des microARN ou en masquant leur site de fixation sur les ARNm cibles, altèrent la fonction des microARN.

c. Interaction lncRNA/protéine

Les lncRNA sont capables d'interagir avec les protéines et d'en moduler leur fonction. L'ARN non codant *gadd7*, qui est induit par la lumière UV, influence la stabilité des ARNm dans la cellule en se fixant sur la protéine TDP43 (*TAR DNA-binding protein 43*) (Geisler & Coller, 2013). Cette protéine est impliquée dans l'épissage alternatif des ARNm, dans le transport des ARNm du noyau au cytosol, dans la traduction des ARNm mais aussi dans le contrôle de leur stabilité. Elle est notamment impliquée dans le saut de l'exon 9 de l'ARNm CFTR, ce qui se traduit par un phénotype de fibrose kystique. L'inactivation de TDP43 permet de restaurer une protéine CFTR fonctionnelle (Ayala et al., 2006). Lorsqu'elle est associée à l'ARN non codant *gadd7*, sa fonction est réprimée. Un second ARN non codant, *MALAT1*, réprime également la fonction de cette protéine en la piégeant dans le noyau. Cibler TDP43 ou favoriser son association avec les ARN non codants *gadd7* et *MALAT1* peut avoir un intérêt thérapeutique chez les individus atteints de la mucoviscidose.

Dans ce précédent cas, l'association lncRNA-protéine réprime l'activité de la protéine dans la cellule. Dans d'autres cas, l'association lncRNA-protéine peut favoriser l'action de la protéine. En effet, l'association de l'ARN non codant *SRP* (signal recognition particule) avec 6 protéines permet la formation d'un complexe ribonucléoprotéique (RNP) qui est nécessaire à la mise en conformation des protéines lors de leur synthèse. L'ARN *SRP* reconnaît le peptide naissant dans le cytosol de la cellule et sa présence dans le RNP permet au peptide naissant d'être conduit au réticulum endoplasmique. Ainsi, la protéine en cours de synthèse pourra être formée correctement et avoir les bonnes conformations et modifications post-traductionnelles (Grotwinkel et al., 2014).

Les longs ARN non codants peuvent aussi s'associer aux protéines du complexe polycomb (EZH2 notamment). Cette association permet alors de cibler ce complexe sur le promoteur de certains gènes et de réguler l'expression de ces gènes par des modifications épigénétiques (voir P. 43).

Un long ARN non codant peut avoir plusieurs modes d'actions dans la cellule. C'est le cas par exemple du *lincRNA-RoR*. Cet ARN non codant peut être une éponge à ARN mais il peut également se fixer sur la protéine P53 et altérer sa fonction (Zhang et al., 2013a).

La fonction des longs ARN non codants dans la cellule a fait l'objet de deux revues dans des journaux soumis à comité de lecture. (Annexe 1 et Annexe 2) (voir P. 317).

III. Les ARN non codants dans la glande mammaire et sa tumorigenèse

Comme nous l'avons décrit précédemment, la glande mammaire est une structure en constante évolution chez la femme. En effet, au rythme des grossesses, les alvéoles se développent et atteignent leur maximum de différenciation lors de la lactation. Lorsque la période de sevrage est terminée, une involution de la glande mammaire est constatée, avec principalement la régression des alvéoles. Ce potentiel de croissance et de différenciation est notamment permis par les cellules souches présentes dans la glande mammaire. Plusieurs études microarray ont mis en évidence l'expression d'ARN non codant dans la glande mammaire de souris (Shore & Rosen, 2014). De plus, dans les tissus néoplasiques, une expression aberrante de ces ARN non codants a été constatée. Bien que peu étudiés, les ARN non codants semblent donc être impliqués dans la formation du tissu mais aussi dans sa cancérisation.

1. ARN non codant et développement de la glande mammaire

La glande mammaire est en constante évolution : pendant la grossesse, la prolifération des acini est maximale ; à la lactation, la différenciation des bourgeons terminaux permet la production de lait. Plusieurs études ont révélé l'implication des ARN non codants dans la prolifération et la différenciation alvéolaire au cours de la grossesse. Par exemple, l'ARN non codant *SRA1* (*Steroid receptor RNA activator 1*) permet la prolifération cellulaire, le développement du réseau de canaux galactophores et la différenciation des cellules alvéolaires (Lanz et al., 1999, 2003). Cet ARN non codant est également impliqué dans la différenciation cellulaire dans d'autres tissus (le muscle par exemple) (Caretto et al., 2006 ; Xu et al., 2010a).

Deux autres ARN non codants sont principalement impliqués dans le contrôle de la différenciation des alvéoles. En effet, l'ARN non codant *PINC* (*Pregnancy induced non-coding RNA*) est surexprimé lors de la grossesse dans les cellules alvéolaires (Hansji et al., 2014). A la naissance, son niveau d'expression diminue, permettant ainsi la différenciation des cellules épithéliales et la production de lait (Ginger et al., 2001 ; Askarian-Amiri et al., 2011). Lors de la phase d'involution, son expression augmente à nouveau. Ainsi, cet ARN non codant permet de maintenir les cellules alvéolaires dans un état indifférencié et par conséquent d'empêcher la lactation.

L'ARN non codant *Zfas1* se comporte de façon similaire à l'ARN *PINC*. Il est exprimé dans les cellules alvéolaires lors de la grossesse et l'involution et son expression diminue à la lactation (Shore & Rosen, 2014).

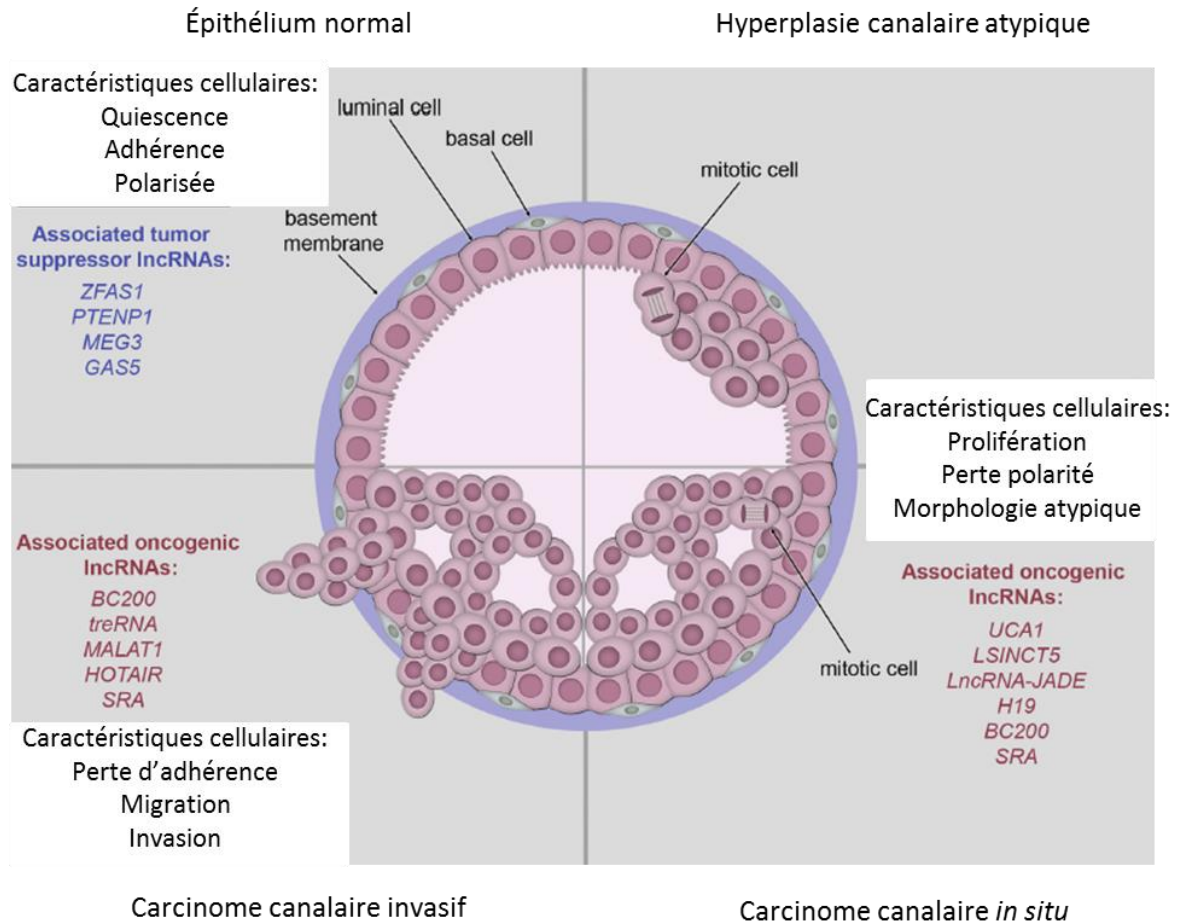


Figure 13. Implication des ARN non codants dans la glande mammaire normale et cancéreuse. Dans la glande mammaire normale, plusieurs ARN non codants agissent comme gène suppresseur de tumeur (*ZFAS1*, *PTENP1*, *MEG3*, *GAS5*). Lors de la tumorigenèse, l'expression de certains ARN non codants favorise la progression tumorale et augmente l'agressivité de la tumeur (*UCA*, *LSINCT5*, *H19*, *SRA*, *HOTAIR*, *treRNA*). (Adapté de Shore & Rosen, 2014).

Enfin, notre équipe a mis en évidence que l'ARN non codant *H19* est également impliqué dans la prolifération des cellules alvéolaires lors de la grossesse (Adriaenssens et al., 1999). Le rôle de cet ARN dans la glande mammaire sera décrit plus tard dans le manuscrit (voir P. 73).

2. ARN non codant et cancer du sein

Dans les cancers du sein, une expression aberrante de plusieurs ARN non codants a été constatée. Comme pour les ARNm, les ARN non codants peuvent être des oncogènes ou des gènes suppresseurs de tumeur. Dans ce paragraphe, quelques exemples de lncRNA impliqués dans la tumorigenèse de la glande mammaire seront décrits.

a. Les ARN non codants suppresseurs de tumeur

L'expression de l'ARN non codant *MEG3* (Maternally expressed gene 3) dans les tumeurs mammaires est diminuée, laissant penser que ce dernier possède une fonction de gène suppresseur de tumeur. En effet, une surexpression ectopique de *MEG3* permet une diminution de prolifération, d'angiogenèse et une augmentation de l'apoptose des cellules cancéreuses mammaires (Zhou et al., 2007; Gordon et al., 2010).

Dans les cancers du sein, l'expression d'un second ARN non codant est diminuée : *GAS5* (*growth arrest-specific transcript 5*). Cet ARN non codant se fixe sur les récepteurs aux glucocorticoïdes, par l'intermédiaire de deux sites de fixation en forme de tige-boucle ; ce qui empêche l'interaction de ces derniers avec leurs cibles usuelles et par conséquent empêche leur activation (Kino et al., 2010). Ces récepteurs sont impliqués dans le contrôle de la croissance cellulaire, de la survie et du métabolisme. Lorsque *GAS5* est surexprimé dans des cellules cancéreuses mammaires, une diminution de prolifération cellulaire ainsi qu'une augmentation de l'apoptose ont été constatées (Mourtada-Maarabouni et al., 2009).

Dans les cellules épithéliales mammaires, l'ARN non codant *Zfas1* empêche la prolifération et la différenciation des cellules, ce qui lui suggère un rôle de suppresseur de tumeur (Askarian-Amiri et al., 2011 ; Ginger et al., 2006). De plus, dans les carcinomes ductaux de la glande mammaire, l'expression de *Zfas1* est diminuée. Ces données, associées au paragraphe précédent, semblent confirmer son rôle de gène suppresseur de tumeur dans la tumorigenèse de la glande mammaire.

L'expression de ces ARN non codants dans les cellules de la glande mammaire protège celle-ci de la tumorigenèse (fig 13).

b. Les ARN non codants oncogènes

Les ARN non codants oncogènes sont plus décrits que les suppresseurs de tumeur car ils induisent une hausse de la prolifération cellulaire ainsi qu'une hausse de la formation de

Introduction

tumeur. La majorité de ces ARN non codants oncogènes est impliquée dans la prolifération et la migration/invasion cellulaire. Ils sont donc associés aux carcinomes *in situ* et invasifs (fig 13).

Par exemple, l'expression de l'ARN non codant *UCA1* (*urothelial carcinoma-associated 1*) est augmentée dans plusieurs cancers : cancer de la vessie et cancer du sein. Cet ARN non codant se fixe sur la protéine hnRNP et empêche la fonction de cette dernière ; ainsi la protéine P27 est moins exprimée. Dans les cellules cancéreuses mammaires, une augmentation de l'expression d'*UCA1* favorise la prolifération cellulaire. A l'inverse des cellules invalidées pour l'ARN *UCA1* prolifèrent moins et forment des tumeurs plus petites lorsqu'elles sont injectées en souris immunodéficientes (Huang et al., 2014; Wang et al., 2006). Par ailleurs, dans les carcinomes de cellules squameuses, il semblerait que cet ARN soit impliqué dans la résistance à la doxorubicine (Tsang et al., 2007). Dans les cancers du sein, cette fonction n'a pas encore été décrite.

L'ARN non codant *HOTAIR* est surexprimé dans les cancers du côlon, du foie et du sein (Hansji et al., 2014). Cet ARN non codant, en s'associant avec les protéines du complexe répresseur PRC2, permet la reprogrammation cellulaire et par conséquent la transformation des cellules normales en cellules cancéreuses. Ainsi, les cellules surexprimant *HOTAIR* ont des capacités de prolifération, de migration et d'invasion qui sont augmentées.

D'autres ARN non codants, impliqués dans l'intégrité du noyau, sont également des oncogènes dans les cancers : *MALAT1* dans les cancers du poumon et du sein par exemple ; *LINSCT5* (*long stress-induced noncoding transcript 5*) dans les cancers ovariens et du sein. Ces deux ARN non codants permettent la formation et la maintenance des *paraspeckles* (sous-domaines nucléaires). Des cellules cancéreuses mammaires invalidées pour l'ARN *LINSCT5* ont des capacités de prolifération diminuées (Silva et al., 2011). De même, lorsque *MALAT1* est réprimé, un défaut d'épissage de certains transcrits est constaté (Lin et al., 2011). Par ailleurs, une perte d'expression de *MALAT1* dans les cellules est corrélée à une augmentation de l'expression des gènes voisins de cet ARN non codant (Zhang et al., 2012).

Certains ARN non codants vont favoriser la transition épithélio-mésenchymateuse (EMT) dans les cancers. Ces ARN seront alors impliqués dans les processus de migration et d'invasion cellulaires ; ils seront donc associés aux métastases. Nous pouvons citer l'ARN non codant *treRNA* (*translational regulatory RNA*). Cet ARN favorise l'expression du gène *Snail* : un régulateur de l'EMT (Orom et al., 2010). De plus, il est capable de réprimer l'expression de la E-cadhérine, ainsi que d'autres marqueurs de cellules épithéliales (Gumireddy et al., 2013). Ces effets moléculaires sont associés à une hausse de la migration, de l'invasion cellulaire

mais aussi du taux de métastases pulmonaires après injection en souris immunodéficientes. Cet ARN non codant est donc associé aux capacités métastatiques des tumeurs.

Enfin, dans les cancers du sein, nous avons mis en évidence que l'ARN *H19* possède un rôle d'oncogène. En effet, cet ARN favorise la transition G1/S du cycle cellulaire dans plusieurs lignées cellulaires cancéreuses mammaires (Berteaux et al., 2005). Le rôle de l'ARN *H19* dans la tumorigenèse sera décrit plus tard dans le manuscrit (voir P. 73).

Les ARN non codants peuvent également induire la résistance des tumeurs aux différents traitements. C'est le cas de l'ARN non codant *ARA* (*Adriamycin resistance associated*) qui est exprimé dans les tumeurs traitées à la doxorubicine (Jiang et al., 2014). Cet ARN non codant régule l'activation de plusieurs voies de signalisation, notamment les voies MAP kinase et d'adhésion focale. Ainsi, lorsqu'il est surexprimé, cet ARN favorise la prolifération et la migration cellulaire, il augmente donc les capacités métastatiques des cellules tumorales.

En conclusion, les ARN non codants sont impliqués dans le développement de la glande mammaire, dans sa cancérisation mais aussi dans les phénomènes d'échappement tumoral aux traitements. Il est donc nécessaire d'identifier les différents mécanismes mis en place par ces derniers pour induire la progression tumorale et/ou les résistances thérapeutiques afin d'identifier de nouvelles cibles thérapeutiques.

Les microARN

Les petits ARN non codants sont divisés en plusieurs groupes : les ARN de transfert (tRNA), les microARN (miRNA), les siARN ou encore les ARN se fixant sur les protéines PIWI (piRNA). Ce chapitre est principalement consacré à l'étude des microARN : leur découverte, leur synthèse et leur fonction dans la cellule.

I. Découverte des microARN

1. Découverte des microARN

Le premier microARN, découvert en 1993 chez le ver nématode *Caenorhabditis elegans*, est le microARN *lin-4* ; il est généré à partir du gène de même nom (Lee et al., 1993). Les auteurs ont montré que l'expression du miRNA *lin-4* empêche la traduction de l'ARNm *lin-14*, un ARNm codant pour une protéine nucléaire. A l'époque, ils pensaient que le microARN *lin-4* était capable d'interagir avec l'ARNm *lin-14* grâce à une complémentarité de séquence avec l'UTR3' de cet ARNm. Cette interaction serait responsable de l'inhibition de l'expression de *lin-14* et permettrait le bon développement de la larve. En effet, à la fin de la première étape du développement de la larve, une diminution d'expression de *lin-14* permet d'initier la seconde étape du développement de la larve (Wightman et al., 1993; Olsen et al., 1999).

Depuis leur découverte, le nombre de microARN identifié n'a cessé d'augmenter. En effet, en 2014, la base de données miRBase répertoriait 2588 microARN chez l'Homme, 466 microARN chez *Drosophila melanogaster* et 434 microARN chez *Caenorhabditis elegans* (Ha & Kim, 2014). Les microARN sont donc les petits ARN les plus abondants dans une cellule somatique.

2. Nomenclature

A leur découverte, les microARN étaient nommés en fonction de leur phénotype associé : *lin-4* ou *let-7* par exemple. Les microARN qui ont ensuite été identifiés par clonage ou par séquençage ont reçu un numéro : miR-21, miR-125. Comme pour les protéines, il existe des familles de microARN. Tous les microARN appartenant à une même famille possèdent la même *seed sequence* (séquence minimale d'interaction avec son ARNm cible); ils sont différenciés par l'ajout d'une lettre : par exemple miR-125a et miR-125b. Les microARN appartenant à la même famille sont redondants, c'est-à-dire qu'ils possèdent les mêmes ARNm cibles. En 2014, 196 familles de microARN, conservées entre les mammifères, étaient identifiées ; ceci reflète de leur importance au cours de l'évolution.

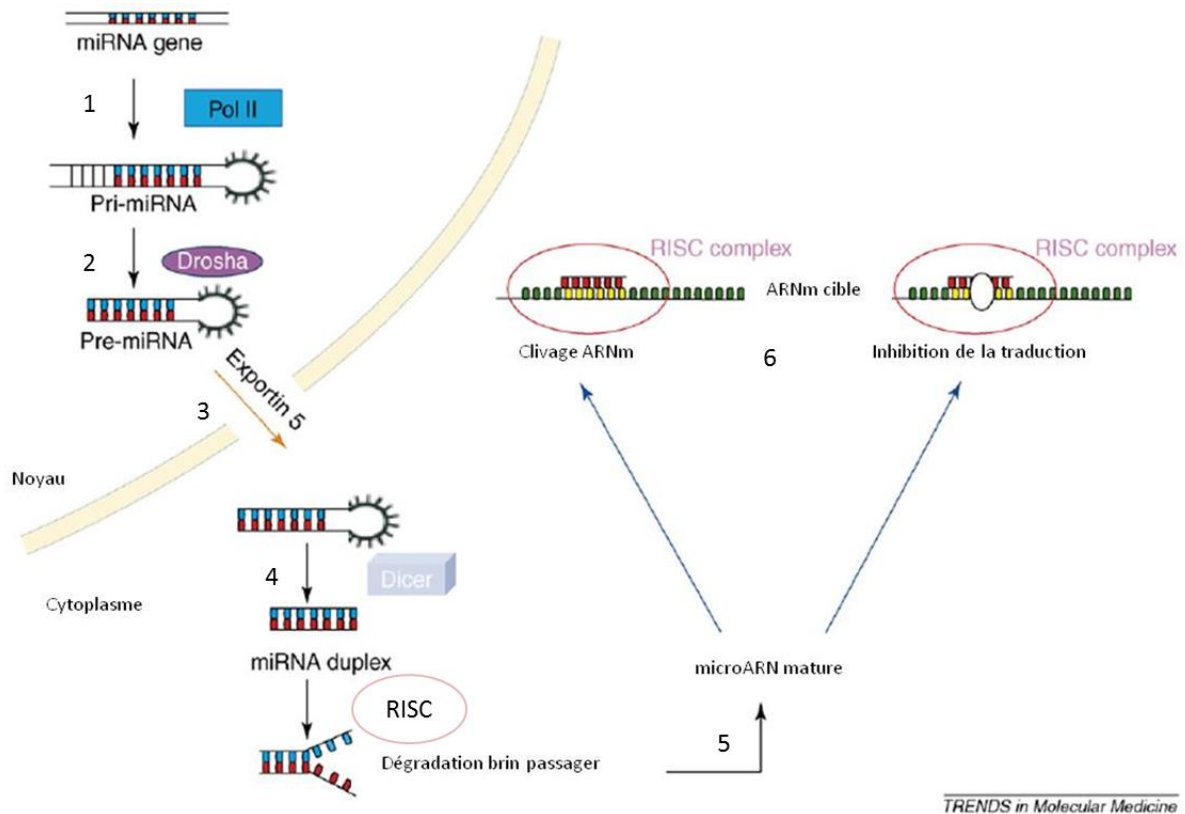


Figure 14. Synthèse des microARN. (1) Les microARN sont transcrits par l'ARN polymérase II sous forme d'un pri-miRNA. (2) Le pri-miRNA est clivé par une première enzyme : Drosha. (3) Obtention d'un pré-miRNA qui est exporté dans le cytosol grâce à l'exportine 5. (4) Dans le cytosol, le pré-miRNA est clivé par Dicer : obtention d'un duplex de microARN qui est pris en charge par le complexe RISC (*RNA-induced silencing complex*). (5) Dans ce complexe, le brin passager du duplex est dégradé. (6) Le microARN, alors mature, se fixera sur ses ARNm cibles pour induire leur clivage ou l'inhibition de leur traduction. (adapté de Garzon et al., 2006).

Un même microARN peut être transcrit à partir de plusieurs loci différents. Ils sont, dans ce cas, différenciés par l'ajout d'un chiffre : miR-125b-1, miR-125b-2. Ils posséderont les mêmes ARNm cibles.

Généralement, un même transcrit peut être précurseur de plusieurs microARN ; il s'agit d'une unité polycistronique. Dans cette unité, les microARN n'appartiennent pas à la même famille, ils ne possèdent pas la même *seed sequence*. L'unité polycistronique la plus conservée est le cluster impliqué dans le développement bilatéral des animaux : miR-100/let-7/miR-125 (Ha & Kim, 2014).

Les gènes permettant la formation des microARN sont notés : mir-xxx tandis que les microARN matures sont notés : miR-xxx (Tetreault & DeGuire, 2013). Les microARN matures sont sous forme de duplex d'ARN. Dans ce duplex, le premier brin d'ARN est appelé brin guide ou miR-xx-5p et le second brin, brin passager ou miR-xx-3p ou miR*. Au cours de la maturation du microARN, le brin passager est souvent dégradé. Dans de rares cas, ce brin miR-xx-3p peut être fonctionnel dans la cellule et avoir des ARNm cibles.

En 2006, il était admis qu'environ 30% des gènes codants pour une protéine pouvaient être régulés par un microARN (Garzon et al., 2006). En 2014, il était estimé qu'environ 60% des gènes codants pour des protéines ont au minimum un site de fixation potentiel à un microARN (Ha & Kim, 2014). Un microARN peut réguler l'expression de plusieurs ARNm. L'expression d'un ARNm peut être régulée par plusieurs microARN différents. Pour la majorité d'entre eux, leur fonction dans la cellule doit encore être identifiée.

II. Les voies de synthèse des microARN

A ce jour, deux voies de synthèse sont à l'origine de la formation des microARN : une voie majoritaire dite canonique (fig 14) et une voie minoritaire dite non-canonique.

1. La voie canonique

a. Formation du pré-miRNA

Les microARN sont transcrits par l'ARN polymérase II sous forme d'un pri-miRNA. Il possède une taille d'environ 1 kilobase (kb) et contient une structure type tige-boucle où la tige possède 30~35 nts. Généralement, le pri-miRNA est porté par les introns des ARNm, les ARN non codants, ou plus rarement, les exons. Lorsqu'il est transcrit, la tige-boucle du pri-miRNA est reconnue par la protéine Drosha (RNase III), qui en association avec une protéine stabilisatrice : DGCR8, va former le complexe microprocesseur (Denli et al., 2004; Gregory et al., 2004 ; Han et al., 2004). Ce complexe va cliver le pri-miRNA pour générer le pré-miRNA. Drosha est une endonucléase, de 160 kDa, spécifique des ARN doubles brins.

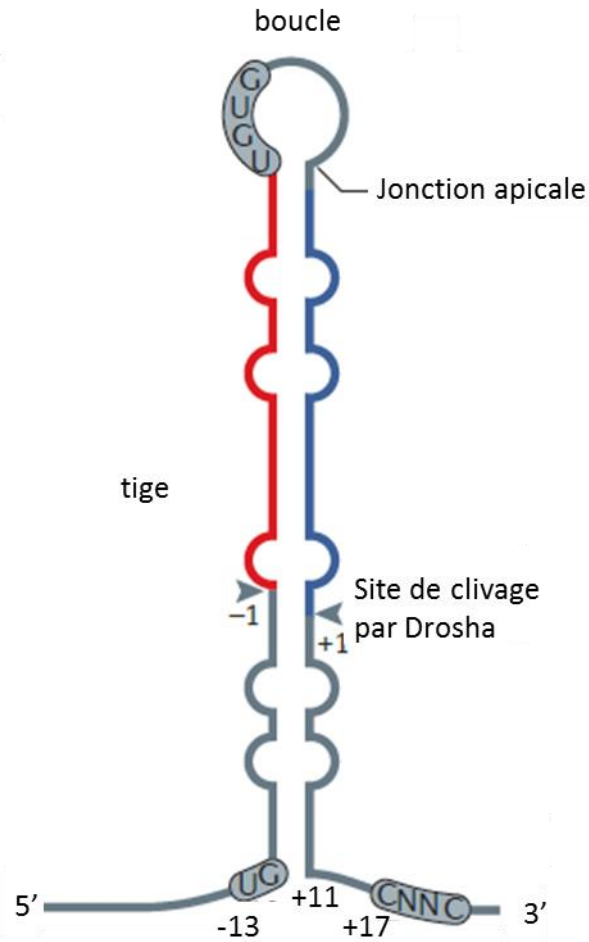


Figure 15. Schéma représentant le pri-miRNA. Sur le pri-miRNA, les séquences UG et CNNC en 5' et 3' de la tige sont reconnues par le complexe microprocesseur constitué de l'enzyme Drosha et son cofacteur DGCR8. Drosha clive alors la tige à 22 nts de l'extrémité apicale. Il génère ainsi le pré-miRNA avec une extrémité 5' phosphorylée et une extrémité 3' débordante de 2 nts. (adapté de Ha & Kim, 2014).

Elle possède trois domaines importants :

- Une partie Nt qui lui permet une localisation nucléaire.
- Une partie intermédiaire qui permet le recrutement de la protéine DGCR8.
- Une partie Ct qui contient un domaine RNase III (RIIID) en tandem et un domaine de liaison aux ARN doubles brins.

Les domaines RIIIDa et RIIIDb constituent la partie catalytique de l'enzyme. Le domaine RIIIDa va cliver l'ARN double à l'extrémité 3' tandis que le domaine RIIIDb va cliver l'ARN doubles brins à l'extrémité 5'. Drosha va générer une extrémité 3' débordante de 2 nts (Blaszczyk et al., 2001 ; Han, et al., 2004 ; Zhang et al., 2004a).

La protéine DGCR8 est un cofacteur essentiel de la synthèse des microARN qui est conservée entre les espèces (Pasha chez *D. melanogaster* ; PASH-1 chez *C. elegans*). Cette protéine se trouve dans le nucléoplasme et possède plusieurs domaines importants pour sa fonction (Yeom et al., 2006 ; Han et al., 2006 ; Sohn et al., 2007):

- Un signal de localisation nucléaire en partie Nt
- Un domaine de dimérisation et de fixation de l'hème en partie intermédiaire
- Un domaine de liaison à la protéine Drosha en partie Ct

Elle possède également deux domaines de liaisons aux ARN doubles brins, ce qui lui permet de reconnaître les pri-miRNA (Yeom et al., 2006).

Un *Knock-out* (KO) pour Drosha ou DGCR8 est léthal au stade embryonnaire, ce qui reflète du rôle essentiel de ces protéines et par conséquent des microARN au cours du développement embryonnaire (Wang et al., 2007b ; Chong et al., 2010). Il a récemment été mis en évidence que des motifs UG et CNNC, respectivement en Nt et Ct sur le pri-miRNA, permettent au complexe microprocesseur de reconnaître ce dernier (fig 15) (Auyeung et al., 2013). Le facteur d'épissage SRp20 (aussi connu sous le nom SRSF3), qui reconnaît également le motif CNNC, favorise la fixation de Drosha sur le pri-miRNA.

Lorsque le complexe microprocesseur est associé au pri-miRNA, Drosha va cliver l'ARN doubles brins à 22 nts de la jonction apicale (fig 15) en générant une extrémité 5' phosphorylé et une extrémité 3' débordante de 2 nts (Zeng et al., 2005 ; Han et al., 2006). Le pré-miRNA sera ensuite reconnu par un certain nombre de protéine et exporté dans le cytosol.

La synthèse du pré-miRNA n'affecte pas l'épissage de l'ARNm puisqu'elle a lieu avant celui-ci (Kim & Kim, 2007). De plus, les microARN pouvant être portés par les introns, le pri-miRNA est reconnu lorsque l'ARNm est épissé. La formation du microARN peut affecter la stabilité

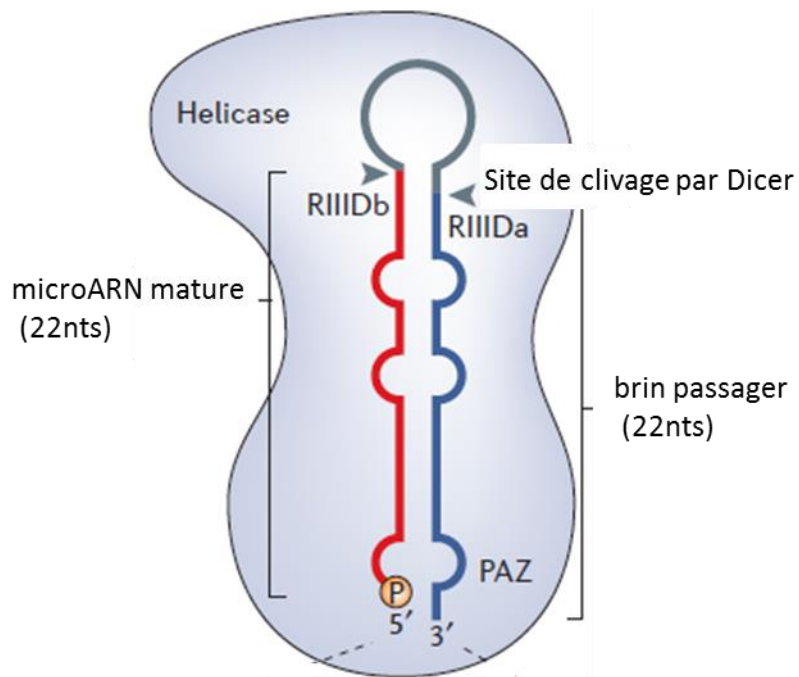


Figure 16. Schéma représentant le pré-miRNA. Les extrémités 5' et 3' du pré-miRNA sont reconnues par l'enzyme Dicer et son cofacteur TRBP. Elle va alors cliver la boucle du pré-miRNA pour former un ARN doubles brins de 22 nts. Ce duplex forme le microARN immature constitué du brin guide (rouge) et du brin passager (bleu). (adapté de Ha & Kim, 2014).

d'un ARNm si celui-ci est contenu dans l'exon de l'ARNm. Dans ce cas, le clivage de l'ARNm par Drosha induit la dégradation de l'ARNm. Ce phénomène se produit lorsque le miR-198 est synthétisé puisqu'il est contenu dans l'UTR3' de *FSTL1* (Sundaram et al., 2013).

b. Export du pré-miRNA dans le cytosol

L'export du pré-miRNA dans le cytosol est permis par l'exportine 5 : EXP5 (protéine codée par le gène *XPO5*). L'exportine 5 va former un complexe avec la protéine nucléaire RAN•GTP et le pré-miRNA (Bohnsack et al., 2004 ; Lund et al., 2004).

Dans le pore nucléaire, le GTP est hydrolysé, le complexe va alors se dissocier et libérer le pré-miRNA dans le cytosol de la cellule.

L'export des pré-miRNA est augmenté lors de dommages à l'ADN de manière ATM dépendante (Melo et al., 2010 ; Wan et al., 2013). Une irradiation ou des UV induisent des cassures doubles brins à l'ADN. Ces cassures vont activer la protéine ATM qui activera, à son tour, la protéine ATR. Cette dernière va phosphoryler la nucléoporeine NUP153. NUP153 phosphorylée interagit avec l'exportine 5, ce qui permet d'augmenter l'export des pré-miRNA dans le cytosol.

c. Maturation du pré-miRNA dans le cytosol

Dans le cytosol, le pré-miRNA est clivé par une seconde endonucléase : Dicer. Cette enzyme va générer un duplex de microARN de 22 nts qui sera pris en charge par le complexe RISC (*RNA-induced silencing complex*) (Bernstein et al., 2001 ; Grishok et al., 2001 ; Hutvagner et al., 2001 ; Ketting et al., 2001 ; Knight & Bass, 2001). Dicer est une endonucléase, de 200kDa, de type RNase III. Cette protéine est composée de trois domaines importants (fig 16) (Macrae et al., 2006 ; Park et al., 2011 ; Tsutsumi et al., 2011 ; Tian et al., 2014) :

- En partie Nt : une hélicase qui reconnaît le pré-miRNA et qui interagit avec la boucle de ce dernier
- En partie intermédiaire : un domaine PAZ (PIWI/AGO/ZWILLE) qui interagit avec la partie terminale du pré-miRNA.
- En partie Ct : un domaine RNase III en tandem qui va former le centre catalytique de la protéine

Dicer, grâce à son domaine PAZ, reconnaît à la fois l'extrémité 5' phosphorylée et l'extrémité 3' débordante du pré-miRNA (Zhang et al., 2004). Le domaine PAZ en se fixant sur ces deux extrémités crée une poche dans laquelle le pré-miRNA sera piégé (fig 16). Cette poche permet aux domaines hélicase et RNase III de la protéine Dicer de se retrouver au niveau de la boucle du pré-miRNA. La distance entre les domaines PAZ et RNase III joue donc un rôle de « règle moléculaire ». Dicer pourra alors cliver le pré-miRNA à la base de cette boucle pour

Introduction

généraliser le duplex de microARN : ARN doubles brins de 22 nts (Park et al., 2011 ; Tian et al., 2014). Dicer clive le pré-miRNA à une distance comprise entre 21 et 25 nts de chacune des extrémités du pré-miRNA.

Chez *D. melanogaster*, il existe deux protéines Dicer : Dicer 1 et Dicer 2. L'association du pré-miRNA avec la protéine Dicer est fonction de la structure de ce pré-miRNA. Dicer 1 est associé aux miRNA tandis que Dicer 2 est associé aux siRNA (Lee et al., 2004). Chez l'Homme, il n'existe qu'une seule protéine Dicer qui est en charge à la fois des miRNA et des siRNA.

Un *knock-out* de la protéine Dicer est létal chez la souris au stade embryonnaire (Bernstein et al., 2003). Ceci montre encore une fois l'importance des microARN au cours du développement.

Comme la RNase Drosha, Dicer peut être associée à des cofacteurs : TRBP (TAR RNA binding protein) et PACT (double strand RNA binding domain) (Chendrimada et al., 2005 ; Haase et al., 2005 ; Lee et al., 2006). Cependant, à l'inverse des cofacteurs associés à Drosha, la fonction de ces derniers est encore mal définie à ce jour.

d. Maturation du duplex de microARN

Le duplex de microARN généré par Dicer va ensuite être pris en charge par une protéine de type Argonaute (AGO) pour former le complexe RISC (Tabara et al., 1999 ; Hammond et al., 2001 ; Mourelatos et al., 2002). Au sein de ce complexe, le microARN poursuit sa maturation : le brin passager du duplex est dégradé.

Les protéines de type Argonaute sont divisées en plusieurs classes : AGO, PIWI et Worm-specific AGO (WAGOs). Les protéines AGO sont des protéines ubiquitaires, présentes dans le complexe RISC. Elles permettent la maturation des microARN, des siARN mais aussi leur fonction sur les ARNm cibles. Les protéines PIWI sont spécifiques des cellules germinales et elles sont principalement associées aux piARN.

La protéine AGO est constituée de plusieurs domaines contenus dans deux lobes principaux (Song et al., 2004 ; Ma et al., 2005 ; Parker et al., 2005 ; Wang et al., 2008b) :

- Lobe Nt : il contient un domaine Nt et un domaine PAZ
- Lobe Ct : il contient un domaine MID (middle) et un domaine PIWI

Le domaine PIWI possède une structure similaire à la RNase H et permet le clivage de l'ARNm cible entre les nucléotides 10 et 11 du microARN (Liu et al., 2004 ; Song et al., 2004 ; Parker et al., 2005).

Chez l'Homme, il existe 4 protéines AGO (1, 2, 3, 4) qui peuvent former le complexe RISC. La protéine AGO2 permet de cliver l'ARNm ciblé tandis que les 3 autres protéines (AGO1, AGO3 et AGO4) empêchent la traduction de l'ARNm.

Comme pour les protéines Drosha et Dicer, un KO AGO2 est léthal au stade embryonnaire (Liu et al., 2004).

Chez *D. melanogaster*, la protéine AGO1 prendra en charge les miRNA possédant un U en position 5'. La protéine AGO2 prendra en charge les siRNA possédant un C en position 5' (Okamura et al., 2004 ; Förstemann et al., 2007 ; Tomari et al., 2007). Chez l'Homme, les 4 protéines AGO prennent en charge les miRNA et siRNA sans aucune distinction (Meister et al., 2004 ; Liu et al., 2004 ; Azuma-Mukai et al., 2008 ; Su et al., 2009 ; Yoda et al., 2010 ; Dueck et al., 2012). Le chargement du complexe RISC est encore mal compris mais il se pourrait que les protéines Dicer et TRBP y jouent un rôle (Chendrimada et al., 2005 ; Gregory et al., 2005 ; Maniataki & Mourelatos, 2005 ; Miyoshi et al., 2005 ; MacRae et al., 2008). Cependant, un KO Dicer n'empêche pas les siRNA de s'associer au complexe RISC ; la protéine Dicer n'est donc pas indispensable au chargement de ce complexe (Kanellopoulou et al., 2005 ; Murchison et al., 2005).

Dans un premier temps, le complexe RISC est immature : pré-RISC puisque le duplex de petits ARN (siRNA ou miRNA) est formé. Après chargement du RISC, le brin passager du duplex est éliminé. S'il y a un mésappariement au centre du duplex, l'endonucléase C3P0 va cliver le brin passager pour générer un RISC mature (Matranga et al., 2005 ; Miyoshi et al., 2005 ; Rand et al., 2005 ; Leuschner et al., 2006 ; Diederichs & Haber, 2007 ; Liu et al., 2009 ; Ye et al., 2011). Un mésappariement entre les nucléotides 2-8 et 12-15 peut également permettre un désassemblage du duplex de microARN de façon ATP-indépendant (Tomari et al., 2004, 2007 ; Yoda et al., 2010). Cependant, il est très rare qu'il y ait un mésappariement dans un duplex de microARN, le brin passager d'un microARN n'est donc pas éliminé de cette manière.

Par ailleurs, dans un duplex de microARN, aucune marque spécifique n'a été identifiée pour discriminer le brin guide du brin passager. A ce jour, nous savons que le brin possédant une extrémité 5' moins stable thermodynamiquement sera le brin guide du duplex ; l'autre brin sera le brin passager, il sera éliminé (Khvorova et al., 2003 ; Schwarz et al., 2003). La faible stabilité du brin guide pourrait faciliter la dissociation du microARN de son ARNm cible (Kawamata & Tomari, 2010). De plus, la présence d'un U en position Nt peut être aussi la caractéristique du brin guide (Lau et al., 2001 ; Czech et al., 2009 ; Hu et al., 2009 ; Okamura et al., 2009 ; Ghildiyal et al., 2010). Après la dissociation du duplex, le brin passager est rapidement dégradé. Le complexe RISC est alors mature, il va pouvoir scanner tous les ARNm dans la cellule pour permettre la fixation du microARN qu'il contient sur ses cibles. Ainsi, les modalités

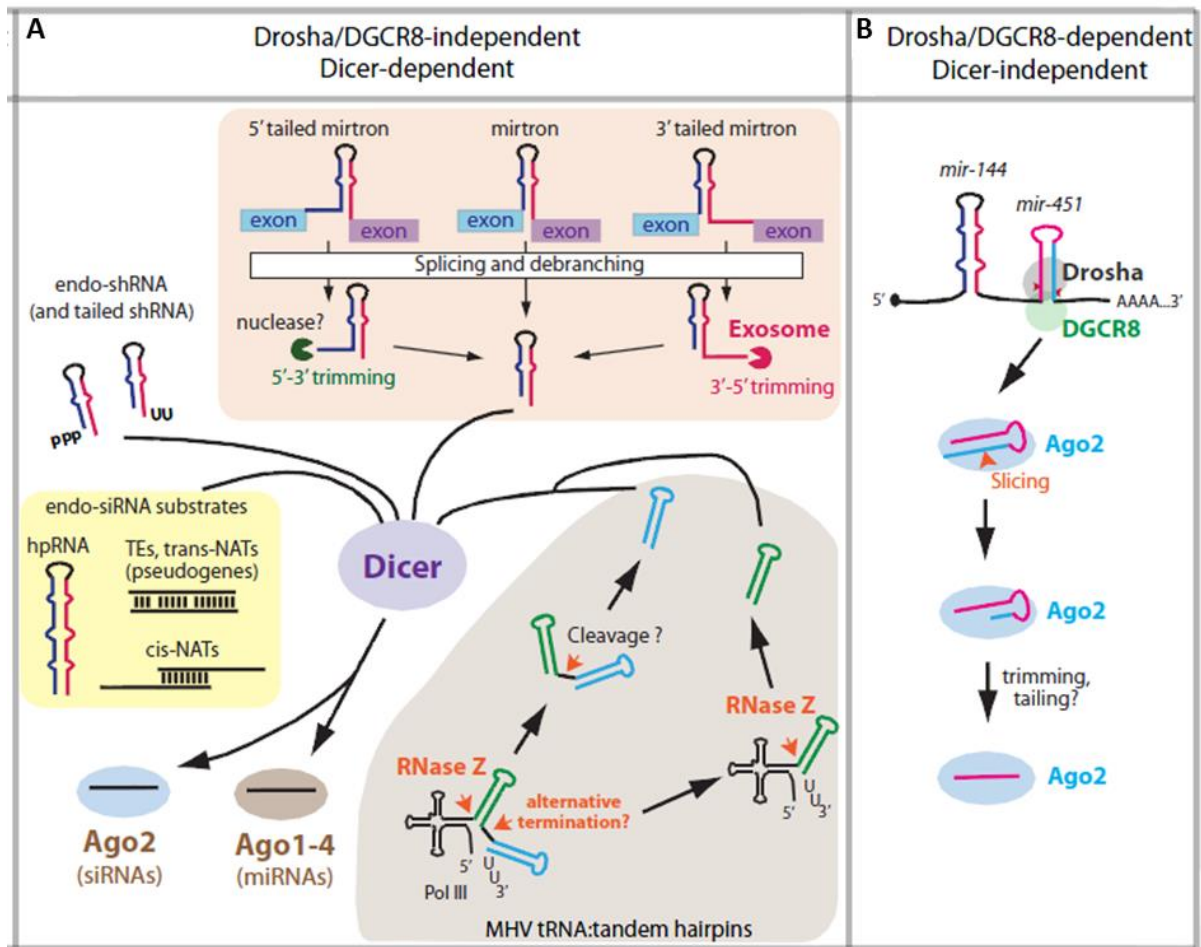


Figure 17. Autres voies de synthèse des microARN: voie non canonique. A. Voie Drosha indépendante. Lors de l'épissage des ARNm, les introns peuvent être reconnus comme des pré-miRNA nommés mirtrons. Ils sont pris en charge par l'enzyme Dicer pour générer un ARN doubles brins. Dans quelques cas, l'action d'exonucléase 5'-3' ou 3'-5' est nécessaire. Les ARN de transfert, après action de RNase Z et P, peuvent aussi former un pré-miRNA reconnu par Dicer. **B. Voie Dicer indépendante.** Lorsque que Drosha clive son substrat, elle génère un pré-miRNA d'une taille de 18 nts ne pouvant pas être reconnu par Dicer. Ce pré-miRNA est reconnu par le complexe RISC qui dégradera la boucle et permettra la maturation du microARN. (adapté de Yang & Lai, 2011).

de maturation (brin guide, brin passager, mécanisme de dissociation du duplex) du pré-RISC sont encore mal définies.

Le brin guide du duplex est prédominant dans la majorité des tissus. Cependant, dans certains cas, le brin passager du duplex peut aussi avoir des ARNm cibles. Nous pouvons, par exemple, citer le miR-675-5p (brin guide) qui agit dans le colon et le miR-675-3p (brin passager) qui possède un rôle dans le placenta (Tsang et al., 2010 ; Keniry et al., 2012). Ces deux microARN sont générés par l'ARN non codant *H19* (Cai & Cullen, 2007).

2. Les voies non canoniques

Les microARN peuvent être synthétisés de façon Drosha ou Dicer-indépendante. Ces voies de synthèse non canoniques sont moins fréquentes.

a. La voie non canonique Drosha-indépendante

Dans le cytosol, lors de l'épissage des ARNm, les introns excisés peuvent être reconnus par l'enzyme Dicer comme étant des pré-miRNA (fig 17). A partir de ces introns, nommés mirtrons, Dicer va générer un duplex d'ARN qui sera ensuite recruté dans le complexe RISC pour y être mûré (Berezikov et al., 2007 ; Okamura et al., 2007 ; Ruby et al., 2007). Cette voie de synthèse ne nécessite donc pas l'action de l'enzyme Drosha. Parfois, l'intervention de RNase est nécessaire pour digérer les extrémités 5' et/ou 3' de l'intron épissé et ainsi permettre sa reconnaissance par l'enzyme Dicer (Flynt et al., 2010).

Les microARN peuvent également provenir d'ARN de transfert (Babiarz et al., 2008). Ces ARN de transfert sont des petits ARN transcrits par l'ARN polymérase de type III. Nous pouvons citer l'ARN de transfert à l'isoleucine (tRNA-Ile) qui peut être digéré par deux RNases : la RNase P et la RNase Z. Ces deux RNases digèrent l'ARNt respectivement aux extrémités 5' et 3'. Cette digestion génère un substrat semblable au pré-miR qui sera reconnu par Dicer. L'action de cette endonucléase permet ensuite la formation du miR-1983 (Babiarz et al., 2008).

Lors de la transcription de certaines parties du génome par l'ARN polymérase III, des petits ARN ayant une structure secondaire type épingle à cheveux peuvent être formés. Ils sont reconnus par l'exportine 1 qui les exportera dans le cytosol de la cellule. Dans ce compartiment cellulaire, ces petits ARN seront reconnus par l'enzyme Dicer qui va alors cliver la boucle de l'épingle et former un duplex de microARN (Ha & Kim, 2014).

b. La voie non canonique Dicer-indépendante

A ce jour, un seul microARN synthétisé par cette voie a été identifié ; il s'agit du miR-451 (Cheloufi et al., 2010 ; Cifuentes et al., 2010 ; Yang et al., 2010). Ce microARN, sous la forme pri-miR-451 est dans un premier temps clivé par l'enzyme Drosha dans le noyau de la cellule

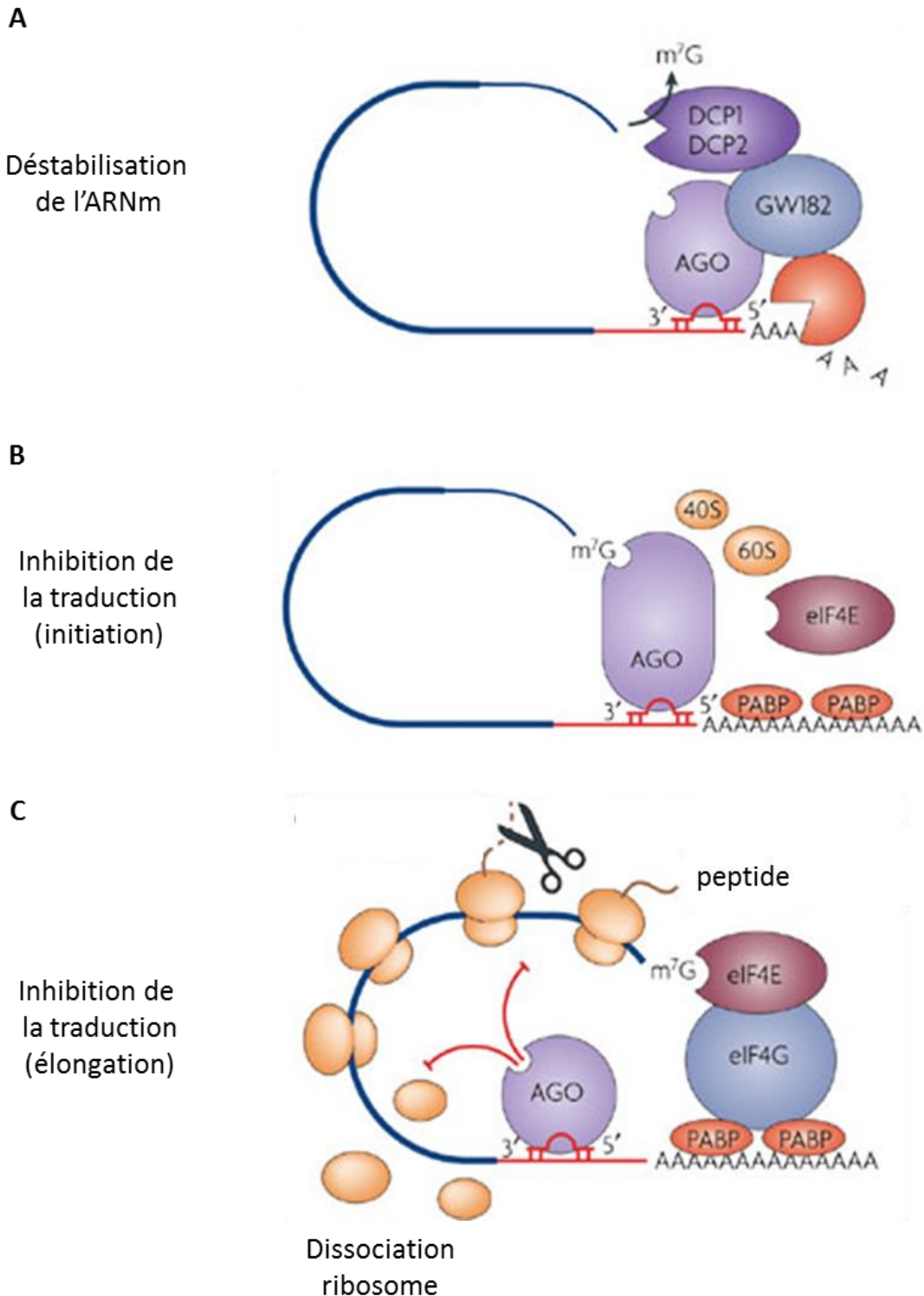


Figure 18. Mode d'action des microARN. (A) La protéine AGO2, associée aux protéines GW182 et DCP1/2, permet la déstabilisation de l'ARNm en dégradant la queue polyA ainsi que la coiffe en 5'. (B) Les microARN associés au complexe RISC se fixent sur l'UTR3' des ARNm et interagissent avec la coiffe située à l'extrémité 5'. Cette association empêche la fixation du facteur eIF4E sur l'UTR5' et empêche donc la fixation du ribosome sur l'ARNm. (C) La protéine AGO2, présente dans le RISC, peut également réprimer la phase d'élongation de la traduction de l'ARNm. (adapté de Stefani & Slack, 2008)

pour former un pré-miR-451 d'environ 18 nts (fig 17). Ce pré-miRNA possède une taille suffisante pour être exporté dans le cytosol *via* l'exportine 5. Dans le cytosol, ce pré-miRNA n'est pas reconnu par Dicer puisqu'il est trop petit. Cependant, il possède une taille correcte pour être pris en charge directement par le complexe RISC. Dans ce complexe, la boucle du pré-miR-451 est clivée par la protéine AGO2 qui possède une activité RNase ; obtention d'un miR-451 mature (Yoda et al., 2013).

L'existence de ces voies alternatives nous montre la flexibilité de la synthèse des microARN. Chez les vertébrés, seulement 1% des microARN sont synthétisés par l'utilisation de ces voies non canoniques.

III. Rôles des microARN

Les microARN ont un rôle biologique important puisqu'ils sont impliqués dans le développement, la prolifération et la différenciation cellulaire. Ils ont également un rôle dans la transduction du signal, l'apoptose ou encore la tumorigenèse. Dans ce paragraphe, nous étudierons les différents mécanismes associés aux microARN. Nous verrons également quelques exemples de microARN impliqués dans des cancers.

1. Mode d'action et régulation des microARN

a. Régulation négative par les microARN

Il est aujourd'hui admis que les microARN exercent un effet négatif sur leurs ARNm cibles. En général, les microARN engendrent la dégradation de leurs ARNm cibles ou l'inhibition de leur traduction en se fixant les UTR3'. Leur action est régulée par l'association du complexe RISC avec des cofacteurs et/ou par la présence de protéines spécifiques sur les ARNm. De plus, le mode d'action des microARN est également déterminé par la complémentarité de séquence entre le microARN et son ARNm cible. En effet, une complémentarité parfaite entre le microARN et son ARNm cible engendre une dégradation de l'ARNm cible (Hutvagner & Zamore, 2002 ; Martinez & Tuschl, 2004). Cette dégradation est permise par l'activité endonucléase de la protéine AGO2. Cette enzyme va cliver le microARN et son ARNm cible entre les nucléotides 10 et 11. Ce clivage entraînera la dégradation de l'ARNm par l'action d'exonucléase 5'→3' et 3'→5'.

Par ailleurs, si la complémentarité du microARN et de son ARNm cible est imparfaite (notamment entre les nucléotides 10 et 11), l'ARNm cible ne sera pas traduit et une modification de sa stabilité, suite à une déadénylation et une perte de la coiffe 5', sera observée (Behm-Ansmant et al., 2006 ; Giraldez et al., 2006 ; Jackson & Standart, 2007 ; Wu et al. 2006) (fig 18 A).

Introduction

Le mécanisme permettant la répression de la traduction de l'ARNm n'est pas clairement défini. Néanmoins, plusieurs études montrent que le microARN empêche la fixation du complexe d'initiation de la traduction tandis que d'autres montrent que le microARN empêche l'élongation de la protéine (Humphreys et al., 2005 ; Pillai et al., 2005 ; Maroney et al., 2006 ; Nottrott et al., 2006) (fig 18 B, C). En effet, la protéine AGO2, lorsqu'elle est présente sur l'UTR3' des ARNm, est capable de fixer l'UTR5' de cet ARNm, et plus précisément la coiffe de celui-ci. Cette interaction empêche le recrutement de la protéine eIF4e, protéine à l'origine de l'initiation de la traduction des ARNm (Humphreys et al., 2005 ; Mathonnet et al., 2007).

De plus, la protéine AGO du complexe RISC est capable d'interagir avec la protéine GW182 : une protéine localisée dans les P-bodies. Cette association engendre la localisation du complexe RISC, du microARN et de son ARNm cible dans ce compartiment cytosolique. Aucune protéines impliquées dans la traduction des ARNm n'ont été retrouvées dans les P-bodies. La localisation du complexe RISC-miR-ARNm dans ce compartiment permettrait donc également la répression de la traduction des ARNm. Cependant, une déplétion des P-bodies n'affecte pas la répression de la traduction des ARNm (Chu & Rana, 2006 ; Jackson & Standart, 2007). Ainsi, la localisation du complexe RISC-miR-ARNm dans ce compartiment cellulaire survient après la répression de la traduction de l'ARNm. Dans les P-bodies, les protéines DCP1/2 et GW182 vont ensuite supprimer la coiffe et déadényler l'ARNm qui sera alors dégradé.

b. Régulation positive par les microARN

En 2007, Vasudevan et collaborateurs ont mis en évidence que le miR-369-3p pouvait exercer un effet positif sur l'ARNm *TNF α* (Vasudevan et al., 2007). L'ARNm *TNF α* contient des séquences riches en AU (ARE). Le microARN miR-369-3p reconnaît ces séquences, ce qui lui permet de se fixer sur cet ARNm. Lorsque les cellules sont en prolifération, le microARN est associé au complexe RISC et induit une répression de la traduction de l'ARNm. Lors d'arrêt du cycle cellulaire, les auteurs montrent que la protéine AGO2 est associée à la protéine FXR1. La présence de FXR1 au niveau des séquences riches en AU induit un effet positif du microARN sur son ARNm cible. En effet, lors de privation nutritive, les auteurs ont observé une augmentation de la traduction de l'ARNm *TNF α* . Ainsi, lorsque les cellules sont arrêtées dans le cycle, les microARN ont un effet activateur sur leurs ARNm cibles.

En 2008, Ørom et al., ont mis en évidence que le microARN-10a se fixe sur l'UTR5' d'ARNm de protéine ribosomale (Ørom et al., 2008). La synthèse des protéines ribosomales est régulée par des motifs 5'TOP qui sont sensibles aux signaux de stress cellulaire et au statut nutritif des cellules. Les auteurs montrent que le miR-10a se fixe sur l'UTR5' de ces ARNm grâce à ces motifs. De plus, ils mettent en évidence qu'en absence de nutriment, le miR-10a, en se

fixant sur l'UTR5' de ces ARNm, y exerce un effet positif. En effet, une augmentation de la traduction des protéines ribosomales a été observée lors de privation nutritive. De plus, les auteurs ont également mis en évidence que le miR-10a pouvait se fixer sur l'UTR3' de ces ARNm. Cette interaction est responsable d'une inhibition de la traduction des ARNm codant pour les protéines ribosomales. Ainsi, l'absence de nutriment, la présence de motif 5'TOP et la fixation du miR-10a sur l'UTR5' permettent l'effet positif de ce microARN.

Par ailleurs, la même année, il a été mis en évidence que le miR-122 exerçait un effet positif sur la traduction des protéines du virus de l'hépatite C (Henke et al., 2008).

c. Autres modes de régulation

L'effet d'un microARN peut également être modulé par la présence de protéine sur l'ARNm cible. En effet, lorsque la protéine DND1 est présente sur les ARNm *NANOS1* et *TDRD7*, elle masque le site de fixation du miR-430, ce qui empêche la régulation négative de ce dernier sur ces ARNm (Mishima et al., 2006 ; Kedde et al., 2007).

La protéine HuR se fixe sur l'UTR3' de l'ARNm *CAT1* et empêche la fonction du miR-122. En effet, la présence de HuR sur *CAT1* induit sa translocation des P-bodies vers le cytosol et permet sa traduction (Bhattacharyya et al., 2006). Cette même protéine peut également exercer un effet négatif puisqu'elle agit en synergie avec let-7 pour réprimer la traduction de l'ARNm *MYC* (Kim et al., 2009). Par ailleurs, la protéine HuR est également capable de réguler la formation des microARN (Keniry et al., 2012). En effet, au cours du développement embryonnaire, son association avec l'ARN non codant *H19* empêche la formation du miR-675 généré par ce dernier.

Ainsi, en fonction de l'état physiologique des cellules, des protéines associées au complexe RISC et des protéines associées aux ARNm, l'action des microARN sur leurs ARNm cibles peut être différente.

2. Les microARN dans le cancer

Les microARN sont impliqués dans tous les processus cellulaires. Une dérégulation de leur expression peut être responsable de cancer. En effet, en 2005, il a été mis en évidence que 29 microARN sont dérégulés dans les cancers du sein (Iorio et al., 2005). Dans les cancers, comme pour les gènes, ils peuvent agir en tant qu'oncogène ou suppresseur de tumeur.

a. Suppresseur de tumeur

Les premiers microARN suppresseur de tumeur ont été mis en évidence dans les Leucémie Lymphocytaire Chronique (CLL) (Garzon et al., 2006). Il s'agit des miR-15a et miR-16-1 (Calin et al., 2002). Ces deux microARN régulent négativement l'expression du gène *B-cell*

Introduction

CLL/lymphoma 2 (BCL2), un gène qui est fréquemment surexprimé dans ce type de cancer. Au cours de cancer, l'expression des miR-15a et miR-16-1 est réprimée à cause d'une délétion génétique. Le gène *BCL2* est donc surexprimé, ce qui induit une prolifération accrue des cellules B matures.

Les microARN de la famille let-7 agissent également comme suppresseur de tumeur dans les cancers. En effet, une diminution d'expression de ces microARN a été observée dans plusieurs cancers : cancer du sein et des poumons par exemple.

b. Oncogène

Le miR-155 est le premier microARN oncogène à avoir été mis en évidence. Ce microARN est surexprimé dans les maladies de Hodgkin, lymphome B diffus à large cellules, CLL à cellules B, cancer du poumon et cancer du sein. Ce microARN est codé par l'ARN non codant *BIC* et réprime l'expression de l'histone désacétylase 4 (*HDAC4*) en se fixant sur l'ARNm (Sandhu et al., 2012). Cette protéine étant un co-répresseur de l'expression de *BCL6 (B-cell lymphoma 6)*, l'expression de *BCL6* est donc augmentée dans ces cancers. Ceci permet une augmentation de la prolifération et de la survie des cellules B.

c. Oncogène ou suppresseur de tumeur ?

Si le rôle pour les microARN 15a, 16-1 et 155 est clairement défini dans les cancers, pour d'autres microARN, leur fonction dans les cancers n'est pas clair. En effet, l'expression du miR-146, par exemple, est augmentée dans beaucoup de cancer, notamment dans les cancers du sein, de la prostate, de la thyroïde ou encore des ovaires (He et al., 2005 ; Volinia et al., 2006 ; Dahiya et al., 2008 ; Wang et al., 2008a). Dans les cancers du sein, le miR-146 régule négativement l'expression de la protéine *BRCA1* (protéine impliquée dans le processus de réparation de l'ADN) (Garcia et al., 2011). Ainsi, dans ce cancer, le miR-146 possède un rôle d'oncogène.

A l'inverse, plusieurs autres études ont mis en évidence qu'une diminution d'expression du miR-146 est associée à une augmentation de survie et de prolifération des cellules hématopoïétiques (Boldin et al., 2011). Dans ces cellules, le miR-146 régule négativement l'expression des activateurs de la voie NF- κ B (*IRAK1*, *TRAF6*). Cette voie étant impliquée dans la survie cellulaire, il semblerait donc que le miR-146 possède un rôle de suppresseur de tumeur dans les cellules hématopoïétiques. Par ailleurs une autre étude a mis en évidence que l'expression du miR-146 est régulée positivement par la voie NF- κ B (Taganov et al., 2006). Il existe ainsi une boucle de rétrocontrôle permettant de réguler l'expression de miR-146. Les auteurs supposent qu'en fonction de taux d'activation de la voie NF- κ B, le miR-146 possède un rôle d'oncogène ou de suppresseur de tumeur.

Un second microARN (miR-29) peut également être oncogène ou suppresseur de tumeur (Kasinski et al., 2011). Le miR-29 induit l'apoptose des cellules de leucémies aigues myéloblastiques (AML) ou de cancer du poumon après xélogreffe en souris immunodéficientes (Fabbri et al., 2007 ; Garzon et al., 2009). Ce qui semble indiquer un rôle de suppresseur de tumeur pour ce microARN. A l'inverse, dans les B-CLL, il semblerait que le miR-29 joue un rôle d'oncogène en réprimant l'expression d'une protéine d'adhésion cellulaire : PXDN (Peroxydase homologue) (Santanam et al., 2010).

IV. Les microARN : une cible thérapeutique ?

Au cours du cancer, les microARN ont donc un rôle majeur. Il est nécessaire d'identifier l'ensemble de leurs ARNm cibles afin de les cibler lors de thérapie.

Plusieurs stratégies thérapeutiques ciblant les microARN ont déjà été réalisées. Par exemple, pour les microARN ayant un rôle d'oncogène, des inhibiteurs spécifiques de microARN (antagomirs) ont déjà été testés. En effet, en 2010, Ma et collaborateurs ont injecté des cellules cancéreuses mammaires 4T1 dans des souris immunodéficientes (Ma et al., 2010). Après l'obtention d'une tumeur, ils ont injecté un inhibiteur spécifique du miR-10b dans la veine de la queue des souris. Une réduction importante du taux de métastases dans les poumons lorsque les souris ont reçu l'inhibiteur a été observée. Cependant, ils n'ont pas observé de diminution de taille de la tumeur primaire. Ainsi, l'inhibiteur du miR-10b empêche l'apparition de métastase.

A l'inverse, des stratégies permettant de restaurer l'expression de certains microARN ont également été réalisées. Par exemple, le microARN-34, dont la transcription est régulée par la protéine P53, est souvent réprimé lors de cancers suite aux mutations de la protéine P53 (Kasinski et al., 2011). Plusieurs équipes ont restauré l'expression de miR-34 par l'utilisation de miR-34 mime. Ces auteurs ont injecté le miR-34 mimétique soit dans la tumeur xélogreffée soit dans le système sanguin de souris et ils ont observé une diminution du volume tumoral (Wiggins et al., 2010 ; Trang et al., 2011). Cet effet négatif sur la tumeur est dû à une diminution de la prolifération cellulaire et à une induction de l'apoptose.

En conclusion, les microARN sont nécessaires au bon développement mais aussi au bon fonctionnement des différents processus cellulaires. Dans une cellule, les microARN exercent majoritairement un effet négatif en se fixant sur les UTR3' de leurs ARNm cibles soit en les dégradant soit en inhibant leur traduction. Cependant, dans certaines conditions, les microARN peuvent exercer un effet positif sur leurs ARNm cibles en augmentant leur traduction. Lors de cancer, les microARN peuvent agir comme des oncogènes ou des suppresseurs de tumeur en fonction de leurs cibles. Au

Introduction

cours de cancer, il est donc nécessaire d'identifier le niveau d'expression mais aussi le rôle des microARN afin de trouver de nouvelles cibles thérapeutiques plus spécifiques.

Le locus *H19/IGF2*

Avant de décrire en détail le locus *H19/IGF2*, je vais vous décrire les mécanismes permettant l’empreinte génomique et quelques locus soumis à cette régulation d’expression. Ensuite, je vous détaillerai la complexité du locus *H19/IGF2* : la mise en place de son empreinte, son organisation tri-dimensionnelle, ses nouveaux transcrits ainsi que la fonction des principaux gènes (*IGF2* et *H19*) dans les cellules normales et cancéreuses.

I. L’empreinte génomique parentale

1. Historique de l’empreinte génomique

Les mammifères sont des organismes diploïdes où chaque paire de chromosome est composée d’un chromosome maternel et d’un chromosome paternel. La majorité des gènes présents sur ces chromosomes possède une expression biallélique. Cependant, en 1970, la notion « d’empreinte chromosomique » est apparue chez les mammifères (Cooper et al., 1971 ; Crouse et al., 1971). Elle est principalement décrite au niveau du chromosome X et permet un dosage d’expression des gènes présents sur ce chromosome, plus couramment nommé, dosage génique. Ce mécanisme permet ainsi un même taux d’expression des gènes entre les individus males et femelle.

A la même époque, dans l’objectif de déterminer les régions chromosomiques nécessaires au développement, des gènes sont transloqués dans le génome. Les auteurs ont observé que les souris possédant certaines mutations ne sont pas viables (Searle & Beechey, 1978). Ceci suggère que ce(s) gène(s) ne sont présent(s) ou exprimé(s) qu’à partir de leur chromosome d’origine. Ces observations ont ensuite été complétées dans les années 1980 par deux équipes différentes. En 1984, les équipes de McGrath, Solter et Surani ont injecté deux génomes maternels (gynogénotes), ou deux génomes paternels (androgénotes) ou un génome paternel et un génome maternel (condition normale) dans des ovocytes fécondés dépourvus de pro-nucleus (fig 19) (McGrath & Solter, 1984a, 1984b ; Surani et al., 1984). Les deux équipes ont observé que les zygotes gynogénotes et androgénotes ne sont pas viables. Ainsi, l’ensemble de ces données montrent que dans les génomes maternel et paternel, il existe des gènes spécifiques, non redondants et nécessaires à la survie et au bon développement de l’embryon.

De plus, en 1989, l’équipe de Wolf Reik a mis en évidence que le syndrome de Prader-Willi est transmis uniquement par l’allèle paternel (Reik, 1989).

Les bases de l’empreinte génomique ont été posées lors de la découverte du premier gène soumis à l’empreinte génomique : le gène *Igf2r* (*Insulin-like growth factor type 2 receptor*) (Barlow et al., 1991). Ce gène est exprimé uniquement à partir de l’allèle maternel. Quelques

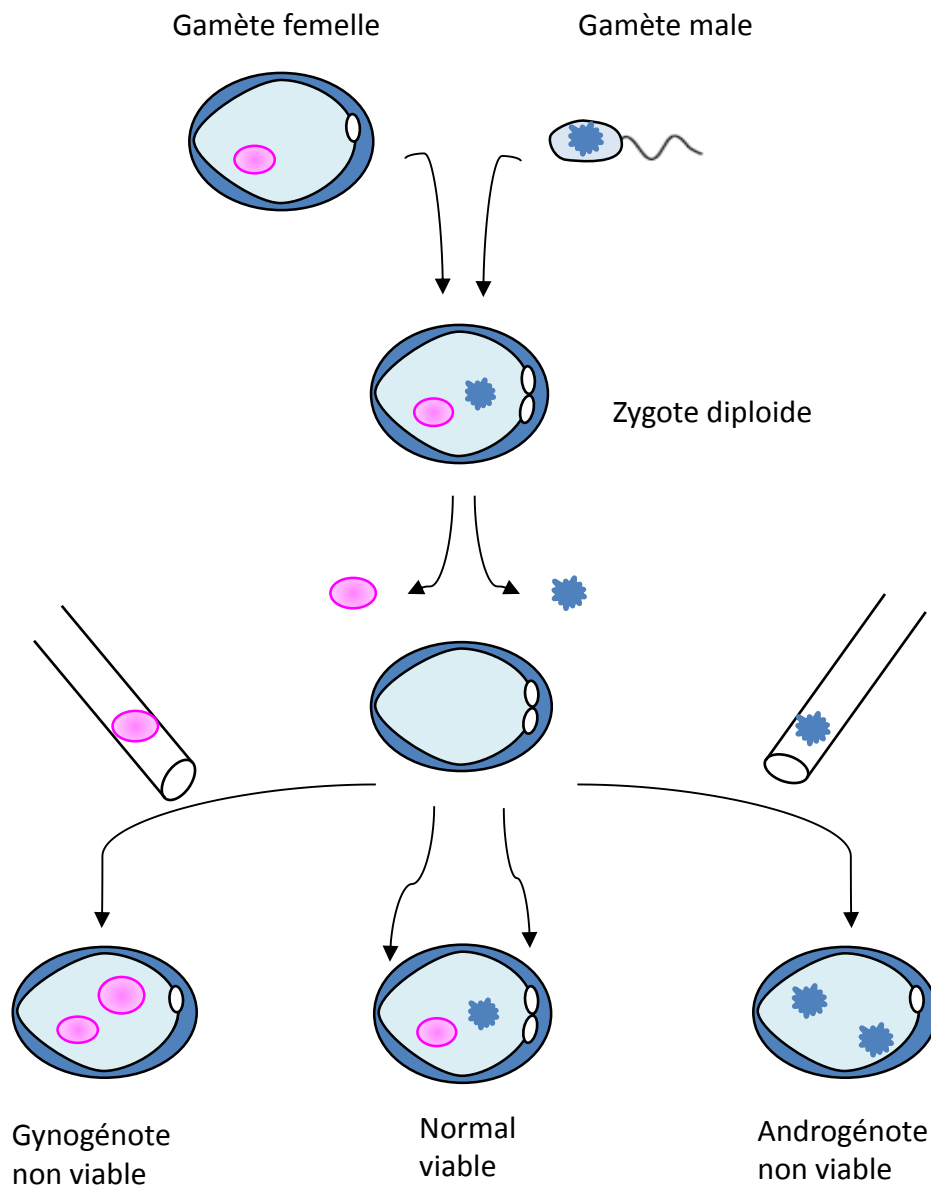


Figure 19. Mise en évidence de l'importance des deux génomes : maternel et paternel.

Dans un zygote fertilisé dépourvu de pro-nucleus, le génome maternel est injecté seul ou en association avec le génome paternel. Le génome paternel est aussi injecté seul ou en association avec le génome maternel. Seuls les individus ayant reçu les deux génomes (maternel et paternel) sont viables. Sur chacun des deux génomes, il existe des gènes essentiels au développement embryonnaire. (adapté de Barlow & Bartolomei, 2014).

mois plus tard, deux autres gènes soumis à l’empreinte génomique et appartenant au même locus ont été identifiés. Ce sont les gènes *Igf2* et *H19* (*clone H19 isolé à partir d’une banque de foie fœtal*) (Bartolomei et al., 1991 ; DeChiara et al., 1991 ; Ferguson-Smith et al., 1991). Ces gènes sont exprimés respectivement à partir de l’allèle paternel et maternel.

L’empreinte génomique parentale est un mécanisme complexe, finement régulé et reposant sur des modifications épigénétiques. Nous en donnons la définition suivante : *expression exclusive d’un seul des deux allèles d’un gène selon son origine parentale et qui résulte de modifications épigénétiques différentes sur chacun des allèles*. Les modifications épigénétiques sont *des modifications chromosomiques réversibles et transmissibles qui s’ajoutent à la séquence nucléotidique apportant une information supplémentaire et se manifestant par une expression phénotypique*.

A ce jour, environ cent gènes sont soumis à l’empreinte génomique parentale. Ces gènes sont regroupés en cluster dont la taille varie de 100 à 4000 kbp. Un cluster contient de 3 à 12 gènes dont, au minimum, un gène produisant un ARN non codant. Le centre de régulation de l’empreinte, IC (*imprinting control*) ou IR (*imprinting region*), est une zone riche en dinucléotides CpG regroupés en îlots, également appelée DMR (*Differentially-methylated region*). Le locus peut également contenir d’autres DMR, ce sont des DMR secondaires dont la méthylation sera fonction de l’état de méthylation de l’IC.

2. Mise en place et maintien de l’empreinte génomique

La mise en place de l’empreinte génomique nécessite la reconnaissance de l’allèle maternel et de l’allèle paternel dans l’embryon. La distinction des deux allèles doit avoir lieu avant la fécondation puisque dans l’embryon, sans marque spécifique, il serait impossible de différencier l’allèle d’origine maternel de l’allèle d’origine paternel. La mise en place de l’empreinte génomique se produit donc lors de la gamétogenèse. Cette empreinte doit être spécifique du gamète d’origine (maternel ou paternel), transmissible aux cellules filles du zygote lors des différentes mitoses et labile dans les gamètes de l’embryon. Ce dernier point est primordial puisque l’individu mâle ou femelle doit produire uniquement un type de gamètes : mâle ou femelle. De manière général, l’allèle maternel est hyperméthylé tandis que l’allèle paternel est hypométhylé. Ces modifications épigénétiques sont maintenues dans les cellules somatiques de l’embryon et de l’individu adulte. Dans les cellules germinales, ces marques épigénétiques sont effacées à un stade précoce ; de nouvelles marques spécifiques du sexe de l’individu sont apposées sur ses gamètes. Ainsi l’empreinte génomique est transmise aux générations suivantes grâce aux gamètes.

Les modifications épigénétiques responsables de l’empreinte génomique sont diverses et variées. Sur l’ADN, une seule modification épigénétique se produit : **la méthylation des**

cytosines. Cette modification épigénétique se produit uniquement sur les cytosines présentes dans des régions riches en CG, dites « îlots CpG ». Ce processus est catalysé par des enzymes appelées *DNA methyltransferase* (Dnmt). La première enzyme identifiée chez l'Homme est la Dnmt1 (Bestor, 1988). Cette enzyme est responsable du maintien de la marque épigénétique lors des divisions cellulaires et est nécessaire au développement normal puisque une délétion du gène *Dnmt1* est létale au stade embryonnaire chez la souris (Leonhardt et al., 1992 ; Li et al., 1992 ; Liu et al., 1998). D'autres enzymes sont, quant à elles, spécifiques d'une méthylation *de novo* de l'ADN : Dnmt3A et Dnmt3B par exemple (Okano et al., 1998). Ces enzymes contiennent dans leur domaine Ct l'activité DNA méthyltransférase. D'autres Dnmts servent de cofacteurs lors de la méthylation de l'ADN : c'est le cas de Dnmt3L qui ne possède pas d'activité méthyltransférase. Cette enzyme contribue à la formation du complexe Dnmt3A/Dnmt3B et permet l'interaction de ce complexe avec la chromatine.

En l'absence de ces enzymes, la méthylation des régions imprimées est modifiée : les gènes possèdent alors une expression biallélique qui est caractérisée par des syndromes du développement. Cependant, dans certains cas, l'absence de méthylation induit une perte d'expression du gène. Ceci suggère que la méthylation de l'ADN réprime mais aussi active l'expression des gènes.

Dans l'embryon précoce, la méthylation générale des allèles est retirée sauf sur certaines régions du génome : les centres de contrôle de l'empreinte génomique (IC/IR). Ces régions doivent conserver leur état de méthylation afin de conserver l'empreinte génomique parentale. Les CpG méthylés sont reconnus par un certain nombre de protéines, notamment les protéines à *methyl-binding domain* (MBD3 par exemple). Ces protéines, en se fixant sur l'ADN méthylé, empêchent la déméthylation de celui-ci dans l'embryon. Ainsi, l'empreinte génomique parentale est conservée au niveau des loci la subissant.

L'empreinte génomique parentale consiste également en des modifications épigénétiques des histones : la méthylation, la désacétylation, la phosphorylation, la sumoylation ou encore l'ubiquitinylation. Ces modifications sont catalysées par différentes enzymes: *histones methyltransferase* (HMT) et *histone deactylase* (HDAC) par exemple. Ces modifications, comme la méthylation de l'ADN, modulent la compaction de la chromatine.

La méthylation des histones peut se produire sur des résidus arginine (modification épigénétique activatrice de l'expression) ou sur des résidus lysine. En fonction de la lysine qui est méthylée, la marque épigénétique est activatrice (H3K4, H3K36, H3K76) ou inhibitrice (H3K9, H3K27, H3K20) de l'expression génique (Voelkel & Angrand, 2007). Ces marques épigénétiques sont permises par plusieurs histones lysines méthyltransférase (HMT) : les protéines des familles SET1, SET2, SUV39 ou encore EZH par exemple. Chacune de ces enzymes sont spécifiques pour la méthylation de chaque lysine, cependant, elles sont

redondantes. L'histone méthyltransférase la plus décrite est la protéine EZH2, appartenant à la famille des EZH (*Enhancer of Zeste homologous*). Cette protéine constitue la sous-unité catalytique du complexe répresseur polycomb PRC2 composé des protéines EZH2, SUZ12 et ESC. L'enzyme EZH2 permet la méthylation des lysines 9 et 27 de l'histone H3 (H3K9, H3K27). Elle est donc impliquée dans la répression de l'expression des gènes. De plus, la marque épigénétique H3K27 est reconnue par les protéines du complexe PRC1. Ce complexe est notamment composé d'une lysine 119 histone H2A E3 ubiquitine ligase. Il permet donc l'ubiquitinylation de ce résidu sur l'histone H2A (Voelkel & Angrand, 2007). Cette modification épigénétique contribue à la répression de l'expression génique en compactant la chromatine. Par ailleurs, certaines lysines méthylées permettent l'expression des gènes (H3K4 par exemple). Cette modification épigénétique est apportée par la MLL1 et supprimée par le complexe LSD1/CoREST par exemple (Voelkel & Angrand, 2007).

La désacétylation des histones, apportée par les *histones deacetylase* (HDAC), permet également la répression génique. Plusieurs HDAC ont été décrites dans la littérature. Elles sont classées en fonction de leur homologie de séquence avec d'autres protéines. On distingue ainsi 4 classes d'HDAC : classes I à IV. Elles ont ensuite été numérotées en fonction de leur découverte. Ainsi, la première histone désacétylase, découverte en 1996, porte le nom HDAC1 (Taunton et al., 1996). A l'heure actuelle, 18 HDAC ont été décrites. Elles sont toutes capables de désacétyler les lysines des histones H3 et/ou H4. Il existe donc une très forte redondance de ces enzymes dans la cellule. Certaines exercent leur action de manière spécifique, c'est le cas de l'HDAC 11, par exemple, qui est impliquée dans la désacétylation des lysines 9 et 14 de l'histone H3. Tandis que d'autres sont plus généralistes : HDAC8 qui permet la désacétylation de toutes les lysines des histones H3 et H4 ou encore HDAC1 qui permet la désacétylation de toutes les lysines sur toutes les histones (H2A, H2B, H3 et H4) (Seto & Yoshida, 2014). L'action de ces enzymes est régulée par trois complexes différents : Sin3, NuRD et CoRest. Ces complexes peuvent être recrutés par des protéines à *methyl-binding domain* (MBD). Ainsi, les HDAC et HMT sont principalement impliquées dans la répression de l'expression des gènes.

Pour finir, plusieurs études ont mis en évidence une relation directe entre la méthylation de l'ADN et les modifications épigénétiques des histones. En effet, les Dnmt1, 3A, 3B et 3L sont capables de recruter les enzymes HDAC1 et EZH2 (Fuks et al., 2000, 2001 ; Deplus et al., 2002 ; Fuks, 2003 ; Viré et al., 2006). De plus, la répression génique apportée par les Dnmt est levée lorsque les cellules sont traitées avec un inhibiteur spécifique des HDAC (la trichostatin A). **Ainsi, les enzymes impliquées dans les modifications épigénétiques (les Dnmt, les HMT et les HDAC) sont conjointement régulées et participent au verrouillage de la chromatine et par conséquent à la répression de l'expression des gènes.**

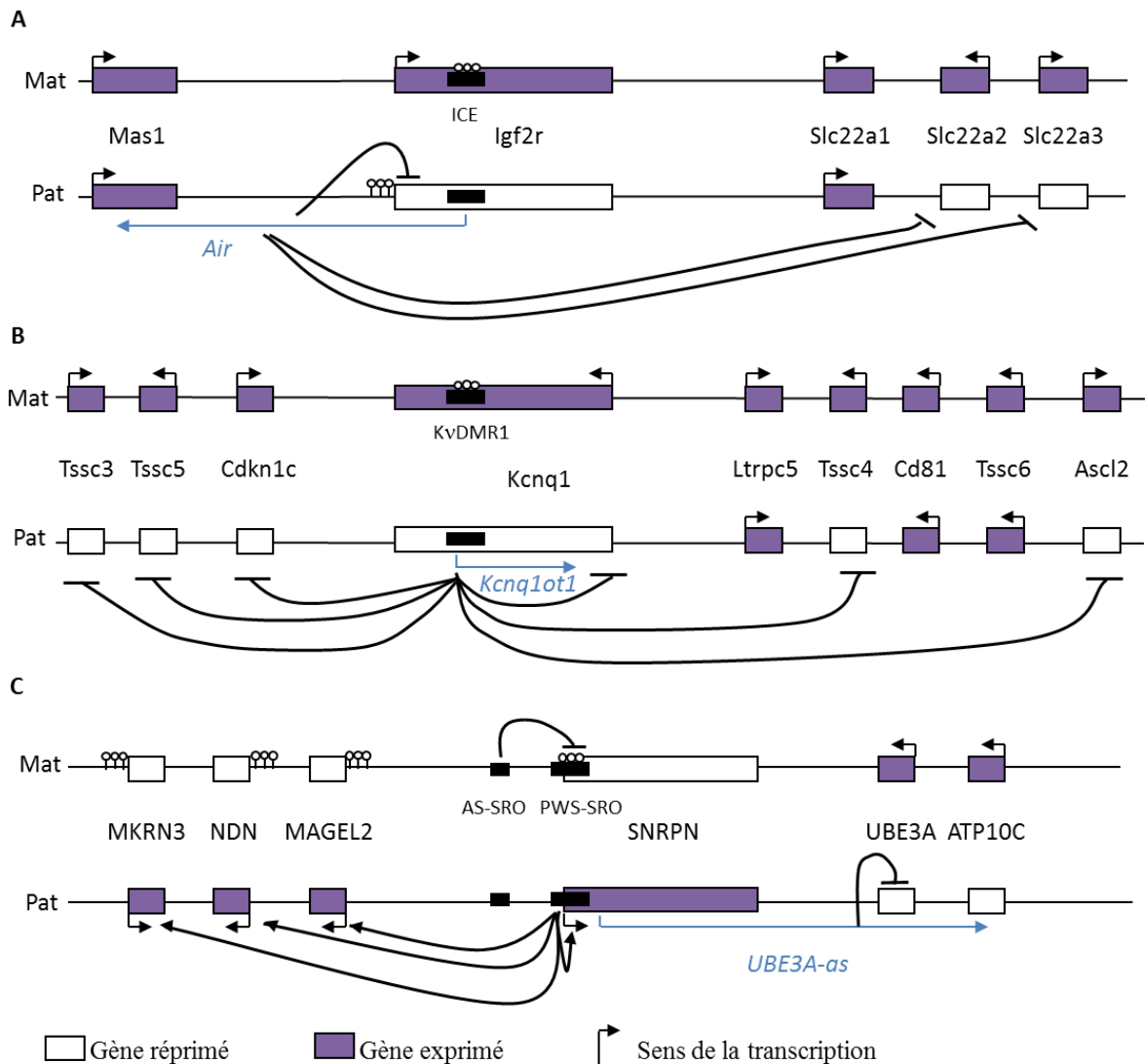


Figure 20. Exemples de locus soumis à l’empreinte génomique parentale. (A) Le locus *Igf2r/Airn*. Sur l’allèle paternel, l’absence de méthylation sur l’ICR permet l’expression de l’ARN non codant *Air*. Celui-ci recrute alors la méthylase G9A qui va réprimer l’expression des gènes du locus (*Igf2r*, *SLC22a2*, *Slc22a3*) à partir de cet allèle par modifications épigénétiques. Sur l’allèle maternel, la méthylation de l’ICR réprime l’expression de *Air*, les gènes *Igf2r*, *SLC22a2*, *SLC22a3* sont alors exprimés. **(B)** Le locus *Kcnq1/Kcnq1ot1*. L’ICR non méthylé sur l’allèle paternel permet l’expression de l’ARN non codant *Kcnq1ot1*. Cet ARN recrute alors les méthyltransférases G9a et EZH2 qui réprimeront l’expression des gènes du locus à partir de cet allèle. Sur l’allèle maternel, l’ICR est méthylé, *Kcnq1ot1* n’est donc pas exprimé, aucune modification épigénétique ne module l’expression des gènes du locus : ils sont donc exprimés. **(C)** Le locus *Ube3A/Ube3A-as*. Sur l’allèle paternel, l’absence de méthylation sur l’ICR permet l’expression de l’ARN non codant *Ube3A-as* et des gènes *MAGEL2*, *NDN*, *MKRN3*. Cet ARN non codant est alors précurseur de snoRNA qui réprimeront l’expression du gène *Ube3A*. La méthylation de l’ICR sur l’allèle maternel permet l’expression du gène *Ube3A* et la répression d’*Ube3A-as*, *MAGEL2*, *NDN*, *MKRN3*. En bleu sont représentés les ARN non codants antisens (adapté de Berteaux et al., 2010).

3. Quelques exemples de locus soumis à l’empreinte génomique

Le premier gène soumis à l’empreinte génomique parentale identifié est le gène *Igf2r* chez la souris (Barlow et al., 1991). Ce gène est localisé sur le chromosome 17 chez la souris et 6 chez l’Homme. L’expression du gène *Igf2r* est régulée par l’expression de l’ARN non codant *Air* (*Antisense Igf2r RNA*). Cet ARN non codant est exprimé à partir du centre de contrôle de l’empreinte (ICE, *imprinting control element*) qui est localisé dans un intron du gène *Igf2r* (fig 20A). Au total, ce cluster de 300kbp contient 4 gènes codants et 1 ARN non codant. Sur l’allèle paternel, l’ICR n’est pas méthylé, l’ARN non codant *Air* est donc exprimé. Il est transcrit en orientation antisens par rapport au gène *Igf2r* (fig 20 A). Lorsqu’il est exprimé, *Air* est capable de fixer un certain nombre de protéines, notamment la protéine G9a. Cette protéine catalyse la méthylation de l’H3K9 ; elle ajoute donc une marque répressive de l’expression des gènes. L’association *Air*/G9a induit la méthylation des gènes du cluster *Igf2r*/*Air* sur l’allèle paternel. Sur l’allèle d’origine maternel, l’ICR étant méthylé, l’ARN non codant *Air* n’est pas exprimé, la G9a n’est donc pas recrutée sur cet allèle, les gènes du locus (*Slc22a3*, *Slc22a2*, *Igf2r*) s’expriment (Sleutels et al., 2002 ; Sanli & Feil, 2015). Des expériences d’invalidation de cet ARN non codant ont mis en évidence des défauts dans la mise en place de l’empreinte génomique. Ainsi, *Air* est essentiel dans la mise en place et le maintien de l’empreinte génomique parentale sur ce locus (fig 20 A).

Cependant, il est à noter que tous les gènes d’un locus imprimé ne sont pas obligatoirement soumis à cette empreinte. En effet, sur ce locus *Igf2*/*Air*, le gène *Slc22a1*, localisé entre deux gènes soumis à ce mécanisme de régulation (les gènes *Igf2r* et *Slc22a2*) possède une expression biallélique. Ainsi, l’empreinte génomique est un processus finement régulé et spécifique de certains gènes.

Le locus *Kcnq1*/*Kcnq1ot1* est également soumis à l’empreinte génomique. Ce locus de 750kbp est localisé sur le chromosome 7 chez la souris et 11 chez l’Homme en position p15 et est constitué de 12 gènes dont 8 sont soumis à l’empreinte génomique (fig 20 B). Le centre de contrôle de l’empreinte, ICR ou encore KvDMR1, est localisé dans un intron du gène *Kcnq1*. Sur l’allèle paternel, l’ICR n’est pas méthylé, l’ARN non codant *Kcnq1ot1* (*Kcnq1 overlapping transcript 1*), antisens à *Kcnq1*, est exprimé. L’expression de cet ARN est associée à plusieurs modifications épigénétiques sur l’allèle paternel du locus : triméthylation de la lysine 27 de l’histone 3 (H3K27me3), diméthylation de l’H3K9 (H3K9me2) ainsi qu’une mono-ubiquitinylation de la lysine 119 de l’histone H2A (H2AK119ub). Ces différentes marques répressives de l’expression géniques sont apportées par différentes enzymes, respectivement EZH2, G9a et RING1B, recrutés par l’ARN non codant *Kcnq1ot1* (Wagschal et al., 2008 ; Zhao et al., 2010 ; Kaneko et al., 2014). La protéine RING1B est un composant du complexe

Introduction

répresseur polycomb PRC1. Ainsi, *Kcnq1ot1*, en recrutant ces enzymes, permet donc de réprimer l'expression des gènes à partir de cet allèle.

Sur l'allèle maternel, l'ICR est méthylé, *Kcnq1ot1* n'est donc pas exprimé, l'ensemble des gènes soumis à l'empreinte sont donc exprimés (*Kcnq1*, *Osbp15*, *Phlda2*, *Slc22a18*, *Cdkn1c*, *Tssc4*, *CD81*, *Ascl2*).

Comme pour le locus *Igf2r/Air*, l'expression de *Kcnq1ot1* est essentielle à la mise en place et au maintien de l'empreinte génomique. Cependant, il a été mis en évidence que le gène *Cdkn1c* reste soumis à l'empreinte génomique dans certains tissus de l'embryon en absence de *Kcnq1ot1* (Shin et al., 2008). Ainsi, l'expression monoallélique de ce gène semble être régulée par un mécanisme indépendant de l'ARN *Kcnq1ot1*. En effet, l'ICR du locus contient des sites de fixation pour la protéine CTCF (CCCTC-binding factor), cette protéine se fixe sur les régions riches en CpG en absence de méthylation (Fitzpatrick et al., 2007). Cette protéine est peut-être impliquée dans la régulation de l'empreinte sur ce locus. Pour les gènes importants dans le développement embryonnaire, il est clair qu'il existe des mécanismes de régulation redondants afin de maintenir leur empreinte génomique et par conséquent leur dosage génique. Ce locus d'empreinte génomique est particulier puisqu'il est associé à un autre locus soumis à l'empreinte génomique, le locus *H19/IGF2*, qui sera décrit et détaillé dans le paragraphe suivant (voir P. 81)

Enfin le troisième locus le plus décrit est le locus *Ube3A/Ube3A-as*. Ce locus, de 2Mbp, est localisé sur le chromosome 7 chez la souris et 15 chez l'Homme en position q11-13 (fig 20 C). Le centre de l'empreinte exerce des fonctions d'enhancers sur ce locus. En effet, sur l'allèle paternel, l'absence de méthylation de l'ICR permet l'expression de l'ARN non codant *Ube3A-as* ainsi que l'expression des gènes *MAGEL2*, *NDN* et *MKRN3*. Le long ARN non codant *Ube3A-as* est alors précurseur de plusieurs petits ARN (snoRNA) qui répriment l'expression du gène *Ube3A* à partir de cet allèle (Rougeulle et al., 1998 ; Runte et al., 2001 ; Landers et al., 2004). Sur l'allèle maternel, la méthylation de l'ICR réprime l'expression de l'ARN *Ube3A-as*, le gène *Ube3A* est donc exprimé. A l'inverse, les gènes *MAGEL2*, *NDN* et *MKRN3* sont réprimés.

Les gènes soumis à l'empreinte génomique parentale possèdent un rôle important dans le développement embryonnaire ou placentaire. De manière générale, il est admis que les gènes exprimés à partir de l'allèle paternel favorisent le développement embryonnaire (*Igf2*, *Peg3*, *Peg1*, *Rasgrf1* et *Dlk1*) tandis que les gènes exprimés à partir de l'allèle maternel ont tendance à réprimer la croissance embryonnaire et/ou placentaire (*Igf2r*, *Gnas*, *Cdkn1c*, *H19* et *Grb10*) (Barlow & Bartolomei, 2014).

Des dérégulations de l’empreinte génomique parentale sont associées à différents syndromes avec des symptômes plus ou moins importants. En effet, la perte de l’empreinte génomique (LOI, *loss of imprinting*) induit généralement des problèmes de croissance de l’embryon, de développement neurologique ou encore des cancers de l’embryon (tératome).

II. Le locus *H19/IGF2*

1. L’empreinte génomique au locus *H19/IGF2*

Le locus *H19/IGF2*, localisé sur le chromosome 11 en position p15, est particulier puisqu’il est adjacent au locus *Kcnq1/Kcnq1ot1*. Ces deux locus imprintés possèdent chacun leur centre de régulation de l’empreinte noté IC1 et IC2 respectivement (fig 21A). Cependant, ces deux locus sont impliqués dans des syndromes du développement suggérant une régulation concomitante de ceux-ci.

Le locus *H19/IGF2* s’étend sur 140kbp environ et contient 2 gènes dont un ARN non codant : *IGF2* et *H19*. Ce locus est conservé entre les espèces et est bordé en 3’ par un gène non soumis à l’empreinte génomique : *MRLP23* (Smits et al., 2008). Comme tous les autres gènes imprintés, les gènes *IGF2* et *H19* sont fortement exprimés dans le développement embryonnaire et réprimés à la naissance dans la majorité des tissus. Le gène *H19* est transcrit en un ARN non codant de 2,3 kb possédant un rôle de suppresseur de tumeur au cours du développement embryonnaire (Brannan et al., 1990 ; Hao et al., 1993). Le gène *IGF2* code pour une protéine impliquée dans la croissance de l’embryon : l’insuline-like growth factor 2 (*IGF2*) (De Chiara et al., 1990). L’empreinte génomique à ce locus est permise par l’ICR (une DMR primaire) qui est localisée 2kbp en amont du gène *H19* (Tremblay et al., 1997) (fig 21B). L’ICR contient 4 sites de fixation pour la protéine CTCF (CCCTC-binding factor) chez la souris et 7 sites chez l’Homme (Hark et al., 2000 ; Szabo et al., 2000 ; Takai et al., 2001). Cette protéine se fixe sur l’ICR lorsque celui-ci n’est pas méthylé. Elle y possède deux rôles : (1) elle empêche la méthylation *de novo* de cette région riche en îlots CpG, (2) elle joue un rôle d’isolateur chromatinien et permet la mise en place et le maintien de l’empreinte génomique. L’expression des gènes de ce locus est également permise par deux sets d’enhancers d’origine mesodermique et endodermique localisés respectivement à 12kbp en amont du gène *H19* et 10,5kbp en aval de ce même gène (Ishihara & Sasaki, 2002 ; Drewell et al., 2002a). L’action de ces enhancers est régulée par la méthylation de l’ICR et la présence de CTCF sur celui-ci. En effet, en 1998, Hark & Tilghman ont mis en évidence une région située en amont du gène *H19* méthylé sur l’allèle paternel (Hark & Tilghman, 1998). Ils montrent que l’allèle maternel est sensible à l’action des nucléases. Leurs travaux mettent en évidence que la méthylation de l’ICR régule l’expression des gènes du locus.

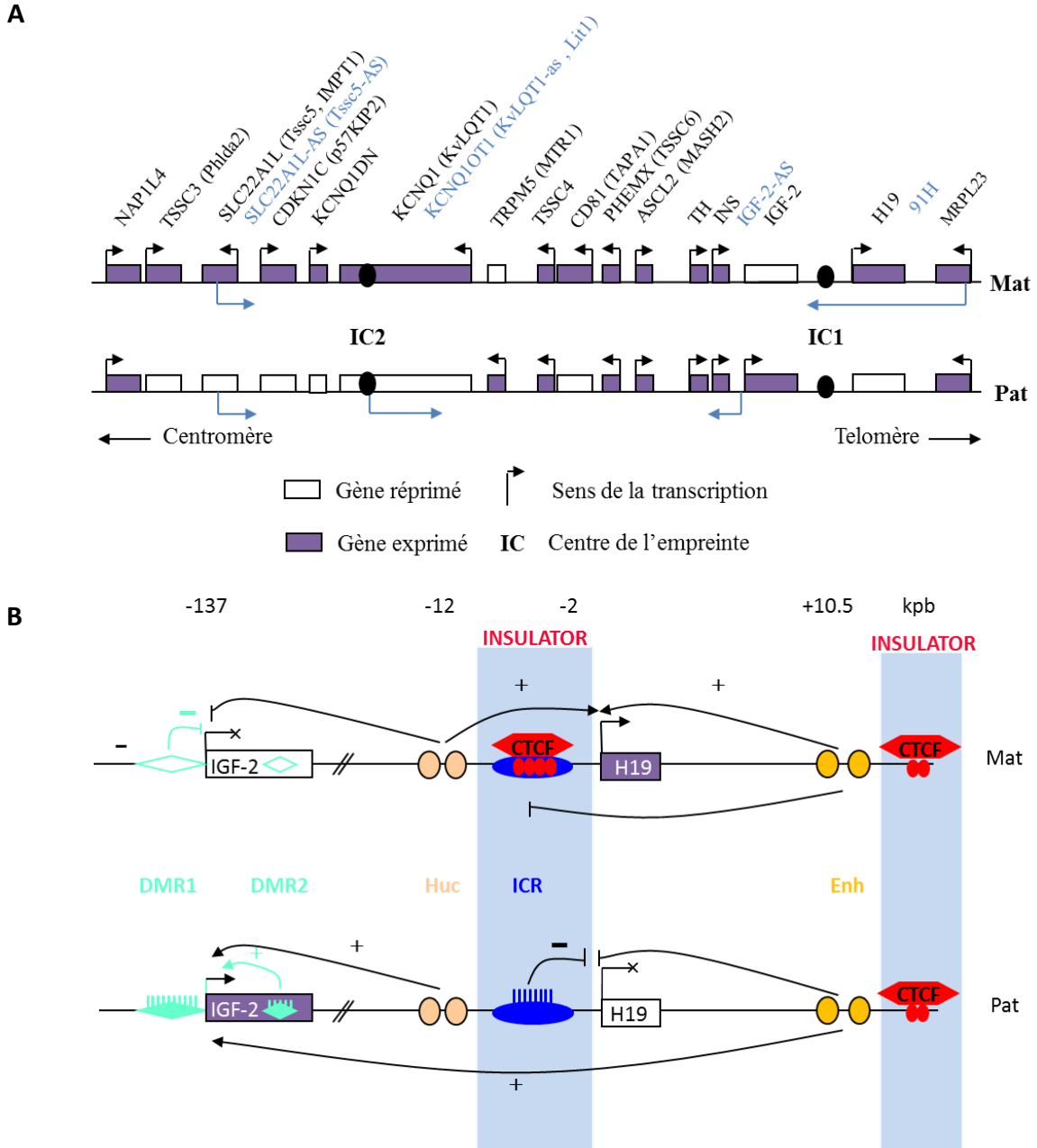


Figure 21. Le locus *H19/IGF2*. (A) Le locus 11p15 chez l’Homme. Ce locus de 1Mbp est constitué de 2 centres de régulation de l’empreinte IC1 et IC2. Ces centres régulent respectivement l’expression des gènes des locus *H19/IGF2* et *Kcnq1/Kcnq1ot1*. Une dérégulation de l’empreinte est responsable de pathologies tel que le syndrome de Beckwith-Wiedemann. En bleu sont représentés les ARN non codants antisens. **(B) Le locus *H19/IGF2* chez l’Homme.** L’expression des gènes *H19* et *IGF2* est régulée par le centre de l’empreinte nommé ICR et deux sets d’enhancers nommés Huc et Enh. En absence de méthylation de l’ICR, la protéine CTCF se fixe sur celui-ci et empêche l’action des enhancers sur le promoteur d’*IGF2* (allèle maternel). Ainsi, à partir de l’allèle maternel, uniquement le gène *H19* est exprimé. Sur l’allèle paternel, la méthylation de l’ICR réprime l’expression d’*H19* et empêche la fixation de la protéine CTCF. Les enhancers peuvent activer l’expression du gène *IGF2*. L’empreinte génomique parentale permet l’expression d’*H19* à partir de l’allèle maternel et l’expression d’*IGF2* à partir de l’allèle paternel. (adapté de Berteaux et al., 2010).

Ainsi sur l'allèle maternel, l'ICR n'étant pas méthylée, la protéine CTCF interagit avec cette région du locus. Elle va alors empêcher l'action des enhanceurs sur le promoteur du gène *IGF2* ; uniquement le gène *H19* sera exprimé (Hark et al., 2000). De plus, l'absence de méthylation des DMR secondaires (DMR1, DMR2) localisées dans le promoteur et le gène d'*IGF2* sur ce même allèle réprime également l'expression du gène *IGF2* (fig 21B) (Constância et al., 2000). Une acétylation des histones sur cet allèle au niveau du promoteur du gène *H19* a également été détectée, permettant ainsi son expression (Grandjean et al., 2001).

A l'inverse, la méthylation de l'ICR sur l'allèle paternel réprime l'expression du gène *H19* en provoquant une hyperméthylation de son promoteur et empêche la fixation de la protéine CTCF (Srivastava et al., 2000). Les enhanceurs vont donc activer l'expression du gène *IGF2* à partir de cet allèle. Par ailleurs, la méthylation de l'ICR permet la méthylation des DMR secondaires (DMR1, DMR2), qui vont également activer l'expression d'*IGF2* sur cet allèle (Murrell et al., 2001). Ainsi *H19* n'est exprimé qu'à partir de l'allèle maternel tandis qu'*IGF2* ne l'est qu'à partir de l'allèle paternel (fig 21B).

Comme la majorité des gènes soumis à l'empreinte génomique, ces deux gènes (*H19* et *IGF2*) sont fortement exprimés et actifs au cours du développement embryonnaire et fœtal. Leur expression a été localisée dans les tissus dérivés des trois feuillets embryonnaires : endodermique, mésodermique et ectodermique (Hemberger et al., 1998 ; Poirier et al., 1991 ; Lustig et al., 1994). Lors du développement embryonnaire, le profil d'expression spatio-temporel de ces gènes est identique, puisqu'ils sont régulés par des enhanceurs commun (Ohlsson et al., 1994). A la naissance, ces deux gènes sont largement réprimés excepté dans certains tissus où un niveau basal d'expression subsiste. En effet, le gène *H19* est fortement exprimé dans le muscle squelettique (Pachnis et al., 1984 ; Leibovitch et al., 1995 ; Milligan et al., 2000) et un niveau basal subsiste dans les poumons, le cœur, le thymus (Poirier et al., 1991), la glande mammaire (Dugimont et al., 1995 ; Adriaenssens et al., 1998), les glandes surrénales (Liu et al., 1995) et l'utérus (Ariel et al., 1997b). A l'inverse du gène *H19* où l'expression est exclusivement d'origine maternelle même à l'âge adulte, une expression biallélique du gène *IGF2* est observée dans le foie, le plexus choroïde et les leptoméninges (Ohlsson et al., 1994 ; Feil et al., 1994). Ainsi l'empreinte génomique peut être régulée de façon spatio-temporelle, il s'agit donc d'un mécanisme actif et finement régulé.

2. Organisation tridimensionnelle de la chromatine

Par 3C (*Chromosome Conformation Capture*), l'organisation tridimensionnelle de la chromatine sur ce locus a été identifiée (Murrell et al., 2004 ; Hagège et al., 2007). Dans ces études, le rôle critique de CTCF dans la mise en place de l'empreinte et la mise en conformation de la chromatine a été mis en évidence. En effet, la protéine CTCF, lorsqu'elle

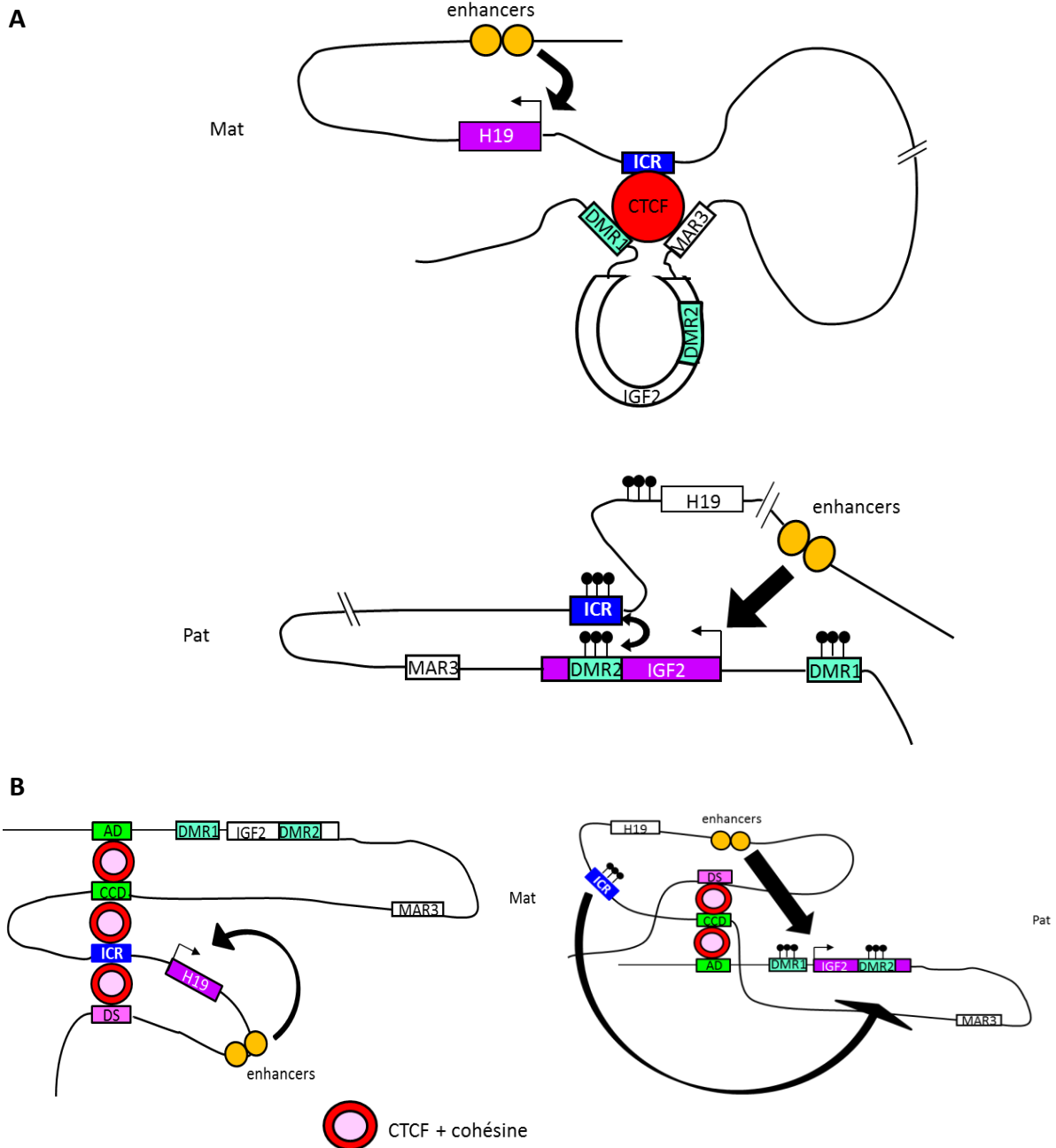


Figure 22. Conformation tridimensionnelle de la chromatine au locus *H19/IGF2* chez l'Homme. (A) Sur l'allèle maternel, la protéine CTCF, fixée sur l'ICR, interagit également avec les domaines DMR1 et MAR3. Ceci permet la formation d'un compartiment silencieux contenant le gène *IGF2* ; seul *H19* est exprimé à partir de cet allèle. Sur l'allèle paternel, l'ICR méthylé interagit avec la DMR2 d'*IGF2*, elle aussi méthylé. Cette interaction permet l'expression d'*IGF2* à partir de cet allèle. **(B) Nouvelle conformation tridimensionnelle du locus *H19/IGF2* chez l'Homme.** Sur l'allèle maternel, la protéine CTCF, associée à la cohésine, se fixe sur 4 domaines différents nommés: CTCF-AD, CCD, ICR et CTCF-DS. Ces interactions permettent la formation de plusieurs boucles. Une boucle est appelée compartiment silencieux, elle empêche l'expression du gène *IGF2*. Une seconde boucle contient le gène *H19* ainsi que les enhancers, ce gène est alors exprimé. Sur l'allèle paternel, l'association CTCF/cohésine se fixe sur les domaines CTCF-AD, CCD et CTCF-DS. Ces interactions permettent l'action des enhancers sur le promoteur d'*IGF2*. (Adapté de Murrell et al., 2004 ; Nativio et al., 2009 ; Ito et al., 2013).

est présente sur l'ICR, est capable d'interagir avec la *matrix attachment region* (MAR3 ; région située en aval du gène *IGF2*) mais aussi la DMR1 du promoteur d'*IGF2* (Weber et al., 2003 ; Kurukuti et al., 2006) (fig 22A). Cette interaction permet la formation d'une boucle piégeant le gène *IGF2*. Cette boucle permet d'une part, l'action des enhancers sur le promoteur d'*H19* et d'autre part, d'éloigner le gène *IGF2* de ces mêmes enhancers (Engel et al., 2008 ; Kato & Sasaki, 2005). De plus, lorsque CTCF est associée aux domaines DMR1 et MAR3, elle recrute le complexe répresseur polycomb PRC2 pour induire la méthylation de la lysine 27 de l'histone H3 sur le promoteur d'*IGF2* (Li et al., 2008). Ainsi, le gène *IGF2* ne peut pas être exprimé à partir de l'allèle maternel.

A l'inverse, sur l'allèle paternel, lorsque l'ICR est méthylé, il interagit avec la DMR2 du gène *IGF2*, elle aussi méthylé. Cette interaction permet un rapprochement des enhancers sur le promoteur d'*IGF2* et donc l'expression de ce gène à partir de cet allèle (Dekker et al., 2002) (fig 22A).

Récemment, plusieurs études ont mis en évidence que la fonction d'isolateur chromatinien de CTCF est permise par son association avec d'autres protéines telle que la cohésine. Cette dernière est constituée de 4 sous-unités : SMC1, SMC3, SCC1/RAD21, SCC3/SA qui sont capables d'interagir avec la protéine CTCF (Rubio et al., 2008 ; Wendt et al., 2008). Dans ces études, les auteurs montrent que l'association CTCF-cohésine est nécessaire pour réguler l'action des enhancers sur le promoteur d'*IGF2* : cette interaction permet à la protéine CTCF sa fonction d'isolateur.

De plus, en 2009, l'équipe de Murrell a montré que le locus *H19/IGF2* possède 3 sites de fixation de la protéine CTCF en plus de l'ICR : CTCF-AD (5kbp en amont de la DMR0 d'*IGF2*), CCD (40kbp en amont de l'ICR) et CTCF-DS (30kbp en aval des enhancers) (fig 22B) (Nativio et al., 2009). Ils montrent que ces trois nouveaux sites de fixation de CTCF sont capables d'interagir entre-eux de façon CTCF-cohésine dépendante. En effet, ils mettent en évidence que le complexe CTCF-cohésine se fixe sur ces sites sur les deux allèles et sur l'ICR en absence de méthylation (fig 22B). Ainsi, les auteurs proposent, une nouvelle organisation tridimensionnelle de la chromatine à ce locus (fig 22B). Cependant, cette nouvelle conformation ne tient pas compte des interactions déjà établies, notamment l'association CTCF/MAR3 et CTCF/DMR1 sur l'allèle maternel.

Ces nouvelles données suggèrent que l'organisation tridimensionnelle de la chromatine est un système dynamique et labile. Par ailleurs, ces différents modèles peuvent être expliqués par l'évolution et l'augmentation de la sensibilité des techniques employées.

3. Modifications épigénétiques au locus *H19/IGF2*

En plus des interactions chromatiniennes régulées par la protéine CTCF, des modifications épigénétiques sont impliquées dans la mise en place de l'empreinte génomique. En effet, la protéine CTCF est capable de recruter la protéine SUZ12 (un composant du complexe répresseur polycomb PRC2) au niveau du promoteur d'*IGF2* (Li et al., 2008 ; Zhang et al., 2011). Le complexe PRC2, contenant également l'histone méthylase EZH2, induit alors la méthylation de la lysine 27 de l'histone H3 (H3K27me) sur le promoteur d'*IGF2*. Cette modification épigénétique réprime l'expression du gène *IGF2* à partir de cet allèle. Ainsi, CTCF, en plus de ses fonctions d'isolateur chromatinien, est également impliquée dans les modifications épigénétiques au locus *H19/IGF2*.

Par ailleurs, sur l'allèle paternel, la méthylation de l'ICR est reconnue par une protéine à *methyl binding domain* : la protéine MeCP2. Cette protéine est constituée de plusieurs domaines : un domaine de fixation sur les îlots CpG méthylé et un domaine de répression transcriptionnel (TRD). Cette protéine est nécessaire au recrutement d'histone désacétylase (HDAC) et d'histone méthyle transférase (HMT) (Drewell et al., 2002b ; Fuks et al., 2003b). Ces deux enzymes permettent la désacétylation et la méthylation de la lysine 9 de l'histone H3 (H3K9) du promoteur d'*H19* sur l'allèle paternel.

Pour résumer, la méthylation du promoteur d'*H19* générée par les Dnmt, ainsi que sa désacétylation et la méthylation des histones générées par les HDAC et HMT permettent un verrouillage de la chromatine conduisant à la répression de l'expression du gène *H19* à partir de l'allèle paternel.

Récemment deux sites de fixation aux facteurs de transcription Sox2 et Oct4, conservés entre l'Homme et la souris, ont été identifiés sur l'ICR (Hori et al., 2012 ; Zimmerman et al., 2013). Ces facteurs de transcription se fixent sur l'ICR non méthylé ; ils se fixent donc sur l'allèle maternel. Ils empêchent la méthylation *de novo* de l'ICR et favorisent l'expression du gène *H19* dans les syndromes de Beckwith-Wiedemann (Zimmerman et al., 2013 ; Abi-Habib et al., 2014). Ces facteurs de transcription sont également impliqués dans le maintien des cellules souches dans les tissus. Venkatraman et collaborateurs ont montré que l'empreinte génomique au locus *H19/IGF2* est nécessaire pour maintenir les cellules souches hématopoïétiques en quiescence (Venkatraman et al., 2013). Ces différentes études nous montrent que les facteurs Sox2 et Oct4 sont impliqués dans la mise en place de l'empreinte génomique au locus *H19/IGF2*. Elles suggèrent également une implication des gènes de ce locus dans les cellules souches.

L'ensemble de ces nouvelles données nous montrent que la mise en place et la régulation de l'empreinte génomique sur ce locus est très complexe : de nouveaux facteurs essentiels à son

contrôle apparaissent et un nouveau modèle d'organisation tri-dimensionnelle de la chromatine se développe.

4. Pathologies associées au locus *H19/IGF2*

Des modifications d'empreinte génomique au locus 11p15 sont responsables de deux syndromes développementaux : les syndromes de Silver-Russel et de Beckwith-Wiedemann. Ces syndromes sont principalement caractérisés par des problèmes de croissance de l'embryon. Le diagnostic pour ces deux pathologies est réalisé pendant la grossesse ou à la naissance.

Le syndrome de Beckwith-Wiedemann affecte une naissance sur 12 000. Il est caractérisé par un gigantisme du nouveau-né à la naissance, une viscéromégalie ainsi qu'une macroglossie pouvant entraîner des troubles de nutrition, d'élocution ainsi que des troubles du sommeil. Ce syndrome est responsable d'une prédisposition de l'enfant aux tumeurs rénales (nommées tumeurs de Wilms) et aux rhabdomyosarcomes. D'un point de vue génétique, il est caractérisé par une expression biallélique du gène *IGF2* ainsi qu'une perte d'expression du gène *H19*. De manière générale, ce syndrome est provoqué par une perte d'expression des gènes exprimés à partir de l'allèle maternel au locus 11p15. Ce locus contient les clusters *Kcnq1/Kcnq1ot1* et *H19/IGF2*. Ainsi au cours de ce syndrome, une perte d'expression des gènes *Kcnq1*, *Cdkn1c* ou encore *H19* est observée. A l'heure actuelle, il est difficile d'associer ce syndrome au cluster *Kcnq1/Kcnq1ot1* ou au cluster *H19/IGF2* puisque bien qu'ils soient indépendants, il semblerait que leur régulation soit associée (Ideraabdullah et al., 2008).

Le syndrome de Silver-Russel est plus rare, il touche 1 naissance sur 75 000 à 100 000. Il est caractérisé par un retard de croissance du fœtus et du nouveau-né avec une dysmorphie faciale et une asymétrie corporelle dans 50 à 60 % des cas. Ce syndrome est associé à une perte d'expression du gène *IGF2* au locus *H19/IGF2*, ainsi qu'une expression biallélique du gène *H19*. Ce syndrome peut également être associé à des mutations sur des gènes présents sur le chromosome 7.

5. Autres transcrits au locus *H19/IGF2*

Au cours des dernières années plusieurs nouveaux transcrits associés au locus *H19/IGF2* ont été identifiés. Ces différents transcrits sont également soumis à l'empreinte génomique parentale et sont, pour la majorité d'entre eux, des transcrits antisens aux gènes *H19* et *IGF2*. Cependant, ils ne sont pas tous conservés entre les espèces ; certains sont spécifiques de la souris tandis que d'autres sont spécifiques de l'Homme.

Le premier transcrit, conservé entre l'Homme et la souris, a été identifié en orientation antisens par rapport au gène *IGF2* : *IGF2-as* (Moore et al., 1997; Okutsu et al., 2000). Ce gène possède

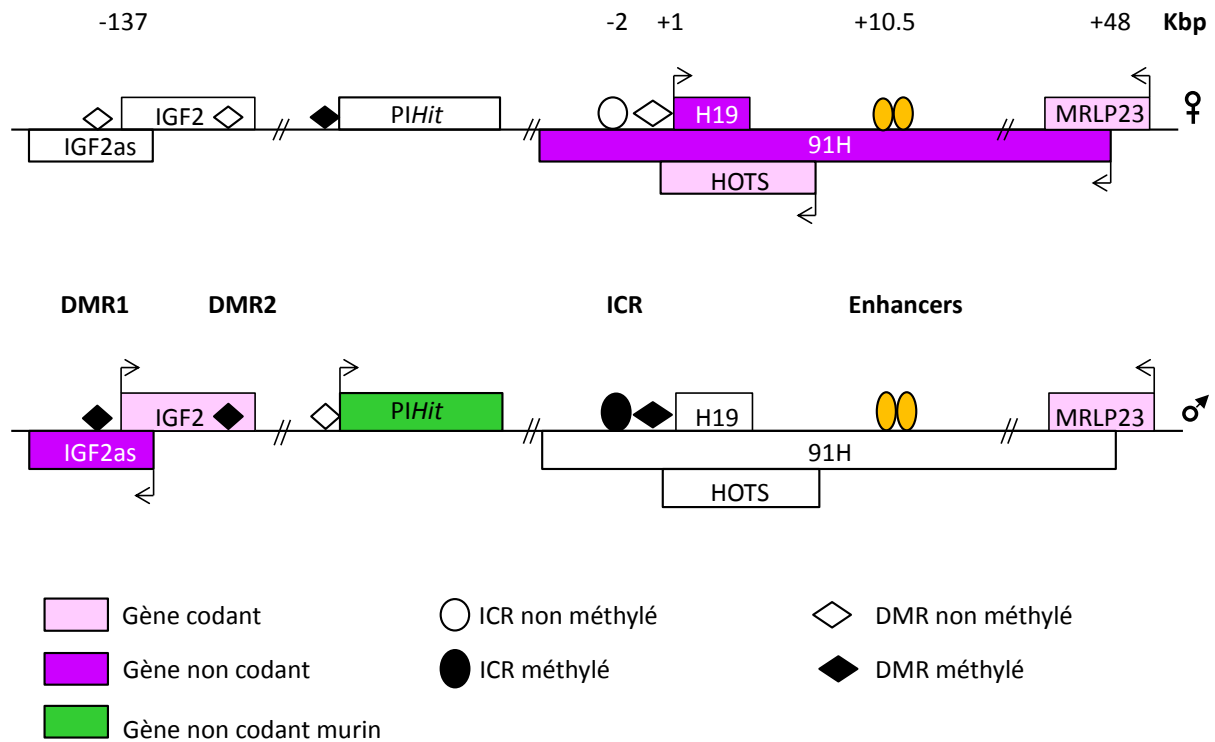


Figure 23. Nouveaux transcrits au locus *H19/IGF2* chez l'Homme. A partir de l'allèle maternel, deux nouveaux transcrits ont été identifiés : *91H*, un ARN non codant de 120kbp et *HOTS*, un gène codant pour la protéine de même nom (Berteaux et al., 2008; Onyango &Feinberg, 2011). Sur l'allèle pternel, deux gènes permettant l'expression d'ARN non codant ont été décrits : *IGF2-as* (Moore et al., 1997 ; Okutsu et al., 2000) et *PIHit* (Court et al., 2011). Cependant, à ce jour, l'ARN non codant *PIHit* a été identifié uniquement chez la souris. Il est localisé à 110 kbp en amont du point d'initiation de la transcription du gène *H19* chez la souris. (Adapté de Angrand et al., 2015).

la même régulation que le gène *IGF2* : il est exprimé uniquement à partir de l'allèle paternel (fig 23). Par ailleurs, aucune phase ouverte de lecture n'a été identifiée sur l'ARNm produit lors de la transcription de ce gène, il s'agirait donc d'un ARN non codant. La fonction de cet ARN non codant n'a pas encore été identifiée. Cependant, il est un bon marqueur dans les tumeurs de Wilms car il y est fréquemment surexprimé.

Néanmoins, récemment, une étude a mis en évidence que l'ARN *IGF2-as* est exporté dans le cytosol et est associé aux polysomes (Duart-Garcia & Braunschweig, 2013). Par conséquent, cet ARN peut coder pour une protéine jusqu'à présent non identifiée.

Plus récemment, nous avons mis en évidence l'existence d'un gène transcrit à partir de l'allèle maternel, en orientation antisens au gène *H19*, que nous avons nommé *91H* (Berteaux et al., 2008). Il s'agit d'un long ARN non codant de 120kbp chez l'Homme (fig 23). Cet ARN est transcrit à partir d'un promoteur contenu dans l'intron 1 du gène *MRLP23*. Cet ARN non codant ainsi que sa fonction sont conservés entre les mammifères, notamment entre l'Homme et la souris (Tran et al., 2012). Cependant, chez la souris, *91H* est un ARN non codant de 90kbp et son promoteur est localisé entre les enhancers et le gène *MRLP23*. Il est connu que les longs ARN non codants antisens sont impliqués dans la régulation de l'empreinte génomique (*Air*, *Kcnq1ot1* ou encore *Tsix*). Nous avons donc étudié les effets d'une invalidation de *91H* sur l'expression des gènes du locus. Nous montrons qu'en absence de *91H*, l'expression d'*IGF2* est diminuée. Le mécanisme responsable de cette inactivation d'expression n'a pas encore été identifié. Cependant, dernièrement, une étude a mis en évidence une corrélation négative entre *91H* et *IGF2* dans des carcinomes œsophagiens (Gao et al., 2015). Les auteurs montrent également que l'expression de *91H* est diminuée dans ces tumeurs, ce qui permet la forte expression d'*IGF2* et par conséquent le développement tumoral. Par ailleurs, dans les carcinomes colorectaux, il a été mis en évidence que l'expression de *91H* favorise la prolifération, la migration et l'invasion cellulaire (Deng et al., 2014). Ce qui lui suggère un rôle d'oncogène dans les cancers. Ces derniers résultats sont en accord avec les données du laboratoire (voir P. 219). En effet, nous avons mis en évidence que les cellules cancéreuses mammaires invalidées stablement pour *91H* ont des capacités de prolifération, migration et invasion diminuées par rapport aux cellules contrôles. De plus, nous avons montré que les capacités de tumorigénèse des cellules sont également affectées lorsque *91H* est réprimé. Cependant, à ce jour, le mécanisme d'action de cet ARN non codant est encore inconnu.

En 2011, une étude a mis en évidence, chez l'Homme uniquement, la présence d'un transcrit maternel antisens au gène *H19* et compris dans l'ARN non codant *91H* : ce transcrit est nommé *HOTS* (*H19 opposite tumor suppressor*) ; il code pour une protéine du même nom (Onyango & Feinberg., 2011) (fig 23). Les auteurs montrent que l'expression de la protéine *HOTS* empêche la formation des tumeurs de Wilms, des rhabdomyosarcomes ou encore des

choriosarcomes. A l'inverse, l'invalidation de *HOTS*, par l'utilisation d'ARN interférant, augmente la croissance tumorale en souris xéno greffées. Il est possible que cette protéine soit traduite à partir de l'ARN *91H*.

Enfin, le dernier transcrit découvert au locus *H19/IGF2* a été identifié chez la souris, il s'agit du gène *PIHit* (*Paternally-expressed IGF2/H19 intergenic transcript*) (Court et al., 2011) (fig 23). Comme son nom l'indique, ce gène est transcrit à partir de l'allèle paternel, il est localisé 110kbp en amont du gène *H19* et code un ARN non codant de 5-6kbp. Par 3C, les auteurs ont mis en évidence l'existence d'une réorganisation de la chromatine sur l'allèle paternel pour induire l'expression de ce gène. Ils suggèrent que la conformation de la chromatine est un système dynamique permettant soit l'expression du gène *IGF2* soit l'expression de l'ARN non codant *PIHit*. Cet ARN non codant est exprimé 8 jours après la naissance chez la souris puis réprimé rapidement à la 3^{ème} semaine après la naissance. L'ARN *PIHit* possède une coiffe à l'extrémité 5' mais n'est pas polyadénylé. Sa fonction dans la cellule n'est pas encore identifiée.

L'ensemble des gènes présents au locus *H19/IGF2* sont soumis à l'empreinte génomique parentale. Une étude a déjà mis en évidence l'existence de deux conformations de la chromatine sur l'allèle paternel pour induire l'expression de ces transcrits (Court et al., 2011). Il n'est donc pas impossible qu'il en soit de même pour l'allèle maternel. Par ailleurs, plusieurs études suggèrent que l'ARN non codant *91H* régule positivement l'expression du gène *IGF2* (Berteaux et al., 2008 ; Tran et al., 2012). Cet ARN est donc probablement impliqué dans la mise en place et la régulation de l'empreinte génomique parentale.

III. Fonctions cellulaires associées au locus *H19/IGF2*

1. Le gène *IGF2*

Comme décrit précédemment, *IGF2* est transcrit à partir de l'allèle paternel. Ce gène, présent sur le chromosome 11 s'étend sur 30kbp. Chez l'Homme, il est exprimé à partir de 4 promoteurs (P1 à P4) et constitué de 10 exons. Les promoteurs P1 à P4 permettent l'expression de 4 transcrits dont uniquement la partie UTR5' sera différente. Ceci n'aura aucune conséquence sur la protéine produite puisque elle est exprimée à partir des exons 8, 9 et 10. Globalement, les promoteurs P2 à P4 sont actifs au cours du développement embryonnaire et fœtal dans tous les tissus somatiques. Le promoteur 3 est principalement actif dans le foie avant la naissance. A la naissance, l'activité de ces 3 promoteurs est réprimée. Seul le promoteur P1 sera actif à l'âge adulte dans le foie, le plexus choroïde et les leptoméninges. De plus, dans ces tissus et à l'âge adulte, l'expression d'*IGF2* à partir du promoteur P1 est biallélique (Bergman et al., 2013). Ainsi l'empreinte génomique parentale n'est présente qu'à certains stades du développement, dans certains tissus et sur quelques

promoteurs. Récemment, une étude a mis en évidence que l'ARN non codant *91H*, transcrit en orientation antisens par rapport à *H19*, permet l'expression du gène *IGF2* à partir de l'allèle paternel à partir d'un nouveau promoteur, noté Pm (Tran et al., 2012). Ce promoteur est situé entre les promoteurs P1 et P2, en amont de la DMR1. Ceci suggère que le mécanisme de régulation de l'empreinte génomique est plus complexe.

La protéine IGF2 est dans un premier temps synthétisée sous forme d'un pré-pro-peptide de 180 acides aminés qui subira ensuite un processus de maturation permettant la formation de l'IGF2 actif de 67 acides aminés (Bergman et al., 2013). Elle possède de fortes homologies de séquence avec l'insuline, d'où son nom. Elle se fixe sur plusieurs types de récepteurs : IGF1R, IGF2R ou encore les récepteurs à l'insuline, avec une affinité plus forte pour le variant A. La fixation d'IGF2 sur son récepteur (IGF2R) n'induit pas de signalisation cellulaire. En effet, ce récepteur est un récepteur « *scavenger* » permettant la dégradation de l'IGF2 lorsque celle-ci se fixe sur ce dernier (Bergman et al., 2012). Ce récepteur permet ainsi de contrôler le taux d'IGF2 circulant dans l'organisme. La fixation de l'IGF2 sur IGF1R ou les récepteurs à l'insuline permet l'activation des voies de signalisation PI3K/Akt et MAP kinase. Ainsi, IGF2 permet la croissance, la migration, la différenciation ou encore la survie cellulaire. Le rôle de l'IGF2 chez l'adulte doit encore être établi. Chez la souris, il a été mis en évidence qu'une perte d'expression d'*IGF2* induit un retard de croissance de 40% des embryons par rapport aux souris témoins. Bien qu'à l'âge adulte ces souris soient viables et fertiles, ce retard de croissance observé chez l'embryon n'est pas compensé lors de l'enfance.

Dans les cancers, il est établi depuis de nombreuses années qu'IGF2 est surexprimé. De nombreuses études montrent que la surexpression du gène *IGF2* dans certains tissus augmente le risque d'apparition du cancer dans celui-ci. En effet, dans des souris transgéniques surexprimant le gène *IGF2*, le risque de développer des tumeurs mammaires ou pulmonaires est augmenté par rapport aux animaux témoins (Bates et al., 1995 ; Moorehead et al., 2003). De plus, des souris ayant un niveau élevé d'expression d'IGF2 à l'âge adulte développent des cancers plus tôt et plus agressif que les souris témoins (Rogler et al., 1994 ; Pravtcheva & Wise, 1998). A l'inverse, une perte d'expression du gène *IGF2* diminue l'incidence tumorale (Bartke et al., 2002). Par ailleurs, la majorité des syndromes de Beckwith-Wiedemann est associée à une perte de l'empreinte génomique conduisant à l'expression biallélique du gène *IGF2*. Cette surexpression d'IGF2 par LOI a également été constatée dans 50% des tumeurs de Wilms (Reeve, 1996).

Dans de nombreux cancers, une LOI est observée : sein (Hartmann et al., 2005), œsophage (Zhao et al., 2009), ovaire (Murphy et al., 2006), ou encore leucémie myéloïde aigue (Wu et al., 1997). Dans les cancers du côlon, la surexpression d'IGF2 après perte d'empreinte est

considérée comme un marqueur potentiel de la tumeur (Cui et al., 2003, 2007 ; Cruz-Correa et al., 2004).

Récemment, il a été mis en évidence que le gène *IGF2* est le précurseur d'un microARN : le miR-483. Ce microARN est généré par l'intron 2 de l'ARNm *IGF2*. Dans les tumeurs de Wilms, où *IGF2* est surexprimé dans 50% des cas, une surexpression de ce microARN a également été constatée (Veronese et al., 2010). Cette corrélation d'expression entre *IGF2* et le miR-483 a également été constatée dans les tumeurs du sein, du colon et du foie. Le miR-483 est suexprimé dans 100% des tumeurs de Wilms. Cependant, dans certaines tumeurs, aucune association entre l'expression du gène *IGF2* et du miR-483 n'a été retrouvée. De plus, les auteurs montrent que l'inhibiteur du miR-483, à l'inverse du siRNA *IGF2*, est capable de réduire la croissance tumorale lorsque les cellules sont xéno greffées en souris immunodéficientes. Ils montrent également que le miR-483 régule négativement l'expression de la protéine pro-apoptotique PUMA, permettant ainsi la résistance des cellules aux apoptogènes. Cette étude suggère que le miR-483 et le gène *IGF2* participent de manière indépendante à la progression tumorale en permettant la résistance à l'apoptose dans un cas et la prolifération cellulaire dans l'autre cas.

Des autres cibles du miR-483 ont été identifiées : *Rho GDP dissociation inhibitor alpha* (RhoGDI1) et *activated leukocyte cell adhesion molecule* (ALCAM) (Song et al., 2014). Ces deux gènes codent pour des protéines impliquées dans l'adhérence des cellules à la matrice extracellulaire. Dans les cellules cancéreuses de poumon, l'expression du gène *IGF2* et du miR-483 est augmentée. Le miR-483, en se fixant sur ses ARNm cibles, favorise alors la transition épithélio-mésenchymateuse des cellules cancéreuses. Par conséquent, il augmente les capacités métastatiques des cellules cancéreuses de poumon.

Enfin, récemment, il a été mis en évidence que l'expression de la protéine MeCP2 (protéine responsable de l'inactivation de l'expression du gène *H19* à partir de l'allèle paternel (voir P. 86)) peut être régulée par le miR-483-5p (Han et al., 2013). En effet, il existe deux sites de fixation du miR-483-5p sur l'UTR3' de l'ARNm MeCP2 mais le miR-483-5p exerce un effet négatif sur le second site uniquement. Le gène *IGF2* est donc également impliqué dans la régulation de l'empreinte génomique parentale.

2. Le gène *H19*

a. Historique

Le gène *H19* a été identifié lors d'un criblage d'une banque de données de foie fœtal chez la souris en 1984 par l'équipe de Tilghman (Pachnis et al., 1984, 1988). Les auteurs ont mis en évidence qu'il est co-régulé avec le gène de l'*alpha*-foetoprotéine. En 1991, Bartolomei et collaborateurs ont mis en évidence que le gène *H19* est soumis à l'empreinte génomique

parentale chez la souris (Bartolomei et al., 1991). En 1992, chez l'Homme, Zhang et Tycko ont isolé le gène *H19* également soumis à l'empreinte génomique parentale (Zhang & Tycko, 1992). Le gène *H19*, exprimé uniquement à partir de l'allèle maternel, est constitué de 5 exons (1325, 134, 112, 122, 613 nucléotides) séparés par des introns très courts. Le gène *H19* est transcrit par l'ARN polymérase II en un ARN de 2,3 kbp qui est exporté dans le cytosol. L'ARN *H19* possède une coiffe à l'extrémité 5', il est épissé et polyadénylé. Il possède donc toutes les caractéristiques d'un ARNm ; cependant à ce jour, il est admis que cet ARN est un ARN riborégulateur.

En effet, bien que plusieurs codons d'initiation de la traduction soient présents dans l'exon 1 de l'ARN, ils sont rapidement suivis par des codons stop : les potentielles phases ouvertes de lecture sont donc très courtes (Brannan et al., 1990). De plus, ces phases ouvertes ne sont pas conservées entre les espèces contrairement à la structure secondaire de l'ARN (Juan et al., 2000). En effet, l'ARN *H19* est composé de 16 structures types épinglées à cheveux conservées entre les mammifères, ce qui suggère que l'ARN *H19* possède ses propres fonctions *per se* (Juan et al., 2000). Enfin, la phase ouverte de lecture la plus longue permettrait la synthèse d'une protéine de 26kDa. En 1996, Joubel et collaborateurs, après délétion et/ou mutation ponctuelle de l'UTR5', ont réussi à obtenir la production de cette protéine dans un système *in vitro* (Joubel et al., 1996). Ils ont donc mis en évidence que le potentiel UTR5' de l'ARN réprime la traduction de celui-ci. En 1990, il a donc été décrit pour la première fois comme étant un ARN riborégulateur (Brannan et al., 1990).

L'ARN *H19* permet de réguler la croissance embryonnaire. En effet, dans des souris « *knock-out* » pour le gène *H19*, une augmentation de la taille et du poids d'environ 10% par rapport aux souris contrôles a été constatée (Leighton et al., 1995 ; Ripoché et al., 1997). Cependant, dans ces souris, l'expression d'*IGF2* est devenue biallélique ; il est donc difficile de déterminer un effet spécifique du gène *H19* sur le développement et la croissance embryonnaire. Une autre équipe a mis en évidence que la délétion de « silencers » spécifique du gène *H19*, permettant l'expression biallélique de celui-ci, permet l'obtention d'animaux plus petits alors que l'expression du gène *IGF2* reste inchangée (Drewell et al., 2000). Ainsi, en condition normale, l'ARN *H19* diminue la croissance de l'embryon.

b. Le gène *H19* dans l'empreinte génomique

Comme largement décrit dans le paragraphe précédent, le gène *H19* est soumis à l'empreinte génomique parentale : son expression est exclusivement d'origine maternelle et il est co-régulé avec le gène *IGF2*, exprimé à partir de l'allèle paternel.

Il est établi que l'expression du gène *H19* module l'expression d'*IGF2* *en trans* (Wilkin et al., 2000 ; Lottin et al., 2002). En effet, ces deux gènes étant régulés par les mêmes séquences

d'enhancers, ils sont en compétition pour ces derniers. De plus, une forte expression du gène *H19* est associée à une diminution de l'expression du gène *IGF2*. Cependant le mécanisme responsable de cette régulation n'a pas encore été identifié.

Par ailleurs, dans des cellules de choriocarcinomes, l'ARN *H19* est associé aux polysomes et une relation inverse entre l'expression d'*H19* et la traduction de l'ARNm *IGF2* a été montrée (Li et al., 1998). *H19* diminuerait la traduction de l'ARNm *IGF2* en empêchant le recrutement de celui-ci aux polysomes. Ainsi *H19* régule *en trans* la transcription et la traduction du gène *IGF2*.

Récemment, l'équipe de Dandolo a mis en évidence que l'ARN non codant *H19* régule l'expression de plusieurs gènes soumis à l'empreinte génomique (*IGF2*, *Slc38a4*, *Dcn*, *Dlk1*, *Peg1*) (Monnier et al., 2013). Dans cette étude, ils montrent que l'ARN *H19* se fixe sur une protéine à *methyl binding domain* : MBD1. La protéine MBD1, en plus de se fixer sur les CpG méthylés, est capable de recruter des HMT, notamment l'histone méthyltransférase Suv39 (Fujita et al., 2003). De plus, ils mettent en évidence que l'association *H19*/MBD1 est nécessaire au maintien de l'empreinte génomique parentale sur les gènes précédemment cités. Cette association *H19*/MBD1 permet ainsi le maintien de la marque répressive H3K9me3 sur ces cinq gènes : *IGF2*, *Slc38a4*, *Dcn*, *Dlk1* et *Peg1*. Il est donc impliqué dans le maintien de l'empreinte génomique parentale sur ces différents locus. Cette étude met donc en évidence le mode d'action employé par l'ARN *H19* pour réguler l'expression du gène *IGF2*. De plus, elle montre que l'ARN *H19* est également impliqué dans la régulation d'expression d'autres gènes, eux aussi soumis à l'empreinte génomique.

c. Régulation de l'expression d'*H19*

En plus de l'empreinte génomique parentale, l'expression du gène *H19* est également régulée par un certain nombre de facteurs : les hormones, les facteurs de croissance ou encore des cytokines.

Dans la glande mammaire, nous avons mis en évidence un niveau basal d'expression du gène *H19* (Dugimont et al., 1995 ; Adriaenssens et al., 1998). Cependant, nous avons montré que l'expression d'*H19* est localisée dans certaines structures de la glande mammaire et qu'elle n'est pas constante au cours de son évolution (Adriaenssens et al., 1999). En effet, par hybridation *in situ*, nous montrons que *H19* est exprimé dans la glande mammaire de façon ubiquitaire lors du développement embryonnaire. A la naissance, il est réprimé dans la majorité des cellules. Une forte expression d'*H19* va être retrouvée dans les nouveaux bourgeons terminaux et dans les cellules épithéliales des acini à la puberté et pendant la grossesse respectivement. Pendant la lactation, aucune expression d'*H19* n'est détectée, il est ré-exprimé dans certaines cellules épithéliales pendant l'involution, pour revenir à un niveau

basal d'expression. Des variations similaires ont été observées dans l'utérus : *H19* est fortement exprimé pendant les phases d'oestrus et de metoestrus et réprimé pendant la phase de latence. Le gène *H19* est donc principalement exprimé dans les cellules épithéliales en prolifération et son expression est corrélée avec les variations hormonales (**œstrogènes, progestérone**) de la Femme. En effet, notre équipe a montré qu'*H19* est régulé positivement par le 17- β -œstradiol et négativement par la progestérone.

Dans la prostate, nous avons montré que *H19* est régulé positivement par la prolactine (une hormone peptidique impliquée dans la lactation mais aussi dans le développement normal de la prostate et la fonction testiculaire) et négativement par la dihydrotestostérone (DHT : forme active de la testostérone) (Berteaux et al., 2004).

De plus, dans les cellules cancéreuses de sein, nous avons montré que l'expression d'*H19* est régulée positivement par un certains nombres de facteurs de croissance : **l'HGF/SF, l'EGF et le FGF-2** (Adriaenssens et al., 2002). Les cellules stimulées avec l'HGF/SF, en plus de montrer une surexpression d'*H19*, ont des capacités de prolifération et de migration augmentées. Nous montrons également que les facteurs TNF- α , TGF β 1 et IGF-2 n'exercent aucun effet sur l'expression du gène *H19*. Ainsi, il n'existe pas de boucle de rétrocontrôle de régulation de l'expression d'*H19* et d'*IGF-2* dans les cellules de sein.

D'autres études ont mis en évidence que l'expression d'*H19* est régulée positivement par les facteurs de croissance : IGF-1, IGF-2 et acide rétinoïque et négativement par le TGF β , le TNF α et l'INF γ dans les cellules surrénales et de testicules (Voutilainen et al., 1994 ; Kopf et al., 1998). Ces études nous montrent que les facteurs régulant l'expression d'*H19* sont différents et peuvent avoir des effets contraires selon le type cellulaire étudié.

Enfin, l'expression d'*H19* peut également être régulée par des protéines ou facteurs de transcription. En effet, dans le promoteur d'*H19* des éléments de réponse à la protéine **P53** ont été identifiés (Dugimont et al., 1998). Lorsque cette protéine est activée, elle réprime l'expression du gène *H19*. Nous avons également mis en évidence que *H19* est régulé positivement par le facteur de transcription **E2F1** (Berteaux et al., 2005). Une autre étude a mis en évidence que le facteur de transcription **c-Myc**, qui agit en hétérodimère avec la protéine Max, est capable de réguler l'expression du gène *H19* (Barsyte-Lovejoy et al., 2006 ; Zhang et al., 2015). Les auteurs montrent une association de l'expression d'*H19* et l'expression de c-Myc dans les tumeurs primaires du sein et dans les cancers des poumons. Ce facteur de transcription se fixe sur des séquences spécifiques, appelées E-boxes, à proximité du centre de contrôle de l'empreinte au locus *H19/IGF2*. En recrutant des histones acétylases, c-Myc active l'expression du gène *H19* uniquement à partir de l'allèle maternel. Ce facteur de transcription ne modifie donc pas l'empreinte génomique au locus. Ces données suggèrent que *H19* participe aux effets oncogènes décrits pour la protéine c-Myc. De la même manière,

Introduction

dans les cancers de la vessie, il est établie que l'expression du gène *H19* est régulée positivement par l'oncogène YAP (*Yes Associated Protein*) (Li et al., 2015). Ainsi, dans ces cancers, les capacités oncogéniques de YAP sont augmentées grâce à l'expression d'*H19*. Par ailleurs, l'expression du gène *H19* est corrélée avec la surexpression du gène *Afp* codant pour l'Alpha-foetoprotéine dans le foie fœtal (Pachnis et al., 1984). En 2005, l'équipe de Perincheri et collaborateurs a cartographié le locus contenant le gène *Afp* et a identifié le gène *Zhx*, orthologue du gène *Zhx2* chez l'Homme (Perincheri et al., 2005). Une relation négative entre l'expression d'*H19* et ce gène *Zhx* a été observée. En effet, des souris déficientes en *Zhx* gardent une expression élevée du gène *H19* après la naissance. De plus, l'expression de *Zhx* est plus importante dans le foie adulte que dans le foie fœtal. L'expression de ce gène serait donc responsable de la répression du gène *H19* observée à la naissance.

Outre les protéines, facteurs de croissance, facteurs de transcription, **le microenvironnement tumoral** est également capable d'influencer l'expression de certains gènes. Nous avons démontré que *H19* est principalement exprimé au niveau des interactions épithélium/mésenchyme dans la glande mammaire normale (Adriaenssens et al., 1999). Dans les cancers du sein, l'expression d'*H19* est localisée préférentiellement au niveau de l'interaction épithélium/stroma tumoral (Dugimont et al., 1995 ; Adriaenssens et al., 1998). De plus, nous avons mis en évidence que les cellules épithéliales mammaires, traitées avec du milieu conditionné de fibroblastes humains normaux (MRC-5), ont une forte expression du gène *H19* (Adriaenssens et al., 2002). Par ailleurs, les cellules ont également des capacités de motilité et de morphogénèse augmentées : propriétés déjà connues pour ce type de milieu. Cependant, les facteurs responsables de l'expression du gène *H19* n'ont pas été identifiés.

Par ailleurs, l'expression d'*H19* en condition hypoxique a été étudiée. En fonction de la proximité des vaisseaux sanguins, les tumeurs sont composées de zones hypoxiques plus ou moins importantes. Les zones hypoxiques sont considérées comme étant les régions responsables de la résistance aux chimiothérapies puisqu'elles sont éloignées des vaisseaux sanguins, elles sont donc inaccessibles aux drogues. En condition hypoxique, deux voies de signalisation sont activées : la voie HIF-1 α et la voie P53. Elles possèdent deux fonctions différentes : la première permet l'adaptation de la cellule à son nouvel environnement tandis que la seconde permet l'induction de la mort cellulaire. Dans leur étude, les auteurs ont montré que le facteur HIF-1 α permet l'expression du gène *H19* uniquement lorsque la protéine P53 est mutée (Matouk et al., 2010). Cette régulation est en accord avec des données obtenues au laboratoire : P53 réprime l'expression du gène *H19* (Dugimont et al., 1998). Ainsi, en condition d'hypoxie, *H19* permet la survie et l'adaptation des cellules ; ce qui leur confère un avantage sélectif de croissance.

En plus de la **régulation** transcriptionnelle classique, il a été mis en évidence que la régulation de l'expression d'*H19* peut aussi être **post-transcriptionnelle**. En 2000, Milligan et collaborateurs ont montré que l'accumulation d'ARN *H19* dans les cellules musculaires de souris est due à une stabilisation de celui-ci par une protéine indéterminée (Milligan et al., 2000). Dans le foie fœtal, ce type de régulation avait déjà été décrit (Jouvenot et al., 1999). D'autres études ont montré une association de l'ARN *H19* avec une famille de protéine. En effet, une association d'expression des gènes *H19/IGF2* et de la protéine CRD-BP (*cMyc mRNA coding Region instability Determinant Binding Protein*) a été retrouvée dans les tumeurs mammaires de souris. Cette protéine CRD-BP, en se liant à l'ARN *H19*, régule positivement son expression en le stabilisant ou en affectant sa localisation (Tessier et al., 2004). De plus, 4 sites de fixation pour la protéine IGF-II mRNA Binding Protein (IMP) ont été identifiés sur les exons 4-5 de l'ARN *H19*. Cette famille de protéine est constituée de trois membres : IMP1, 2 et 3 qui sont localisés dans les lamellipodes et les régions périnucléaires des fibroblastes (Runge et al., 2000). Ces protéines sont capables de fixer l'ARNm *IGF-2*. Chez l'Homme, la protéine IMP1 est un orthologue de la protéine CRD-BP (Nielsen et al., 2001). En 2005, il a été mis en évidence que la protéine IMP3 se fixe sur l'ARN *H19 in vivo*. Cependant, l'inactivation de celle-ci par un siRNA n'affecte pas l'expression d'*H19* (Liao et al., 2005). Outre l'empreinte génomique parentale, les mécanismes de régulation de l'expression du gène *H19* sont donc complexes et variés.

d. *H19* dans la tumorigenèse

A l'heure actuelle, la fonction du gène *H19* dans le processus de tumorigenèse est controversée. Cette controverse peut être expliquée par le fait que pendant de nombreuses années, le(s) mécanisme(s) d'action de cet ARN n'étaient pas décrits. De plus, son étude est assez complexe puisque aucune protéine associée à cet ARN n'est identifiée et qu'il est soumis à un mécanisme de régulation complexe et finement régulé. Cependant, les premières études ont mis en évidence que l'absence d'expression du gène *H19* est associée à l'apparition du syndrome de Beckwith Wiedemann ; ce qui lui suggère un rôle de suppresseur de tumeur dans cette pathologie (Reik et al., 1995). De plus, l'absence d'expression d'*H19* induit une prédisposition de l'enfant aux tumeurs rénales : les tumeurs de Wilms (Rainier et al., 1993). Ces deux pathologies sont issues d'une dérégulation de l'empreinte génomique parentale conduisant à une perte d'expression du gène *H19* et à une expression biallélique du gène *IGF2*. Cependant, en parallèle de ces études, il était montré que *H19* est surexprimé dans différents cancers et qu'il participe à la progression tumorale. L'ensemble des études essayant d'identifier le rôle d'*H19* dans les pathologies est regroupé dans le tableau 5.

ONCOGENE		SUPPRESSEUR DE TUMEUR	
organe	références	organe	références
vessie	Biran et al., 1994 Ariel et al., 1995	BWS	Reik et al., 1995
sein	Dugimont et al., 1995 Adriaenssens et al., 1998 Lottin et al., 2002	muscle strié enfant, Rabdomyosarcome	Casola et al., 1997 Lynch et al., 2002
utérus	Douc-rasy et al., 1996 Hashimoto et al., 1997 Tanos et al., 2004 Lottin et al., 2005	tumeurs de Wilms	Raineir et al., 1993 Moulton et al., 1994
ovaire	Kim et al., 1998 Tanos et al., 1999	Foie enfant, hépatoblastomes	Li et al., 1998 Fukuzawa et al., 1999
testicules	Verkerk et al., 1997 Ariel et al., 2000b		
poumon	Kondo et al., 1995		
œsophage/colon	Hibi et al., 1996 Cui et al, 2002 Tsang et al., 2010		
foie	Kim & Lee, 1997 Ariel et al., 1998		

Tableau 5. Liste non exhaustive des travaux montrant le rôle oncogène ou suppresseur de tumeur de l'ARN *H19*.

Récemment, afin d'identifier le rôle d'*H19* dans les cancers, une équipe a utilisé différentes souris transgéniques pour le locus *H19/IGF2* : *H19Δ3* (délétion du gène *H19* sur l'allèle maternel) et *H19ΔEnh* (délétion des enhanceurs en aval du gène *H19* sur l'allèle maternel) qu'ils ont ensuite croisé avec des lignées de souris formant des cancers spontanées : tératome, colon, foie (Yoshimizu et al., 2008). Dans cette étude, les auteurs clarifient le rôle d'*H19* dans les cancers embryonnaire et de l'enfant. En effet, ils mettent en évidence que l'absence d'expression du gène *H19* favorise l'apparition des tumeurs et permet également l'obtention de tumeurs plus grosses pour les trois types de tumeurs étudiées : tératome, tumeur du colon et tumeur du foie. Cependant, ces tumeurs sont obtenues au stade embryonnaire ou 4 mois après la naissance et il est à noter qu'une légère surexpression du gène *IGF2* a été constatée dans les tissus étudiés. Néanmoins, les auteurs ont également réalisé ces études dans des souris *IGF2*^{-/-}, ils ont obtenu des résultats similaires ; ce qui confirme que l'effet observé sur la tumorigénèse est due à la perte d'expression du gène *H19*. Ainsi, il est clairement défini que *H19* possède un rôle de suppresseur de tumeur au stade embryonnaire et après la naissance.

Cependant, dans les cancers de l'adulte, plusieurs études prônent en faveur d'un rôle d'oncogène pour l'ARN *H19*. En effet, il est montré que *H19* est fréquemment surexprimé dans les cancers : poumon (Kondo et al., 1995), foie (Kim & Lee, 1997 ; Ariel et al., 1998), vessie (Biran et al., 1994 ; Ariel et al., 1995), œsophage (Hibi et al., 1996) et ovaire (Kim et al., 1998). Nous avons également mis en évidence que *H19* est surexprimé dans 70% de cancers du sein (Dugimont et al., 1995, Adriaenssens et al., 1998) mais aussi dans les cancers de l'utérus (Ariel et al., 1997b ; Lottin et al., 2005). Dans les cancers du foie, *H19* est même considéré comme étant un marqueur de tumeur (Ariel et al., 1998). Par ailleurs, nous avons montré que *H19*, régulé positivement par le facteur E2F1, favorise la progression du cycle cellulaire et augmente les capacités tumorigènes des cellules cancéreuses mammaires (Lottin et al., 2002 ; Berteaux et al., 2005). De plus, certains auteurs qualifient l'ARN *H19* d'ARN oncofoetal (Ariel et al., 1997a ; 2000a).

Pour finir, plusieurs modes d'action associés à l'ARN *H19* ont été décrits dans les cancers chez l'adulte (voir paragraphe suivant). L'ensemble des modes d'action et des cibles de son microARN permettent la progression tumorale et métastatique : augmentation de la prolifération cellulaire, de la transition épithélio-mésenchymateuse, de la migration/invasion ou encore de la résistance des cellules aux stress environnementaux (l'hypoxie par exemple). Depuis de nombreuses années, il est donc clairement établi que l'ARN *H19* possède un rôle d'oncogène dans les cancers chez l'adulte. Avec l'identification des modes d'action de cet ARN, cette fonction d'oncogène est d'autant plus vraie.

Ainsi, la fonction d'*H19* est différente en fonction du stade de développement, comme pour la régulation de son expression : il est exprimé au stade embryonnaire, il y possède un rôle de suppresseur de tumeur ; chez l'adulte, il est réprimé dans la majorité des tissus, sa ré-expression est associée à un phénotype oncogène. Cette régulation de fonction de l'ARN *H19* est également suggérée par Raveh et collaborateurs (Raveh et al., 2015).

e. Mode d'action de l'ARN *H19*

Comme décrit dans le chapitre précédent, les longs ARN non codant ont différents mode d'action pour exercer leurs fonctions : régulation épigénétique, précurseur de microARN, éponge à microARN ou encore interactant protéique. Pour exercer sa fonction d'oncogène dans la majorité des tumeurs, *H19* possède l'ensemble de ces modes d'action (fig 24).

➤ Précurseur de microARN :

En 2007, Cai et Cullen ont mis en évidence que l'ARN *H19* est le précurseur de microARN : hsa-miR-675-5p et hsa-miR-675-3p (Cai & Cullen, 2007) (fig 24-1). Ces deux microARN sont générés à partir d'une structure tige-boucle contenue dans l'exon1 et conservée entre les espèces. La synthèse de ces microARN utilise la voie canonique de formation des microARN : elle est donc Drosha- et Dicer-dépendante. A l'heure actuelle, plusieurs cibles de ces microARN ont été identifiées dans le tissu sain mais aussi dans du tissu cancéreux. La première cible du miR-675 a été décrite dans des cellules cancéreuses de colon : il s'agit de l'ARNm *Rb* (Tsang et al., 2010). Les auteurs ont mis en évidence que le miR-675-5p est capable de se fixer sur l'UTR3' de cet ARNm, ce qui permet une diminution de protéine Rb dans les cellules. Ils mettent également en évidence une régulation négative entre la protéine Rb et le miR-675 dans les tumeurs colorectales.

Par ailleurs, dans les cellules du placenta, le miR-675-3p est capable de réguler négativement l'expression de l'ARNm *IGF1R* en se fixant sur l'UTR3' (Keniry et al., 2012). Dans cette étude, les auteurs montrent que la synthèse du miR-675 n'est pas constante lors du développement embryonnaire bien que l'expression d'*H19* le soit. En effet, le miR-675-3p n'est exprimé qu'à partir du 11^{ème} jour de développement embryonnaire et dans les cellules du placenta uniquement sans modification de l'expression d'*H19*. L'expression du miR-675 est régulée négativement par la protéine HuR. Cette protéine se fixe sur l'ARN *H19* 55bp en amont du pri-miR-675. En masquant le site de clivage, elle empêche ainsi l'action de Drosha. Au 11^{ème} jour de développement embryonnaire, l'expression de la protéine HuR diminue fortement, le miR-675 est donc synthétisé, il pourra interagir avec ses ARNm cibles dans ce tissu. Les auteurs montrent que le miR-675 régule négativement l'expression de l'*IGF1R* (le récepteur de l'*IGF2*) dans le placenta. Cette régulation permet une diminution de la croissance du placenta : une étape nécessaire permettant de préparer l'accouchement.

Ainsi le miR-675, généré par l'ARN non codant *H19*, contribue au bon développement de l'embryon mais aussi à la progression tumorale. A ce jour, environ 20 cibles du miR-675 ont été identifiées dans plusieurs types cellulaires. Ces cibles est résumé dans le tableau 6.

Les cibles du miR-675 déjà identifiées permettent le développement tumoral. En effet, ces cibles régulent la prolifération cellulaire (Rb, Cdc6), la transition épithélio-mésenchymateuse (Twist1, Slug, Cadherin 11) ou encore la signalisation de facteurs de croissance (Smad1, Smad5, TGFβ1). Grâce à son microARN, *H19* est également capable de réguler l'expression des gènes de façon épigénétique. En effet, le miR-675 réprime l'expression de la protéine HP1α dans les cellules cancéreuses du foie (Li, 2015). Cette régulation est associée à une diminution des marques répressives H3K9me3, H3K27me3 ainsi qu'à une augmentation de la marque activatrice H3K27Ac sur le promoteur du gène codant pour l'EGR1. Cette dernière permet à son tour l'expression du gène *H19*.

Ainsi, grâce à son microARN, *H19* est également impliqué dans la régulation épigénétique et favorise sa propre expression. Bien que *H19* soit impliqué dans la croissance de la glande mammaire et sa tumorigenèse, aucune cible du miR-675 n'a été identifiée dans ce tissu.

cible	position	miR-675	tissus	papier
Rb	3'UTR	5p	cancer colorectal	Tsang, 2010
COL2A1	?	?	chondrocytes	Dudek, 2010
Igf1r	3'UTR	3p	placenta	Keniry, 2012
NOMO1	3'UTR	5p	trophoblaste placentaire	Gao, 2012
Twist1	3'UTR	?	carcinome hépatocellulaire	Hernandez, 2013
MITF	3'UTR	3p	melanome	Kim, 2013
RUNX1	3'UTR	5p	cancer gastrique	Zhuang, 2014
Smad1	3'UTR	3p	muscle squelettique (Souris)	Dey, 2014
Smad5	3'UTR	3p		
Cdc6	3'UTR	5p		
CALN1	3'UTR	?	cancer gastrique	Li, 2014
SLUG	Cible indirecte		cancer ovaire	Matouk, 2014
TGFBI	3'UTR	5p	cancer prostate	Zhu, 2014
Cadherin 11	3'UTR	3p	Mélanome	Kim, 2014
TGFBI	CDS/5'UTR	3p/5p	cellule souches mesenchyme	Huang, 2015
HP1	3'UTR	5p	cancer du foie	Li, 2015
CDK6	3'UTR	3p	gliome	Li, 2015

Tableau 6. Liste des cibles des miR-675-5p et miR-675-3p déjà identifiées.

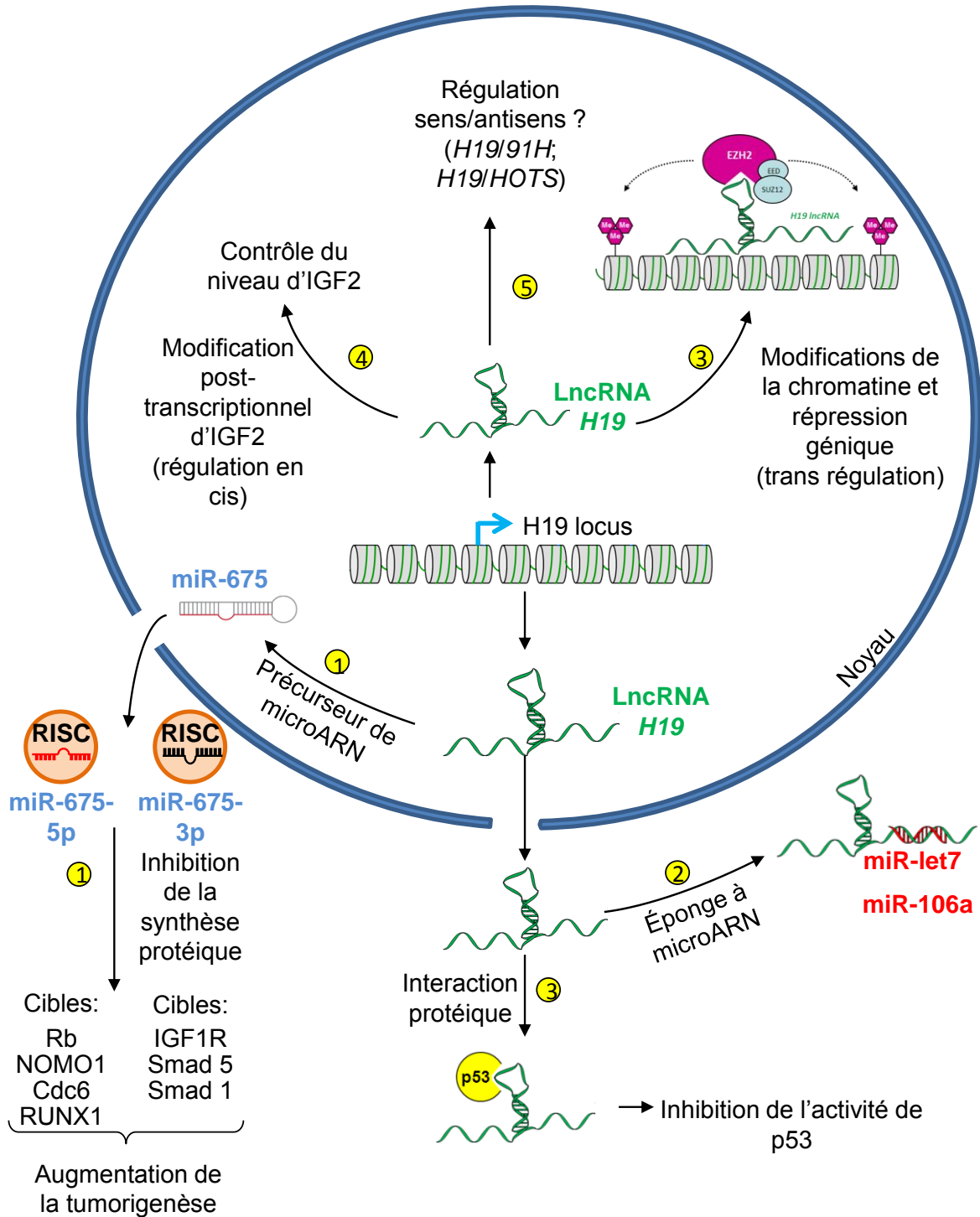


Figure 24. Les différents modes d'action de l'ARN non codant H19. (1) L'ARN H19 est un précurseur de microARN : miR-675-5p et miR-675-3p (Cai et cullen, 2007). Plusieurs cibles ont été identifiées mais aucunes dans les cellules mammaires. (2) L'ARN H19 peut servir d'éponge à microARN pour inhiber leur fonction. (3) L'ARN H19 est capable d'interagir avec des protéines : P53, EZH2. Il empêche la fonction de P53 et il participe à la régulation de l'expression des gènes par modifications épigénétiques. (4) H19 est capable de réguler l'expression d'IGF2 au niveau transcriptionnel et post-transcriptionnel (Vennin et al., 2013). (5) H19 pourrait interagir avec les ARN 91H ou HOTS pour être précurseur de petits ARN ou réguler l'empreinte génomique. (Adapté de Angrand et al., 2015).

➤ Eponge à microARN

Récemment, il a été mis en évidence que l'ARN *H19*, en plus d'être précurseur de microARN, possède également des fonctions d'éponges à microARN (fig 24-2). En 2013, Kallen et collaborateurs ont identifié 4 sites potentiels de fixation pour les microARN let-7 sur l'ARN *H19* (Kallen et al., 2013). Les microARN let-7 sont notamment impliqués dans la régulation de la différenciation des cellules musculaires. Les auteurs montrent que l'interaction *H19/let-7* empêche la fonction de ce dernier. En effet, ils ont observé une augmentation de la différenciation des cellules musculaires en absence de l'ARN *H19*. Ainsi, *H19*, en fixant let-7, régule la fonction de celui-ci dans les cellules. *H19* permet donc de titrer la quantité de microARN let-7 présent et actif dans un tissu. Il contribue à la régulation de l'expression des gènes cibles de ces microARN et par conséquent au bon développement tissulaire.

Par ailleurs, il a été mis en évidence que l'ARN *H19* possède des fonctions d'éponges à microARN dans les cellules cancéreuses de l'ovaire et les myoblastes (Imig et al., 2015). Les auteurs ont utilisé la technologie du miR-CLIP afin d'identifier tous les ARNm cibles du miR-106a. Cette technique consiste à transfecter un miR-106a biotinylé sur l'extrémité 3'. Ensuite, grâce au complexe streptavidine/biotine, les ARNm associés au miR-106a sont immunoprécipités et séquencés. Cette technologie a permis de mettre en évidence que l'ARN *H19* est capable d'interagir avec le miR-106a. Deux éléments de réponse au miR-106a (MRE), fonctionnels dans des cellules ovariennes et des myocytes, ont ainsi été identifiés sur l'exon 1 de l'ARN *H19*. Par ailleurs, ces MRE sont situés à proximité d'un site de fixation des protéines appartenant à la famille IGF2BP. Les auteurs montrent également que l'ARN *H19* est stabilisé en présence du miR-106a ; ils supposent un effet synergique de ce microARN avec les protéines IGF2BP pour le stabiliser. De plus, dans les cellules cancéreuses ovariennes et dans les myocytes, une augmentation de l'expression des cibles du miR-106a a été observée ; ceci permet une augmentation de la prolifération cellulaire. Ainsi, *H19* en fixant les membres de la famille miR-106a, possède un rôle d'éponge à microARN et favorise la prolifération cellulaire.

L'ARN *H19* est également une éponge aux microARN miR-138 et miR-200a dans les cancers colorectaux et miR-141 dans les cancers gastriques (Liang et al., 2015 ; Zhou et al., 2015). Ces trois microARN sont impliqués dans la régulation de l'expression des protéines ZEB1 et ZEB2. Grâce à ces interactions, *H19* favorise la prolifération cellulaire ainsi que la transition épithélio-mésenchymateuse dans ces cancers.

➤ Interactant protéique :

Enfin, l'ARN *H19* est capable de se fixer sur plusieurs protéines et d'orienter leur fonction (fig 24-3). En effet, en 2012, une interaction entre l'ARN *H19* et la protéine P53 a été mise en évidence dans les cancers gastriques et elle est associée à une prolifération cellulaire accrue (Yang et al., 2012). De plus, ils supposent que la fixation d'*H19* sur la protéine P53 réduit la fonction de celle-ci. La protéine P53, activée lors de dommages à l'ADN, possède deux rôles majeurs : (1) induire un arrêt du cycle cellulaire si les dommages sont réparables ; (2) induire l'apoptose cellulaire lors de dommages trop importants. Il s'agit donc d'une protéine « gardien de l'intégrité génomique ». Ainsi, lorsque *H19* empêche la fonction de la protéine P53 dans les cancers gastriques, il favoriserait l'instabilité génomique et par conséquent la progression tumorale.

Par ailleurs, une autre étude a mis en évidence que l'ARN *H19* est capable de se fixer sur la protéine EZH2 qui est présente dans le complexe répresseur polycomb PRC2 (Luo et al., 2013). Cette interaction a été mise en évidence dans les cancers de la vessie. Les auteurs montrent que l'association *H19*/EZH2 permet le recrutement du complexe PRC2 au promoteur du gène codant la E-cadherine. Cette protéine permet l'attachement des cellules à la matrice extracellulaire. En absence d'E-cadherine, les cellules sont capables de réaliser la transition epithelio-mésenchymateuse : une transformation nécessaire à la dissémination des cellules cancéreuses dans l'organisme. Lorsque *H19* est associé à la protéine EZH2, il permet la méthylation de la lysine 9 de l'histone H3 sur le promoteur du gène codant la E-cadhérine. Ce gène n'est donc plus exprimé dans ces tissus, les cellules acquièrent donc des capacités métastatiques. En permettant le recrutement du complexe répresseur polycomb PRC2, l'ARN *H19* est donc impliqué dans la régulation épigénétique de l'expression des gènes.

L'ensemble des données acquises nous montrent que l'ARN *H19* possède plusieurs modes d'action différents dans un même type cellulaire. En effet, dans les cancers gastriques, *H19*, par l'intermédiaire de son microARN, réprime l'expression de la protéine Rb. Il est une éponge aux microARN miR-138 et miR-200a. De plus, il se fixe sur la protéine P53 et réprime sa fonction. Cependant, tous ces modes d'action lui permettent de promouvoir la progression tumorale.

Tous les modes d'action décrits pour l'ARN H19 participent au phénotype oncogène associé à cet ARN. Il est à noter qu'aucun mode d'action n'a encore été décrit dans les cellules mammaires normales ou cancéreuses. Par conséquent, l'ARN H19 ou son microARN possède d'autres modes d'actions et/ou d'autres cibles pour induire la tumorigenèse dans les cellules mammaires.

OBJECTIFS

Les longs ARN non codants possèdent différents mécanismes, spécifiques de leur localisation cellulaire, pour exercer leurs fonctions. En effet, les lncRNA nucléaires régulent l'expression des gènes par modifications épigénétiques, ils régulent également le transport des ARNm dans le cytosol. Dans le cytoplasme, les lncRNA peuvent être des précurseurs ou des éponges à microARN, ils peuvent également se fixer sur des protéines pour en moduler leur fonction ou leur localisation cellulaire. Le locus *H19/IGF2* est constitué, entre autres, de deux longs ARN non codants, *H19* et *91H* dont l'un a été découvert par notre équipe (Berteaux et al., 2008). A mon arrivée au laboratoire, nous savions que *H19* augmentait les capacités tumorigènes des cellules cancéreuses mammaires mais son mode d'action dans ce cancer n'était pas décrit. Par ailleurs, les fonctions et le rôle de *91H* dans les cellules cancéreuses mammaires n'étaient pas encore identifiés.

Mon travail de thèse a donc consisté à **mettre en évidence le(s) mode(s) d'action des ARN non codants *H19* et *91H* au cours de la tumorigenèse de la glande mammaire.**

Pour répondre à cet objectif, dans un premier temps, j'ai recherché des **cibles du microARN** généré par *H19* dans les cellules cancéreuses mammaires. J'ai identifié plusieurs cibles (Cbl, Cbl-b et FADD) que j'ai validées dans les cellules cancéreuses mammaires. J'ai ensuite étudié les phénotypes associés à la modification d'expression de ces cibles. J'ai également mis en évidence que l'ARN ***H19* est capable d'interagir avec la protéine P53**. J'ai donc étudié l'effet de cette association *H19/P53* dans des cellules cancéreuses mammaires. De plus, en collaboration avec une équipe située à l'Institut de Biologie de Lille, j'ai mis en évidence que l'ARN *H19* est impliqué dans l'apparition et le **maintien des cellules souches normales** prostatiques et mammaires.

Pour finir, j'ai étudié les effets d'une invalidation de *91H* dans les cellules cancéreuses mammaires. J'ai, ainsi, mis en évidence les **propriétés oncogéniques de l'ARN non codant *91H***. J'ai également mis en évidence que ***91H* est impliqué dans le maintien de l'empreinte génomique au locus *H19/IGF2***.

RESULTATS

I. Identification de cibles du miR-675 dans les cellules cancéreuses mammaires.

1. Les ARNm c-Cbl et Cbl-b : deux nouvelles cibles du miR-675

Introduction

Dans de nombreux cancers, plusieurs études ont mis en évidence le rôle oncogène de l'ARN *H19* ; notamment dans les cancers de la vessie, des poumons, de l'œsophage. Cependant, pendant de nombreuses années, la fonction de cet ARN était encore inconnue. A ce jour, plusieurs modes d'action associés à l'ARN *H19* ont été mis en évidence dans les cancers. En effet, en 2007, il a été montré que *H19* est le précurseur de microARN (miR-675-5p et miR-675-3p) et plusieurs cibles de ces microARN ont été identifiées dans les cancers. *H19* peut également être une éponge à microARN ou encore interagir avec des protéines pour en moduler leur fonction.

Dans le cancer du sein, nous avons mis en évidence que *H19* est surexprimé dans 70% des cas. La régulation positive du gène *H19* par des facteurs de croissance (HGF par exemple) et des facteurs de transcription (E2F1) favorise les capacités d'invasion et de prolifération des cellules cancéreuses mammaires. Ces données confirment le phénotype oncogène d'*H19* dans le cancer du sein. Cependant, le mode d'action de cet ARN au cours de la tumorigenèse est encore inconnu.

Dans cette étude, nous avons identifié des cibles du microARN généré par *H19* dans les cellules cancéreuses mammaires. Ces cibles sont impliquées dans le maintien de l'activation des récepteurs aux facteurs de croissance et de leurs voies de signalisation. Nous montrons également le rôle oncogène du miR-675 dans les cellules cancéreuses mammaires.

Article 1.

Constance Vennin, Nathalie Spruyt, Fatima Dahmani, Sylvain Julien, François Bertucci, Pascal Finetti, Thierry Chassat, Roland P. Bourette, Xuefen Le Bourhis and Eric Adriaenssens. (2015) *H19* non coding RNA-derived miR-675 enhances tumorigenesis and metastasis of breast cancer cells by downregulating c-Cbl and Cbl-b. *Oncotarget*, 6, 29209–29223.

***H19* non coding RNA-derived miR-675 enhances tumorigenesis and metastasis of breast cancer cells by downregulating c-Cbl and Cbl-b**

Constance Vennin^{1,2}, Nathalie Spruyt³, Fatima Dahmani³, Sylvain Julien^{1,2}, François Bertucci⁴, Pascal Finetti⁴, Thierry Chassat⁵, Roland P. Bourette³, Xuefen Le Bourhis^{1,2} and Eric Adriaenssens^{1,2*}

¹ INSERM U908, 59655 Villeneuve d'Ascq, France.

² University of Lille, 59655 Villeneuve d'Ascq, France.

³ CNRS UMR 8161, 59021 Lille, France.

⁴ Paoli-Calmettes Institute, Aix-Marseille University, 13009, Marseille, France.

⁵ PLETHA, Institut Pasteur Lille, 59019 Lille, France.

* Correspondance: INSERM U908, Cell plasticity and Cancer, Bâtiment SN3, University of Lille, 59655 Villeneuve d'Ascq, France. eric.adriaenssens@univ-lille1.fr

Tel: 00 33 (0)3 20 43 40 37

Fax: 00 33 (0)3 20 43 40 38

Key Words: *H19*, miRNA, Breast cancer, CBL, Tyrosine Kinase Receptor

ABSTRACT

H19 is a long non-coding RNA precursor of miR-675 microRNA. *H19* is increasingly described to play key roles in the progression and metastasis of cancers from different tissue origins. We have previously shown that the *H19* gene is activated by growth factors and increases breast cancer cell invasion. In this study, we established *H19*/miR-675 ectopic expression models of MDA-MB-231 breast cancer cells to further investigate the underlying mechanisms of *H19* oncogenic action. We showed that overexpression of *H19*/miR-675 enhanced the aggressive phenotype of breast cancer cells including increased cell proliferation and migration *in vitro*, and increased tumor growth and metastasis *in vivo*. Moreover, we identified ubiquitin ligase E3 family (c-Cbl and Cbl-b) as direct targets of miR-675 in breast cancer cells. Using a luciferase assay, we demonstrated that *H19*, through its microRNA, decreased both c-Cbl and Cbl-b expression in all breast cancer cell lines tested. Thus, by directly binding c-Cbl and Cbl-b mRNA, miR-675 increased the stability and the activation of EGFR and c-Met, leading to sustained activation of Akt and Erk as well as enhanced cell proliferation and migration. Our data describe a novel mechanism of protumoral action of *H19* in breast cancer.

INTRODUCTION

The *H19* gene, located in human in 11p15.5 locus, is submitted to genomic imprinting. It is expressed only from the maternal allele [1]. It is transcribed by the RNA polymerase II and the transcript is spliced, polyadenylated, capped and exported into the cytosol. However, no protein associated to this transcript has been discovered and Brannan et al. proposed that *H19* RNA functions as a riboregulator [2]. *H19* is highly expressed in the extraembryonic tissues (placenta), the embryo proper, and the fetal tissues. After birth, its expression is repressed even if a basal expression subsists in several tissues including mammary gland, adrenal gland and uterus [3-6]

The role of *H19* in cancer is still matter of debate. It has been proposed that *H19* acts as a tumor suppressor in Wilm's tumors, embryonic rhabdomyosarcoma, and the Beckwith-Wiedemann syndrome [7]. In addition, using *in vivo* mice models of tumorigenesis, a role of tumor suppressor gene has been ascribed to *H19* [8]. However, numerous studies have shown that *H19* is an oncogene in many types of cancers. Indeed, *H19* overexpression is often correlated with poor prognosis in bladder, lung, oesophageal and gastric cancers [9-13]. *H19* exerts its oncogenic activity through different mechanisms. For example, it has been reported that *H19* functions as a Myc-up-regulated gene to potentiate the tumorigenic phenotype of breast and lung cancer cells [14]. More recently, *H19* was described to act as a molecular sponge to regulate the let-7 family of miRNAs [15]. In addition, *H19* is also a precursor for microRNA-675 (miR-675) and generates two mature miRNAs, miR-675-5p (miR-675) and miR-675-3p (miR-675*) [16].

MicroRNAs (miRNAs) are 19- to 25-nucleotide regulatory non-coding RNAs that are initially expressed as hairpin transcripts of primary miRNA under the control of RNA polymerase II. These primary miRNA hairpins are cleaved by two enzymes, Drosha and Dicer, to generate mature miRNAs. Although several mechanisms of gene expression regulation by miRNAs have been demonstrated [17], they mainly repress gene expression at the post-transcriptional level by interacting with 3'UTR of target mRNA.

Recent data indicate that *H19*-derived miR-675 favours tumor progression by repressing the expression of several target genes, including *Rb* in colorectal cancer [12], *Twist1* in hepatocellular carcinoma [18], and *RUNX1* in gastric cancer [19].

We have previously shown that *H19* is overexpressed in 70% of breast cancer [3]. *H19* gene overexpression in mammary epithelial cells promotes tumorigenesis by upregulating thioredoxin, a modulator of signal transduction and potentiator of tumorigenesis [20]. *H19* gene is up-regulated by growth factors such as HGF and by transcription factors such as E2F1 to enhance cell invasion and cell cycle progression [21, 22]. Altogether these findings are in favor of a role of *H19* as an oncogene in breast cancer [23].

In this study, we have examined the role of *H19*-derived miR-675 in controlling the properties of breast cancer cells. Using *in silico* prediction and functional assays, we identified c-Cbl and Cbl-b as direct targets of miR-675. *H19*-miR-675-Cbl increased the expression level of tyrosine kinase receptors and sustained their activation of down-stream signaling pathways. Moreover, miR-675 overexpression increased the aggressive phenotype of breast cancer cells both *in vitro* and *in vivo*. Our findings provide novel mechanistic insights into a critical role for *H19* RNA in breast cancer development and reveal a previously unknown link between *H19*/miR-675, Cbl and tyrosine kinase receptors to enhance breast cancer cell aggressiveness.

RESULTS

***H19*-derived miR-675 targets c-Cbl and Cbl-b through their coding sequences in breast cancer cells**

We have previously demonstrated the oncogenic role of *H19* gene in breast tumorigenesis [20]. *H19* is a precursor of miR-675-5p/miR-675-3p [16], and *H19*-derived miR-675 has been reported to promote tumorigenesis of several cancers including colon and gastric cancers [12, 19]. To investigate the molecular mechanism of oncogenic miR-675 in breast cancer cells, we performed alignment prediction and found that miR-675-5p was aligned with coding sequences of 2 proteins belonging to the ubiquitin ligase E3 protein family: c-Cbl and Cbl-b (Fig 1A and B). Interestingly, putative seeds are located on coding sequence of these two mRNAs and are conserved between human and mouse. Furthermore, analyses of *H19* and *Cbl* family gene expression in breast cancer cell lines [24] showed a negative correlation between *H19* and c-Cbl or Cbl-b (Fig 1C). We then verified the expression of miR-675-5p and c-Cbl/Cbl-b in breast cancer cells overexpressing *H19*. As shown in Fig 1D, enhanced *H19* expression in MDA-MB-231 and MCF-7 breast cancer cell lines was correlated with an increased level of miR-675-5p. Moreover, the levels of c-Cbl and Cbl-b expression decreased significantly in *H19*-overexpressing cells (Fig 1D and E). Together, these data indicate that c-Cbl and Cbl-b may be negatively regulated by *H19* in breast cancer cells.

To confirm the direct regulation of c-Cbl and Cbl-b by *H19*, we cloned coding sequences (CDS) of c-Cbl and Cbl-b mRNAs, each in two fragments (CDS1 and CDS2) into the *Firefly* luciferase reporter vector (pMIR-REPORT) (Fig 1A, 1B). We also cloned CDS mutated on seed sequence and 3'UTR of these mRNAs in the same vector. The mutation is represented on Fig 2A. Each of these DNA constructs were transfected into various breast cancer cells together with either miR-675 mimic, anti-miR-675 (miR-675 inhibitor) or their corresponding controls. As shown in Fig 2B (left panel), miR-675 mimic decreased relative luciferase activity of pMIR-CDS1 c-Cbl of 35-40% compared to microRNA control in MCF-7 and T47D cells. This effect was abolished when seed sequence was mutated. Furthermore, miR-675 had no effect on luciferase activity of pMIR-CDS2 c-Cbl or pMIR-3'UTR c-Cbl plasmids (Fig 2B, left panel). Similar results were

obtained for Cbl-b, (Fig 2B, right panel). MiR-675 mimic decreased relative luciferase activity of pMIR-CDS2 Cbl-b of 30% in the two tested cell lines. The effect of miR-675 mimic on CDS2 was abolished after mutation in seed sequence (mut CDS2). MiR-675 mimic had no effect on luciferase activity of pMIR-CDS1 Cbl-b or pMIR-3'UTR Cbl-b plasmids (Fig 2B, right panel). Contrary to miR-675 mimic, the miR-675 inhibitor was found to increase relative luciferase activity of pMIR-CDS1 c-Cbl (50% increase in T47D, 100% increase in MCF-7 and 90% increase in MDA-MB-231 cells) (Fig 2C, left panel). The miR-675 inhibitor had no effect on luciferase activity of the other constructs including pMIR-mut CDS1, pMIR-CDS2 or pMIR-3'UTR of c-Cbl. Similar results were obtained with Cbl-b plasmid (Fig 2C, right panel). The miR-675 inhibitor increased luciferase activity of only pMIR-CDS2 plasmid in the three tested cells lines, but had no effect on luciferase activity of the other constructs (Fig 2C, right panel). Taken together, our results confirm that miR-675 decreases both c-Cbl and Cbl-b by interacting with coding sequence of these mRNAs and precisely, with seed sequences predicted in the alignment.

***H19*/miR-675 induce up-regulation of tyrosine kinase receptors and activation of the downstream AKT and ERK pathways**

c-Cbl and Cbl-b are well known to be involved in the degradation of tyrosine kinase receptors after their activation by growth factors [25]. To evaluate the role of *H19*/miR-675 in this process, starved MDA-MB-231 cells overexpressing or not *H19* were cultivated in the presence of EGF for different periods of time and then analyzed for the expression levels of EGF receptor (EGFR) by Western blot (Fig 3A). In control cells (mock), EGFR levels decreased progressively upon EGF treatment with nearly half (56%) of the initial EGFR expression present after 7h of culture in the presence of EGF (Fig 3A, left panel). By contrast, the expression levels of EGFR were not modified in cells overexpressing *H19* (Fig 3A, right panel). This suggests that *H19*/miR-675 induced-down regulation of c-Cbl and Cbl-b may contribute to impair the degradation of tyrosine kinase receptors. We then evaluated cell surface expression of EGFR by flow cytometer analysis (Fig 3B). Cells overexpressing *H19* clearly exhibited higher levels of cell surface EGFR compared to mock cells. Interestingly, in contrast to parental (not shown) and mock cells (Fig 3B), *H19*-overexpressing cells, exhibited two subpopulations expressing medium EGFR level (referred as peak **a**) and high EGFR level (referred as peak **b**) (Fig 3B). The two cell subpopulations were sorted by flow cytometry and analysed by qRT-PCR for *H19* expression (Fig 3C). Cells exhibiting a higher level of cell surface EGFR (peak **b**) also expressed higher *H19* expression levels as compared to peak **a** (Fig 3C) suggesting that the expression level of *H19* correlated with cell surface EGFR expression. To further investigate this relationship, we delivered siRNA to knock down *H19* expression in these *H19*-overexpressing MDA-MB-231 cells. Efficiency of siRNA was determined by qRT-PCR (Fig 3E).

As shown in Fig 3D, transient *H19* siRNA expression decreased cell surface EGFR level, thus confirming the ability of *H19* to upregulate EGFR.

We then treated cells with EGF and studied activation of EGFR and its downstream signaling pathways over the time by Western blot and ALPHAscreen®. In MDA-MB-231 breast cancer cells stably overexpressing *H19*, EGF strongly activated EGFR, Akt and Erk compared to parental cells (Fig 4A and B). Interestingly, HGF treatment induced also a stronger activation of its receptor c-Met, Akt and Erk in MDA-MB-231 cells stably overexpressing *H19* compared to parental cells (Supplementary Fig 1A-C). Similarly, *H19*-overexpressing cells exhibited stronger activation of Akt and Erk upon proNGF and NGF treatments (Supplementary Fig 1D and E). When cells were transiently transfected with a plasmid encoding the *H19* gene and then treated with EGF or HGF (Supplementary Fig 2), similar results were obtained in terms of Akt and Erk activation (Supplementary Fig 2).

To further verify if the increased activation of AKT and Erk was mediated by miR-675, we first transfected MDA-MB-231 parental cells with miR-675 mimic and cells overexpressing *H19* with the miR-675 inhibitor or the corresponding controls. Then, we treated cells with EGF and studied Akt and Erk phosphorylation. As shown in Fig 4C, cells transfected with mimic and then stimulated with EGF exhibited an increase of Akt and Erk activation compared to control vector transfected cells. Conversely, in *H19*-overexpressing cells, the miR-675 inhibitor decreased Akt and Erk activation (Fig 4D).

All together, our results indicate that *H19*/miR-675-5p increase the activation of several tyrosine kinase receptors and their downstream signaling pathways in breast cancer cells.

***H19*/miR-675 increase migration and proliferation of breast cancer cells and enhance the effects of growth factors**

To further determine the functional impact of Akt and Erk activation in cells overexpressing *H19* and miR-675, we first evaluated the migratory capacity of *H19*-overexpressing cells using transwell and wound-healing assays. As shown in Fig 5A and B, *H19*-overexpressing MDA-MB-231 cells exhibited an increased migratory capacity compared to control cells. Moreover, *H19* overexpression further increased HGF-induced migration of MDA-MB-231 cells (Supplementary Fig 3A). Similarly, *H19* gene overexpression increased also EGF-induced proliferation of MCF-7 cells (Supplementary Fig 3B).

To evaluate the role of miR-675 independently of *H19* mRNA, we first evaluated the migratory ability of parental MDA-MB-231 transiently transfected with miR-675 mimic. We found that miR-675 mimic increased cell migration as revealed by wound healing assay (Fig 5C). This indicates that miR-675 *per se* is able to increase cell mobility. We then generated cells stably overexpressing miR-675. For this, we introduced in MDA-MB-231 cells a plasmid encoding Green Fluorescent Protein (GFP) fused with miR-675 precursor sequence. As shown in Fig

5D, the two selected clones (miR cl1 and miR cl2) expressed more than 5-fold of miR-675 when compared to mock cells. As expected, the protein levels of c-Cbl and Cbl-b were decreased in the two selected clones (Fig 5E). Importantly, these clones exhibited stronger migratory and proliferative capacities when compared to control cells (Fig 5F and G). Together, these data confirm that miR-675 increases cell migration and proliferation independently of *H19*.

The miR-675 enhances the tumorigenicity and metastatic potential of breast cancer cells

Since ectopic expression of miR-675 increased migration and proliferation of breast cancer cells *in vitro*, we examined whether overexpression of miR-675 could enhance tumor growth and metastasis *in vivo* by subcutaneously injecting cells into immunodeficient SCID mice. As shown in Fig 6A, tumor volume was increased in mice injected with miR-675-overexpressing clones as compared to control mice injected with parental cells. Accordingly, an increased cell proliferation (PCNA staining) and a decreased apoptosis (TUNEL detection) were found in sections of tumors formed by miR-675-overexpressing cells (Fig 6B, middle and right panels) as compared to those formed by parental cells (Fig 6B, left panels). Of note, the protein levels of c-Cbl and Cbl-b were decreased in tumor formed by miR-675-overexpressing cells compared to control (Fig 6C). We then analyzed metastases in xenografted mice thanks to GFP expression in MDA-MB-231 cells. As shown in Fig 6D and E, more GFP-positive cells were found in brain, liver and lungs of mice xenografted with miR-675 overexpressing cells compared to control (Fig 6D, E). Taken together, these results indicate that miR-675 increases primary tumor formation and promotes metastasis.

DISCUSSION/CONCLUSION

It has become increasingly clear that *H19* RNA plays essential role in tumor development and that the *H19* gene is regulated by a complex interplay of both extrinsic and intrinsic factors. We have already demonstrated that *H19* expression is positively regulated by several growth factors such as HGF, EGF and FGF-2 [21]. HGF induces *H19* expression *via* the ERK/MAPK and phospholipase C pathways. Here, we demonstrated that *H19* induced up-regulation of tyrosine kinase receptors including EGFR and c-Met as well as the activation of their downstream Akt and ERK signaling pathways. Our findings indicate that the positive feedback between *H19* expression and growth factors may be of importance in promoting breast cancer development.

More recently, *H19* has been found to encode miR-675. Several targets of miR-675 have been identified in cancers from different tissues. These include Twist1, CALN1, TGF β 1, and Cadherin11 [13, 18, 26, 27]. While this study was underway, it was shown that miR-675

regulates negatively Rb expression by interacting with its 3'UTR mRNA in colorectal cancer and hepatocellular carcinoma [12, 18]. However, we did not find any regulation of Rb by miR-675 in breast cancer cells (Supplementary Fig. 4). Our results together with the previous identification of miR-675 targets indicate that miR-675 may function in a tissue-specific manner and that the increased proliferation of *H19*-overexpressing breast cancer cells cannot be explained by a negative regulation of Rb. Indeed, we demonstrated that miR-675 decreased the expression of c-Cbl and Cbl-b. MiR-675 interacted directly with c-Cbl and Cbl-b coding sequences to prevent mRNA translation. Among the reported targets of miR-675, c-Cbl and Cbl-b were the first ones found to interact with miR-675 *via* their coding sequences.

In mammals, three homolog of Cbl exist: c-Cbl, Cbl-b and Cbl-c. The protein structures of c-Cbl and Cbl-b are similar; they consist of a N-terminal tyrosine kinase-binding domain, a ring finger motif, a proline-rich region and a c-terminal ubiquitin-associated domain that overlap leucine zipper motif. The Cbl-c lacks the c-terminal domain (proline-rich region and ubiquitin-associated domain). Cbl proteins are involved in the regulation of actin skeleton, lymphocyte signaling and downregulation of tyrosine kinase receptors [28]. Cbl proteins negatively regulate tyrosine kinase receptors by interacting with Grb2 to prevent Grb2/SOS association. Cbl can also recognize activated receptors and induce their downregulation by lysosomal degradation. In breast cancer cells, it has been demonstrated that c-Cbl is associated with two other proteins to form the tripartite complex Cdc42, p85Cool-1/ β Pix, and c-Cbl [29]. Upon EGFR activation, Cdc42 interacts with c-Cbl within the tripartite complex to repress the degradation of the activated receptor initiated by c-Cbl. In lung cancer cells, c-Cbl inactivating-mutation increases viability and migration of cells [30]. Similarly, in acute myeloid leukemia, c-Cbl mutations decrease EGFR ubiquitinylation, leading to signaling activation, cell proliferation and survival. Moreover, c-Cbl has been proposed to act as a tumor suppressor since c-Cbl null mice develop invasive cancer (notably, juvenile myelomonocytic leukemia) with complete penetrance [31]. In accordance with these data, we demonstrated that miR-675-induced decrease of c-Cbl and Cbl-b was associated with a more aggressive phenotype of breast cancer cells; miR-675-overexpressing cells presented increased levels of EGFR and c-Met (not shown), increased activation of the downstream signaling pathways as well as increased tumor growth and metastasis. In agreement with our data, Matouk et al demonstrated that miR-675 indirectly targets slug leading to increase of cell invasion and *in vivo* metastasis [32].

On the other hand, Berberine (an isoquinoline alkaloids used in gastroenteritis, type II diabetes, hypertension and arrhythmia) inhibits proliferation of colon cancer cells by stimulating c-Cbl activation and enhancing EGFR degradation [33]. Similarly, another drug, icotinib, activates Cbl-b protein to downregulate EGFR, and induces apoptosis and G1 phase arrest in non-small-cell lung cancer cells [34]. Thus, cancer therapeutic agents can induce downregulation of

tyrosine kinase receptors by activating Cbl proteins. However, cancer resistance to chemotherapy is also frequently associated with the overexpression of tyrosine kinase receptors, which can serve as a link between tumor cells and the microenvironment [35]. The activation of tyrosine kinase receptors by specific growth factors in tumor microenvironment can protect tumor cells from drug-induced damages. Our findings suggest that drug resistance of *H19*-overexpressing cancer cells could be due to Cbl downregulation by *H19*-derived miR-675. Supporting this hypothesis, we found that high expression of miR-675 combined with low expression of c-Cbl identified a subset of ER-negative breast cancer tumors of poorer prognosis (Supplementary Fig. 5). This was found by exploiting the matched mRNA and microRNA global expression profiling of a cohort of 207 primary tumors [36] and, given the rarity of such cohort, will need to be confirmed by future similar studies.

An antisense transcript, *91H*, has been found in the *H19* locus [37, 39]. We verified if this long non coding RNA can act as a miR sponge and impair miR-675 effects. *91H* expression does not vary in miR-675 overexpressing cells. In addition, in *91H* knockdown cells, c-Cbl and Cbl-b expression is stable (Supplementary Fig 6), indicating *91H* is not involved in c-Cbl and Cbl-b regulation.

In conclusion, we identified ubiquitin ligase E3 family (c-Cbl and Cbl-b) as direct targets of miR-675 in breast cancer cells. Moreover, by directly targeting c-Cbl and Cbl-b, miR-675 increased the stability and the activation of EGFR and c-Met leading to subsequent activation of Akt and Erk as well as enhanced cell proliferation and migration. Our findings provide novel mechanistic insights into a critical role for *H19* RNA in breast cancer development, and reveal a link between *H19*/miR-675, Cbl and tyrosine kinase receptors to enhance breast cancer cell aggressiveness.

MATERIALS AND METHODS

Cells culture

The MCF-7 and T47D estrogen-sensitive and the MDA-MB-231 estrogen-insensitive breast cancer cell lines were obtained from the American Type Culture Collection and maintained routinely in Roswell Park Memorial Institute medium (RPMI, Gibco) containing 10% of foetal bovine serum (FBS) and 0.01% of Zell Shield (Minerva Biolabs). Cell lines were cultured at 37°C with 5% CO₂, 95% of air in humidified atmosphere.

Migration assays

Cell migration was determined by transwell assay and wound healing assay. For transwell assay, 3x10⁴ cells were seeded on collagen (1/100, Millipore) coated insert (0.8µm, BD Bioscience) of 6 well-plates in RPMI containing 10% FBS or 0.1 % FBS in the presence of

Résultats

HGF. At the end of the experiments, cells migrating to the other side of the filter were stained with 0.5% crystal violet and counted. Wound healing test was performed by using Culture-inserts for Live Cell Analysis (Ibidi). A total of 3×10^4 cells were plated on each compartment of insert and cultured for 5h. The insert was then removed (0h) and cells were further cultured for 24h. Cells were photographed at 0h and after 24h of culture to record the wound width.

Proliferation assays

A total of 500 cells were plated in each well of a 96-well plate. Cell proliferation was determined using MTT test every day as previously described [12].

Western blot analysis

Cells were lysed in RIPA buffer containing protease and phosphatase inhibitors (protease inhibitor P8340; phosphatase inhibitor cocktail 2 P5726, Sigma-Aldrich). Proteins were quantified with BCA protein assay (Pierce) and then reduced in NuPAGE LDS Sample buffer (Invitrogen) with NuPAGE Reducing Agent (Invitrogen) at 70°C for 10 min. Proteins were separated on SDS-PAGE 4-12% (Invitrogen) and transferred onto PolyVinylidene Fluoride (PVDF) membrane (Millipore). After saturation in PBS 0.2% of casein, membranes were incubated with primary antibodies overnight at 4°C. References of antibodies are listed in the supplemental Table S1. Membranes were washed with PBS 0.5% Tween for 30 min and incubated with secondary antibodies conjugated with Horse Radish Peroxidase (HRP) for 2 h at room temperature. Membranes were analysed with SuperSignal west Dura Chemiluminescence Substrate (Pierce).

Quantification of protein phosphorylation by ALPHAscreen®

Quantification of protein phosphorylation was performed using ALPHAscreen® SureFire technology (Perkin Elmer). This technology allows protein activation quantification in low volume of lysates. Briefly, cells were lysed in appropriate buffer completed with protease and phosphatase inhibitors and lysates were clarified by centrifugation. Lysates were incubated with different antibodies and acceptor beads for 2 h at room temperature on 384-well plates. Then donor beads were added for 2 h and light emission was measured on EnSpire Alpha (Perkin Elmer). The anti-p-Erk Thr202/Tyr204 (TGRESHV100), anti-total Erk TGRTES500, anti-p-Akt ser473 (TGRA4S500), anti-total Akt (TGRTAPS500), anti-p-Met Tyr1234/1235 (TGRCMS500), and anti-total Met (AL281C) antibodies were from Perkin Elmer.

RNA extraction, reverse transcription and Real-time RT-PCR

RNA extraction and qRT-PCR was performed as previously described [37]. Primers used for qRT-PCR are described in supplemental Table S2.

MiRNA was extracted with mirVANA™ miRNA isolation kit (Life Technologies). For the detection, a total of 50 ng RNA was used in the reverse transcription reaction (miRCURY LNA™ Universal RT) (Exiqon). Quantitative PCR was performed by using Exilent SYBR Green master mix (Exiqon) and hsa-miR-675-5p LNA PCR primer set, UniRT (Exiqon). Human U6

RNA was used as an internal control. Fold change of miRNA expression was calculated by the equation $2^{-\Delta\Delta Ct}$.

Construction of plasmids

Coding DNA Sequence (CDS) and 3'UTR of c-Cbl and Cbl-b mRNA were cloned in the pMIR-REPORT luciferase vector (Ambion). Because of their length, CDS were cloned in two fragments named CDS1 and CDS2. All fragments of Cbl-b and 3'UTR of c-Cbl were cloned between *SpeI* and *HindIII* enzyme site. CDS1 of c-Cbl was cloned between *SpeI* and *PmeI* and CDS2 between *SpeI* and *MluI* (New England Biolabs). Fragments were generated by PCR (primers are listed in supplemental Table S2). Point mutations in seed sequence were generated by PCR.

H19 gene was cloned in pcDNA3.1 (-) (Invitrogen) between *NotI* and *BamHI* (New England Biolabs). The normalizing vector pRL-null has no promoter sequence to drive expression of the *Renilla* luciferase gene and was purchased from Promega.

MiR-675-5p was amplified by PCR and cloned in pEGFP-C1 plasmid (Clontech) between *HindIII* and *BamHI* (New England Biolabs). Plasmid productions were performed in *E. Coli* TOP 10 (Invitrogen). Plasmid extraction was performed by using Nucleobond PC100 (Macherey-Nagel).

H19 gene and siH19 transient transfection

For transient transfection, a total of 1.3×10^5 cells were plated on 6-well plates. After 24h, cells were transfected with $1 \mu\text{g}$ of pcDNA3.1-*H19* DNA using Exgen-500 (Euromedex) according to the manufacturer's instructions. Cells were incubated at 37°C for 6h, and then cultured in medium used routinely. *H19* siRNA (Table S2) were transfected with DharmaFECT Duo according to manufacturer's guidelines (Thermo Fischer Scientific Dharmacon).

Establishment of cells overexpressing H19 or miR675

To establish cells overexpressing *H19*, MDA-MB-231 and MCF-7 breast cancer cell lines were transfected with $1 \mu\text{g}$ of pcDNA3.1-*H19* or pcDNA3.1 empty vector as a control using Exgen-500 and cells were allowed to recover for 48h. Cells were then selected in the presence of 1mg/ml G418 (Sigma) for at least one month. To establish cells overexpressing miR675, MDA-MB-231 cells were transfected with $1 \mu\text{g}$ of pEGFP-C1 /miR675 or pEGFP-C1 empty vector using Exgen-500 for 48h. Cells were then selected in the presence of 1mg/ml G418 (Sigma) for at least one month. Among numerous clones obtained, two of them were randomly chosen for *in vitro* and *in vivo* experiments.

MicroRNA transfection and luciferase activity assay

A total of 1×10^5 cells were plated on 12-well plates for 24h. Hsa-miR-675 mimic or its hairpin inhibitor (Thermo Fischer Scientific Dharmacon) were transfected with DharmaFECT-Duo according to manufacturer's guidelines (Thermo Fischer Scientific Dharmacon). Briefly, 75 nM mimic or hairpin inhibitor (anti-miR) were transfected with 500 ng of *Firefly* luciferase plasmid

and 12 ng of pRL-null plasmid expressing Renilla luciferase to monitor the transfection efficiency. The luciferase activity was measured 24h after transfection by using the dual-luciferase reporter assay system according to manufacturer's instructions (Promega). *Firefly* luciferase was normalized with *Renilla* luciferase.

Flow cytometer analysis of membrane EGFR

A total of 2×10^5 cells were plated on 6 well-plates and cultured in RPMI containing 1% FBS. After 48h, cells were isolated by trypsinization using trypsin-versene EDTA solution, then incubated with an anti-EGFR antibody (EGFR-AF488, 1/100, SC-120, Santa Cruz) for 30 min at 4°C before analysis on flow cytometer (Calibur II or ARIA II, Becton Dickinson).

Tumorigenesis in SCID mice

Female SCID mice (8-weeks-old) were purchased from Pasteur Institute, Lille and kept under pathogen-free condition. Animals were handled in accordance with the European Communities Recommendations for Animals Experimentation. Exponentially growing cells were harvested, resuspended in PBS, and subcutaneously (2×10^6 cells in 150 μ l) injected into the flank fat pad of each mouse. Tumor volume and metastasis were monitored as previously described [38].

Analysis of cell proliferation and apoptosis in xenografted tumors

Tumors were fixed, paraffin embedded and cut on 5 μ m tumor sections. Cell proliferation was measured by immunostaining with an anti-PCNA antibody (Santa Cruz, sc-56 HRP). Apoptosis was measured using *in situ* cell death detection kit, POD (Roche), according to manufacturer's instructions.

Statistical analysis

Data are expressed as mean values \pm standard error of the mean of at least 3 independent experiments. The statistical analysis was done by using Student's *t*-test and *p* value < 0.05 was considered significant.

ACKNOWLEDGMENTS

We thank Helene Bauderlique-Le Roy from the flow cytometry platform of the BiImaging Center of Lille (BICeL). We also thank the animal facility at the Pasteur Institute of Lille (PLETHA, Dir: J.P. de Cavel) for animal housing and technical advices.

DISCLOSURE OF POTENTIAL CONFLICTS OF INTEREST

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

GRANT SUPPORT

This work was supported by INSERM and grant from INCA (PLBio 2010-180). C.V. was supported by doctoral fellowships from the University of Lille2 and by Fondation ARC fellowships. F.D. was supported by fellowships from INCA (PLBio 2010-180). S.J. was supported by post-doctoral fellowships from the “Région Nord-Pas de Calais”.

REFERENCES

1. Bartolomei MS, Zemel S, Tilghman SM. Parental imprinting of the mouse H19 gene. *Nature* 1991; 351: 153–155.
2. Brannan CI, Dees EC, Ingram RS, Tilghman SM. The product of the H19 gene may function as an RNA. *Mol Cell Biol* 1990; 10: 28–36.
3. Adriaenssens E, Dumont L, Lottin S, Bolle D, Leprêtre A, Delobelle A, Bouali F, Dugimont T, Coll J, Cury JJ. H19 overexpression in breast adenocarcinoma stromal cells is associated with tumor values and steroid receptor status but independent of p53 and Ki-67 expression. *Am J Pathol* 1998; 153: 1597–1607.
4. Adriaenssens E, Lottin S, Dugimont T, Fauquette W, Coll J, Dupouy JP, Boilly B, Cury JJ. Steroid hormones modulate H19 gene expression in both mammary gland and uterus. *Oncogene* 1999; 18: 4460–4473.
5. Liu J, Kahri AI, Heikkilä P, Ilvesmäki V, Voutilainen R. H19 and insulin-like growth factor-II gene expression in adrenal tumors and cultured adrenal cells. *J Clin Endocrinol Metab* 1995; 80: 492–496.
6. Ariel I, Weinstein D, Voutilainen R, Schneider T, Lustig-Yariv O, de Groot N, Hochberg A. Genomic imprinting and the endometrial cycle. The expression of the imprinted gene H19 in the human female reproductive organs. *Diagn Mol Pathol* 1997; 6: 17–25.
7. Soejima H, Higashimoto K. Epigenetic and genetic alterations of the imprinting disorder Beckwith-Wiedemann syndrome and related disorders. *J Hum Genet* 2013; 58:402-9
8. Yoshimizu T, Miroglio A, Ripoché MA, Gabory A, Vernucci M, Riccio A, Colnot S, Godard C, Terris B, Jammes H, Dandolo L. The H19 locus acts in vivo as a tumor suppressor. *Proc Natl Acad Sci U S A*. 2008; 105: 12417-12422.

9. Ariel I, Lustig O, Schneider T, Pizov G, Sappir M, De-Groot N, Hochberg A. The imprinted H19 gene as a tumor marker in bladder carcinoma. *Urology* 1995; 45: 335–338.
10. Kondo M, Suzuki H, Ueda R, Osada H, Takagi K, Takahashi T, Takahashi T. Frequent loss of imprinting of the H19 gene is often associated with its overexpression in human lung cancers. *Oncogene* 1995; 10: 1193–1198.
11. Hibi K, Nakamura H, Hirai A, Fujikake Y, Kasai Y, Akiyama S, Ito K, Takagi H. Loss of H19 imprinting in esophageal cancer. *Cancer Res* 1996; 56: 480–482.
12. Tsang WP, Ng EKO, Ng SSM, Jin H, Yu J, Sung JJY, Kwok TT. Oncofetal H19-derived miR-675 regulates tumor suppressor RB in human colorectal cancer. *Carcinogenesis* 2010; 31: 350–358.
13. Li H, Yu B, Li J, Su L, Yan M, Zhu Z, Liu B. Overexpression of lncRNA H19 enhances carcinogenesis and metastasis of gastric cancer. *Oncotarget* 2014; 5: 2318-29.
14. Barsyte-Lovejoy D, Lau SK, Boutros PC, Khosravi F, Jurisica I, Andrulis IL, Tsao MS, Penn LZ. The c-Myc oncogene directly induces the H19 noncoding RNA by allele-specific binding to potentiate tumorigenesis. *Cancer Res* 2006; 66: 5330-7.
15. Kallen AN, Zhou XB, Xu J, Qiao C, Ma J, Yan L, Lu L, Liu C, Yi JS, Zhang H, Min W, Bennett AM, Gregory RI, et al. The imprinted H19 lncRNA antagonizes let-7 microRNAs. *Mol Cell* 2013; 52: 101-12.
16. Cai X, Cullen BR. The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA* 2007; 13: 313–316.
17. Cipolla GA. A non-canonical landscape of the microRNA system. *Front Genet.* 2014 5; 337.
18. Hernandez JM, Elahi A, Clark CW, Wang J, Humphries LA, Centeno B, Bloom G, Fuchs BC, Yeatman T, Shibata D. miR-675 Mediates Downregulation of Twist1 and Rb in AFP-Secreting Hepatocellular Carcinoma. *Ann Surg Oncol* 2013. doi:10.1245/s10434-013-3106-3.
19. Zhuang M, Gao W, Xu J, Wang P, Shu Y. The long non-coding RNA H19-derived miR-675 modulates human gastric cancer cell proliferation by targeting tumor suppressor RUNX1. *Biochem Biophys Res Commun* 2014; 448: 315–322.

20. Lottin S, Adriaenssens E, Dupressoir T, Berteaux N, Montpellier C, Coll J, Dugimont T, Curgy JJ. Overexpression of an ectopic H19 gene enhances the tumorigenic properties of breast cancer cells. *Carcinogenesis* 2002; 23: 1885–1895.
21. Adriaenssens E, Lottin S, Berteaux N, Hornez L, Fauquette W, Fafeur V, Peyrat JP, Le Bourhis X, Hondermarck H, Coll J, Dugimont T, Curgy JJ. Cross-talk between mesenchyme and epithelium increases H19 gene expression during scattering and morphogenesis of epithelial cells. *Exp Cell Res* 2002; 275: 215–229.
22. Berteaux N, Lottin S, Monté D, Pinte S, Quatannens B, Coll J, Hondermarck H, Curgy JJ, Dugimont T, Adriaenssens E. H19 mRNA-like noncoding RNA promotes breast cancer cell proliferation through positive control by E2F1. *J Biol Chem* 2005; 280: 29625–29636.
23. Vennin C, Dahmani F, Spruyt N, Adriaenssens E. Role of long non-coding RNA in cells: Example of the H19/IGF2 locus. *Advances in Bioscience and Biotechnology* 2013; 04: 34–44.
24. Charafe-Jauffret E, Ginestier C, Monville F, Finetti P, Adélaïde J, Cervera N, Fekairi S, Xerri L, Jacquemier J, Birnbaum D, Bertucci F. Gene expression profiling of breast cell lines identifies potential new basal markers. *Oncogene* 2006; 25: 2273-2284.
25. Mohapatra B, Ahmad G, Nadeau S, Zutshi N, An W, Scheffe S, Dong L, Feng D, Goetz B, Arya P, Bailey TA, Palermo N, Borgstahl GE, et al. Protein tyrosine kinase regulation by ubiquitination: critical roles of Cbl-family ubiquitin ligases. *Biochim Biophys Acta* 2013; 1833:122-39.
26. Zhu M, Chen Q, Liu X, Sun Q, Zhao X, Deng R, Wang Y, Huang J, Xu M, Yan J, Yu J. lncRNA H19/miR-675 axis represses prostate cancer metastasis by targeting TGFBI. *FEBS J* 2014; 281: 3766-75.
27. Kim NH, Choi SH, Lee TR, Lee CH, Lee AY. Cadherin 11, a miR-675 target, induces N-cadherin expression and epithelial-mesenchymal transition in melasma. *J Invest Dermatol* 2014; 134: 2967-76.
28. Schmidt MHH, Dikic I. The Cbl interactome and its functions. *Nat Rev Mol Cell Biol* 2005; 6: 907–918.

Résultats

29. Hirsch DS, Shen Y, Wu WJ. Growth and motility inhibition of breast cancer cells by epidermal growth factor receptor degradation is correlated with inactivation of Cdc42. *Cancer Res* 2006; 66: 3523–3530.
30. Tan Y-HC, Krishnaswamy S, Nandi S, Kanteti R, Vora S, Onel K, Hasina R, Lo FY, El-Hashani E, Cervantes G, Robinson M, Hsu HS, Kales SC et al. CBL is frequently altered in lung cancers: its relationship to mutations in MET and EGFR tyrosine kinases. *PLoS ONE* 2010; 5: e8972.
31. Ogawa S, Shih L-Y, Suzuki T, Otsu M, Nakauchi H, Koeffler HP, Sanada M. Deregulated intracellular signaling by mutated c-CBL in myeloid neoplasms. *Clin Cancer Res* 2010; 16: 3825–3831.
32. Matouk IJ, Raveh E, Abu-lail R, Mezan S, Gilon M, Gershtain E, Birman T, Gallula J, Schneider T, Barkali M, Richler C, Fellig Y, Sorin V et al. Oncofetal H19 RNA promotes tumor metastasis. *Biochim Biophys Acta*. 2014; 1843: 1414-26.
33. Wang L, Cao H, Lu N, Liu L, Wang B, Hu T, Israel DA, Peek RM-Jr, Polk DB, Yan F. Berberine inhibits proliferation and down-regulates epidermal growth factor receptor through activation of Cbl in colon tumor cells. *PLoS ONE* 2013; 8: e56666.
34. Mu X, Zhang Y, Qu X, Hou K, Kang J, Hu X, Liu Y. Ubiquitin ligase Cbl-b is involved in icotinib (BPI-2009H)-induced apoptosis and G1 phase arrest of EGFR mutation-positive non-small-cell lung cancer. *Biomed Res Int* 2013; 2013: 726375.
35. Corso S, Giordano S. Cell-autonomous and non-cell-autonomous mechanisms of HGF/MET-driven resistance to targeted therapies: from basic research to a clinical perspective. *Cancer Discov* 2013; 3: 978–992.
36. Buffa FM, Camps C, Winchester L, Snell CE, Gee HE, Sheldon H, Taylor M, Harris AL, Ragoussis J. microRNA-associated progression pathways and potential therapeutic targets identified by integrated mRNA and microRNA expression profiling in breast cancer. *Cancer Res* 2011; 71: 5635–5645.
37. Berteaux N, Aptel N, Cathala G, Genton C, Coll J, Daccache A, Spruyt N, Hondermarck H, Dugimont T, Curgy JJ, Forné T, Adriaenssens E. A novel H19 antisense RNA

overexpressed in breast cancer contributes to paternal IGF2 expression. *Mol Cell Biol* 2008; 28: 6731–6745.

38. Lagadec C, Meignan S, Adriaenssens E, Foveau B, Vanhecke E, Romon R, toillon RA, Oxombre B, Hondermarck H, Le Bourhis X. TrkA overexpression enhances growth and metastasis of breast cancer cells. *Oncogene* 2009; 28: 1960–1970.

39. Tran VG, Court F, Duputié A, Antoine E, Aptel N, Milligan L, Carbonell F, Lelay-Taha MN, Piette J, Weber M, Montarras D, Pinset C, Dandolo L et al. H19 antisense RNA can up-regulate Igf2 transcription by activation of a novel promoter in mouse myoblasts. *PLoS One*. 2012;7: e37923.

LEGENDS

Figure 1. *H19/miR675* downregulated *c-Cbl* and *Cbl-b* expression in breast cancer cells.

(A, B) Alignment prediction of miR-675-5p on *c-Cbl* and *Cbl-b* mRNA. Relative positions are indicated in bp. Note interaction of miR-675 on *Cbl* mRNA is conserved in human and mouse. Coding sequences of these mRNA are too long, so we cloned them in pMiR-REPORT luciferase in two parts named CDS1 and CDS2. The artificial break is represented by the dotted line. (C) Negative correlation between *H19* and *c-Cbl/Cbl-b* expression in breast cancer cell lines [24]. Relative expression of *H19* and *c-Cbl* or *Cbl-b* in (1) MDA-MB-361, (2) MDA-MB-134, (3) SUM225, (4) T47D, (5) S68, (6) SUM159, (7) MCF-7, (8) ZR-75-30 and (9) BT483. (D) QRT-PCR analysis of expression of *H19*, miR-675-5p, *c-Cbl* and *Cbl-b* in breast cancer cell lines. Results are presented as relative levels compared to MDA-MB-231 mock cells (indexed to 1). Data represent mean of three independent experiments and error bar sem.* $p < 0.05$; ** $p < 0.005$; *** $p < 0.001$. (E) Western blot analysis of *c-Cbl* and *Cbl-b* levels in MDA-MB-231 and MCF-7 breast cancer cells. β -actin was used as a loading control.

Figure 2. MiR-675 directly targeted on *c-Cbl* and *Cbl-b* through interaction with coding sequences.

(A) Alignment prediction of miR-675-5p on *c-Cbl* and *Cbl-b* mRNA. Mutation in seed sequence is represented in bold. Coding sequence (CDS) and 3'UTR of *c-Cbl* and *Cbl-b* mRNA were cloned into pMIR-REPORT luciferase. Coding sequence of these mRNA are so long that we cloned them in pMIR-REPORT luciferase in two parts named CDS1 and CDS2. (B and C) The firefly luciferase activity in breast cancer cells after cotransfection with reporter construct and miR-675 mimic or miR-675 inhibitor (α mir). The luciferase activity was measured by dual-luciferase reporter assay (Promega) and was normalized to Renilla luciferase activity. Plasmids were transfected with mimic or antimir, or their controls in breast cancer cell lines

(T47D, MCF-7 and MDA-MB-231). Data represent mean of three independent experiments versus their respective controls in percentage and error bar sem. * $p < 0.05$; ** $p < 0.005$.

Figure 3. EGF-induced EGFR downregulation was prevented in *H19* overexpressing cells. (A) MDA-MB-231 control (mock) and *H19*-overexpressing cells (H19) were treated with EGF (50 ng/ml) and proteins were extracted at indicated times. Total EGFR levels were determined by western blot. The intensities of bands were quantified by densitometry (multigauge, Fujifilm), and the results obtained for EGFR expression during the time-course were compared to those obtained in control (indexed to 1) after normalization to actin expression. (B) Flow cytometer analysis of membrane EGFR in MDA-MB-231 control cells (mock) and *H19*-overexpressing cells (H19). (C) Subpopulations of *H19* overexpressing cells with medium and high expression levels of cell surface EGFR (peaks a and b in Fig 3B) were sorted by FACS, and *H19* expression was then quantified by qRT-PCR. Results are presented as relative levels in cells overexpressing *H19* compared to control. * $p < 0.05$; ** $p < 0.005$. (D) *H19*-overexpressing MDA-MB-231 cells were transiently transfected with siRNA-*H19*, and membrane EGFR levels were then determined by flow cytometry analysis. (E) Relative expression of *H19* determined by qRT-PCR in breast cancer cells overexpressing *H19* transfected with siRNA-*H19* or siRNA-GFP as a control. * $p < 0.05$

Figure 4. EGF-induced Akt and Erk phosphorylation was enhanced in *H19*/miR675-overexpressing cells. (A, B) Control (mock) and *H19*-overexpressing cells (H19) were treated with 10 ng/ml EGF and proteins were extracted at indicated times. Akt and Erk activation was determined by Western blot analysis (A) and ALPHAscreen® (B). (C) Parental MDA-MB-231 cells (ctrl) were transfected with miR-675 (mimic) for 72 h. Then, cells were treated with 10 ng/ml EGF. Proteins were extracted at indicated times and Akt and Erk activation was performed by ALPHAscreen® analysis. (D) *H19*-overexpressing cells (H19) were transfected with miR-675 specific inhibitor (amiR) for 72 h. Then cells were treated with 10 ng/ml EGF. Proteins were extracted at indicated times and Akt and Erk activation was determined by ALPHAscreen® analysis. These experiments were performed three times in triplicate. Data represent one representative experiment. For ALPHAscreen® error bars represent SEM.

Figure 5. *H19*/miR675 increased migration and proliferation of MDA-MB-231 breast cancer cells. (A) Control (mock) or *H19*-overexpressing (H19) cells were cultured in transwells for 24 h. Migrated cells were then colored with violet crystal and counted. Results are presented as the percentage of control. (B) Wound healing assay performed on control (mock) and *H19*-overexpressing (H19) cells. (C) Wound healing assay performed on control (ctrl) and

transiently overexpressing miR-675 (mimic) cells. **(D)** QRT-PCR analysis of expression of miR-675-5p in control (mock) and miR-675-overexpressing cells (miR cl1, miR cl2). Results are presented as relative levels compared to MDA-MB-231 control cells (indexed to 1). **(E)** Western blot analysis of c-Cbl and Cbl-b levels in control and miR-675-overexpressing cells. **(F)** Transwell migratory assay performed on control (mock) and miR-675-overexpressing cells after 6 h of culture. **(G)** Cell proliferation determined by MTT test. Data represent mean of three independent experiments and error bar sem. * $p < 0.05$; ** $p < 0.005$; *** $p < 0.001$.

Figure 6. MiR-675 promoted tumor growth and metastasis. **(A)** MDA-MB-231 control (mock) and overexpressing miR-675 (miR cl1, miR cl2) cells were subcutaneously injected into SCID mice (7 mice per group). Tumor growth curve represents the mean of tumor volumes in each group and error bar sem. MiR cl1 growth curve statistical analysis is surrounded. **(B)** Detection of tumor cell proliferation (PCNA) and apoptosis (TUNEL) of paraffin-embedded sections of tumors 65 days after injections. Scale bar is 100 μ m. **(C)** Western blot analysis of c-Cbl and Cbl-b levels in xenografted tumors. Tumors from three mice for each group were analyzed (mock, miR cl1, miR cl2). **(D and E)** Detection of GFP-positive cells by flow cytometry analysis in brain, liver and lung of mice xenografted with control or miR-675-overexpressing cells (mock, miR cl1, miR cl2). Histograms in E represent the mean of results from six xenografted mice * $p < 0.05$; ** $p < 0.005$; *** $p < 0.001$.

SUPPLEMENTARY LEGENDS

Supplementary Figure 1. Growth factors induced Akt and Erk phosphorylation in MDA-MB-231 cell stably overexpressing H19. **(A, B, C)** Control (mock) and *H19* overexpressing cells (*H19*) were treated with 10 ng/ml HGF. Proteins were extracted at indicated times and Met, Akt and Erk activation was determined by ALPHAscreen® **(A, B)** and Western blot analysis **(C)**. **(D, E)** Cells were treated with 10 ng/ml pro-NGF **(D)** or 100 ng/ml NGF **(E)**. Proteins were extracted at indicated time and Akt and Erk activation were determined by Western blot analysis. These experiments were performed three times in triplicate. Data represent one representative experiment. For ALPHAscreen® error bars represent the SEM.

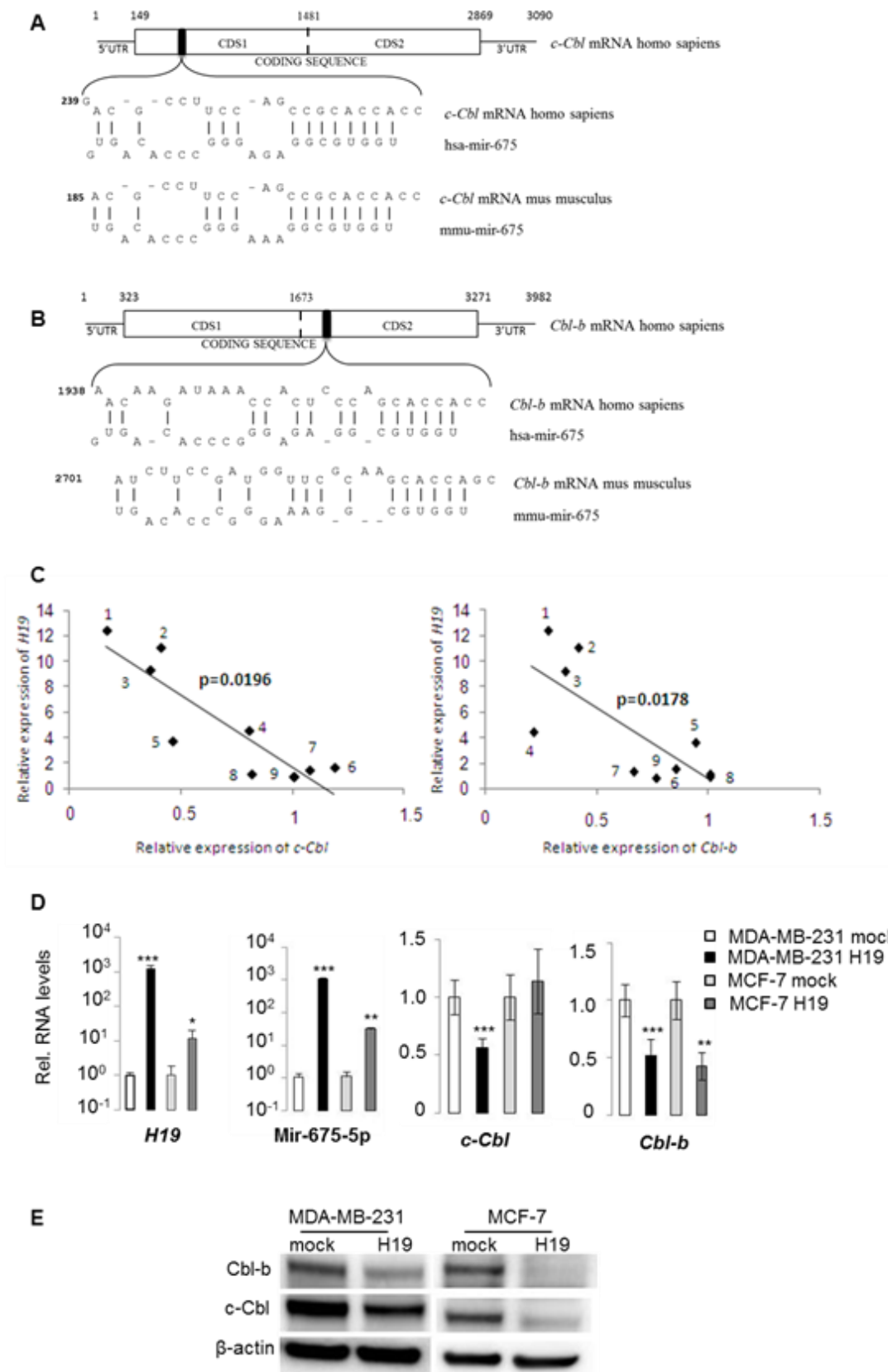
Supplementary Figure 2. HGF and EGF induced Akt and Erk phosphorylation in MDA-MB-231 cell transiently overexpressing H19. Cells transiently transfected with GFP plasmid or *H19* plasmid for 72 H were treated with 10 ng/ml EGF **(A, B)** or 10 ng/ml HGF **(C, D)**. Proteins were extracted at indicated times and activation of Akt and Erk was determined by ALPHAscreen®. These experiments were performed three times in triplicate. Data represent one representative experiment and error bars the SEM.

Supplementary Figure 3. H19 increased migration and proliferation of breast cancer cells. (A) Control (mock) or *H19*-overexpressing (H19) MDA-MB-231 cells were cultured in transwells in the presence or absence of 10 ng/ml HGF for 24 h. Migrated cells were then colored with violet crystal and counted. Results are presented as the percentage of non-treated control cells. (B) Control (mock) and *H19*-overexpressing (H19) MCF-7 cells were cultured in the presence of 10 ng/ml EGF. Cell proliferation was determined by MTT test at indicated days. ** $p < 0.005$; *** $p < 0.001$

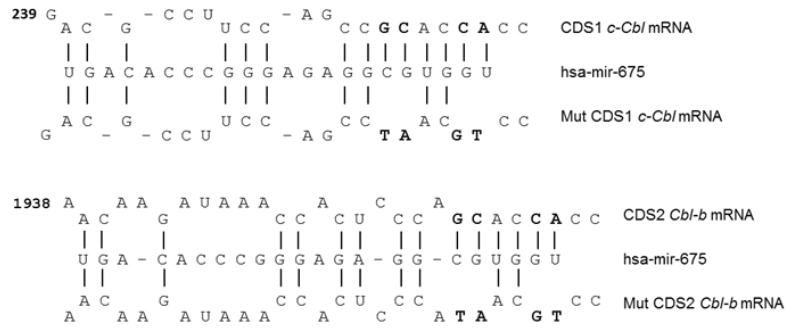
Supplementary Figure 4. Rb is not regulated by H19 in breast cancer cells. The firefly luciferase activity in breast (A) and colorectal (B) cancer cells after cotransfection with reporter construct and miR-675 mimic or miR-675 inhibitor (amiR). The luciferase activity was measured by dual-luciferase reporter assay (Promega) and was normalized to Renilla luciferase activity. Plasmids were transfected with mimic or antimir, or their controls in breast and colorectal cancer cells lines (MDA-MB-231, T47D and HT-29). Data represent mean of three independent experiments *versus* their respective controls in percentage and error bar, sem. ** $p < 0.01$

Supplementary Figure 5. MiR-675 expression is associated to poor free survival. microRNA and mRNA expression data were collected from a previously published cohort [36], NCBI accession number GSE22220. We used Graphpad Prism to analyze the data corresponding to miR-675 (probe ILMN_3167407), c-Cbl (probe 1230102) and Cbl-b (probe 4010563). Tumors were classified as 'low' when \leq to the median and 'high' otherwise. (A) Kaplan-Meier analyses were performed for each marker individually after ER stratification. (B) Kaplan-Meier analyses were performed by comparing the patients expressing both a high level of miR-675 and a low level of its putative target (c-Cbl or Cbl-b). Number of patients included is indicated next to each curve. P values were calculated using Grehan-Breslow-Wilcoxon test.

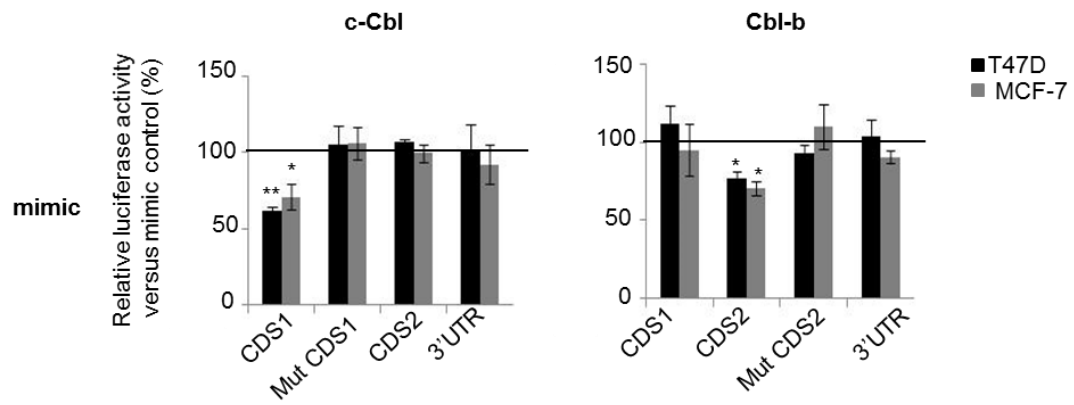
Supplementary Figure 6. 91H RNA is not a sponge of miR-675. (A) Relative expression of *91H* in miR-675 overexpressing cells (miR cl1, miR cl2) determined by qRT-PCR. Data represent mean of three independent experiments and error bar, sem. (B) Western blot analysis of c-Cbl and Cbl-b in breast cancer cells, MDA-MB-231, transfected with *91H* shRNA (sh) compared to control cells (mock).



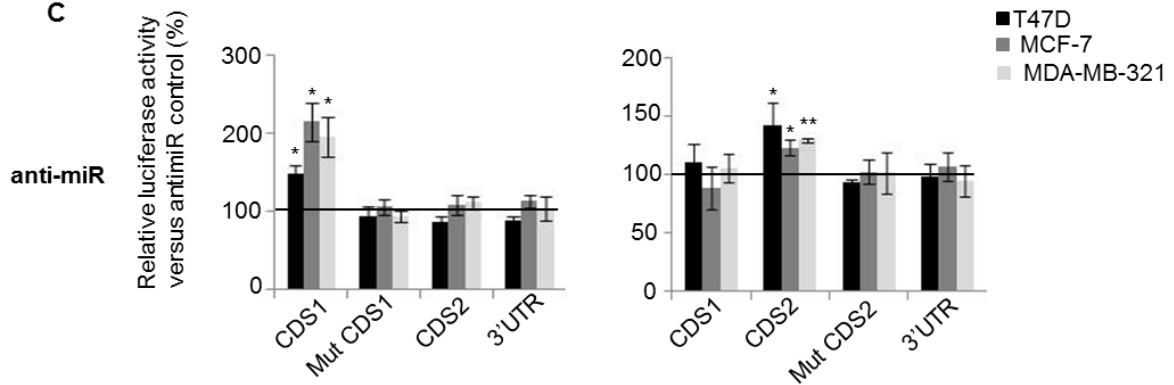
A

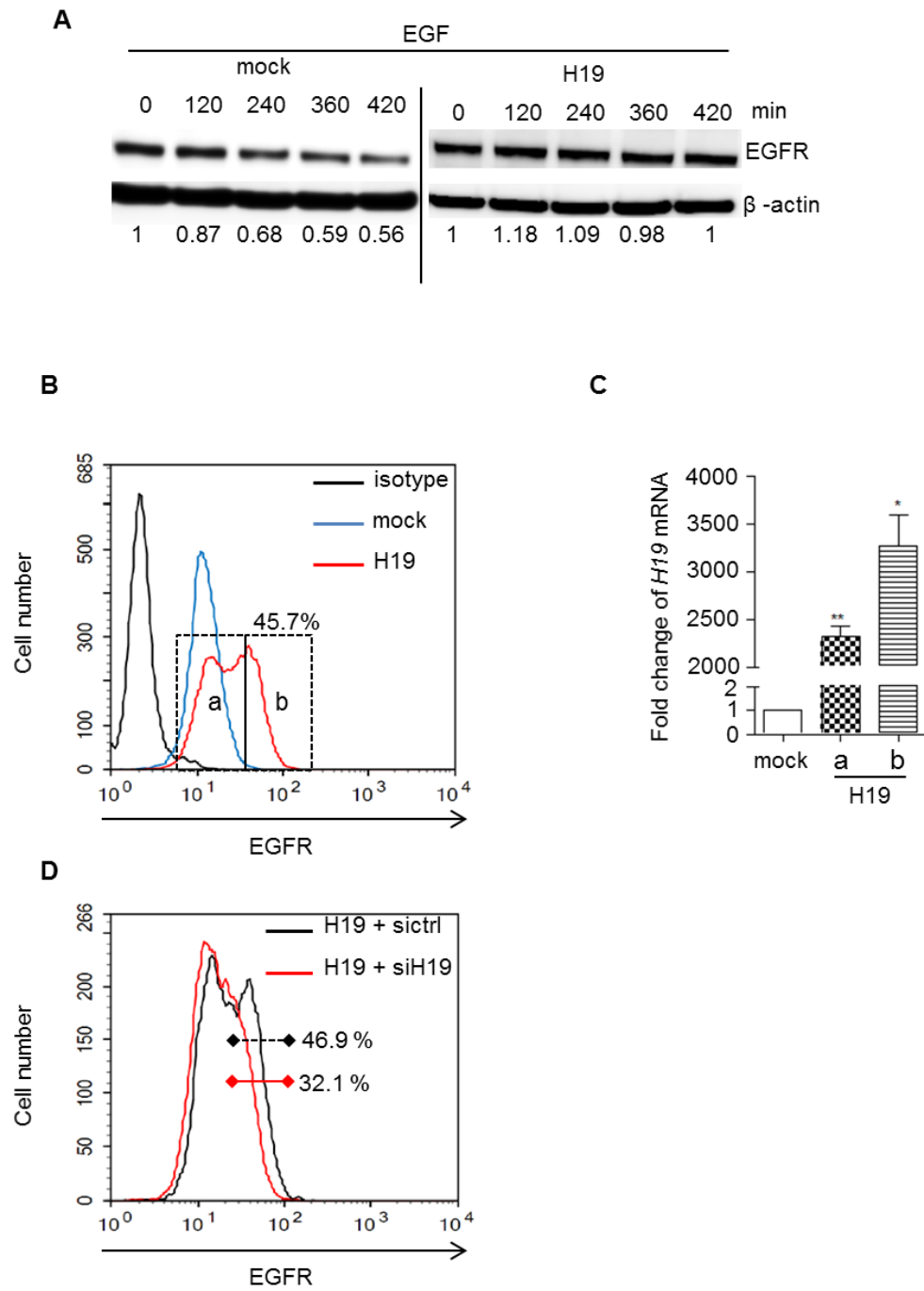


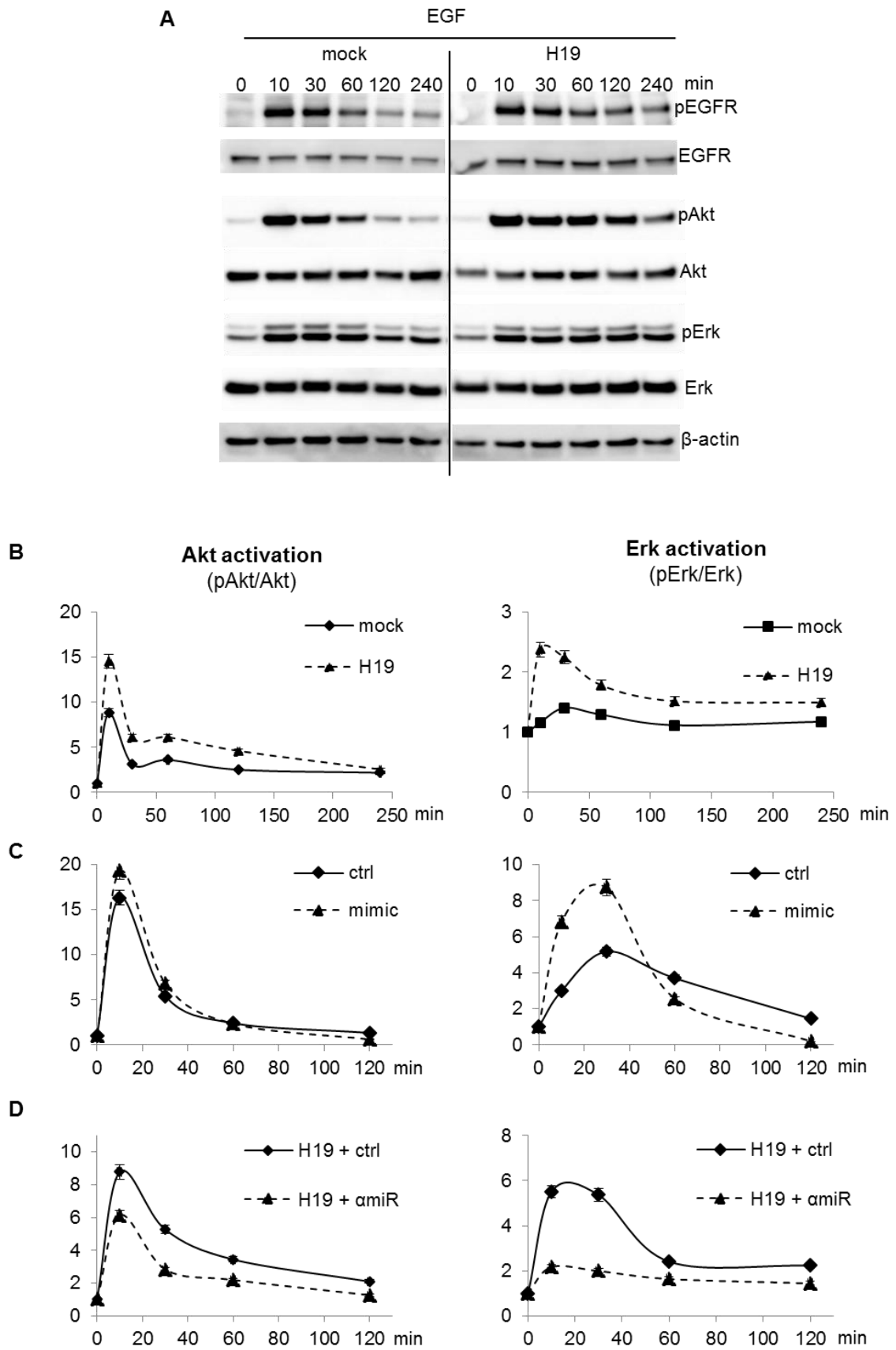
B

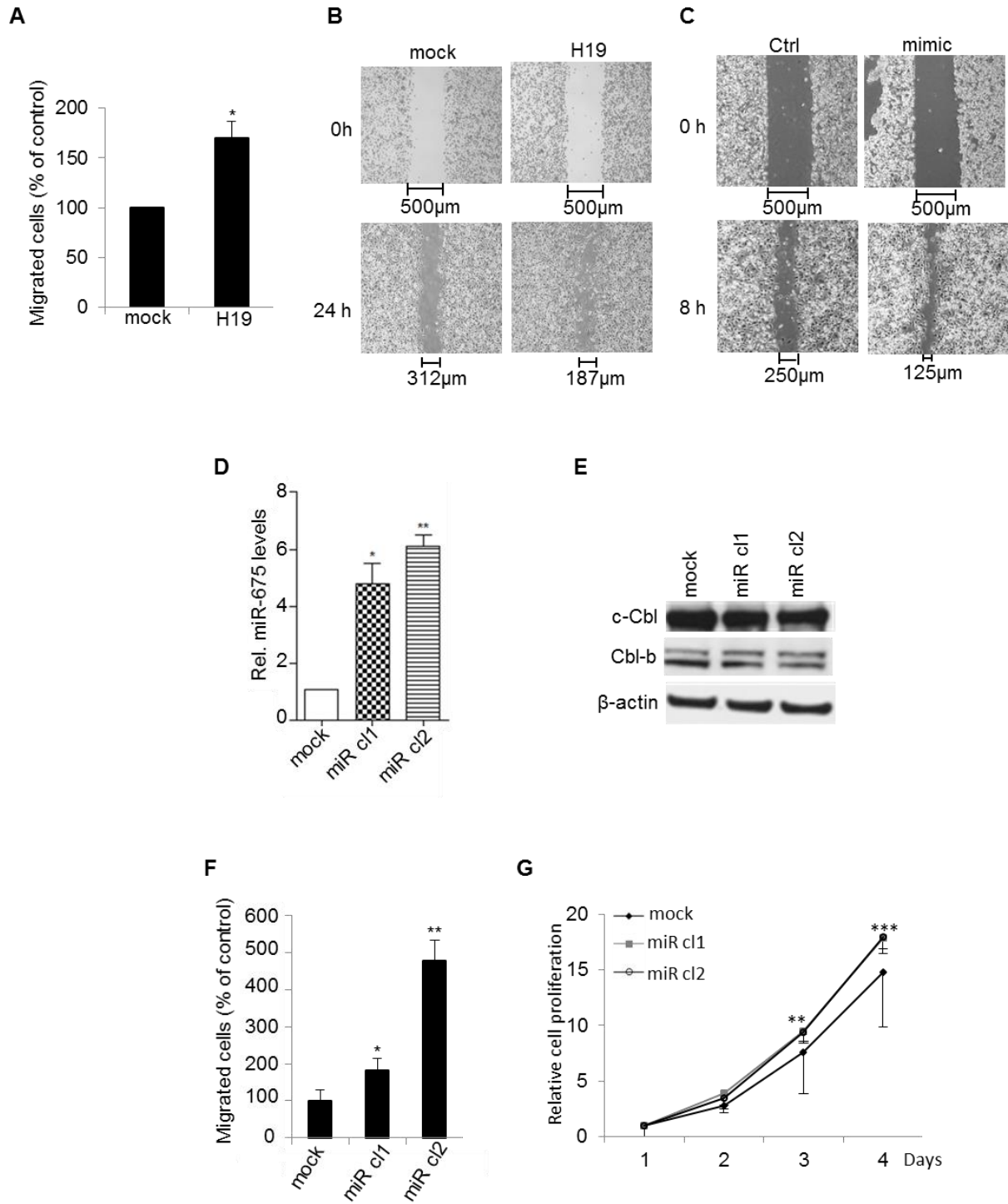


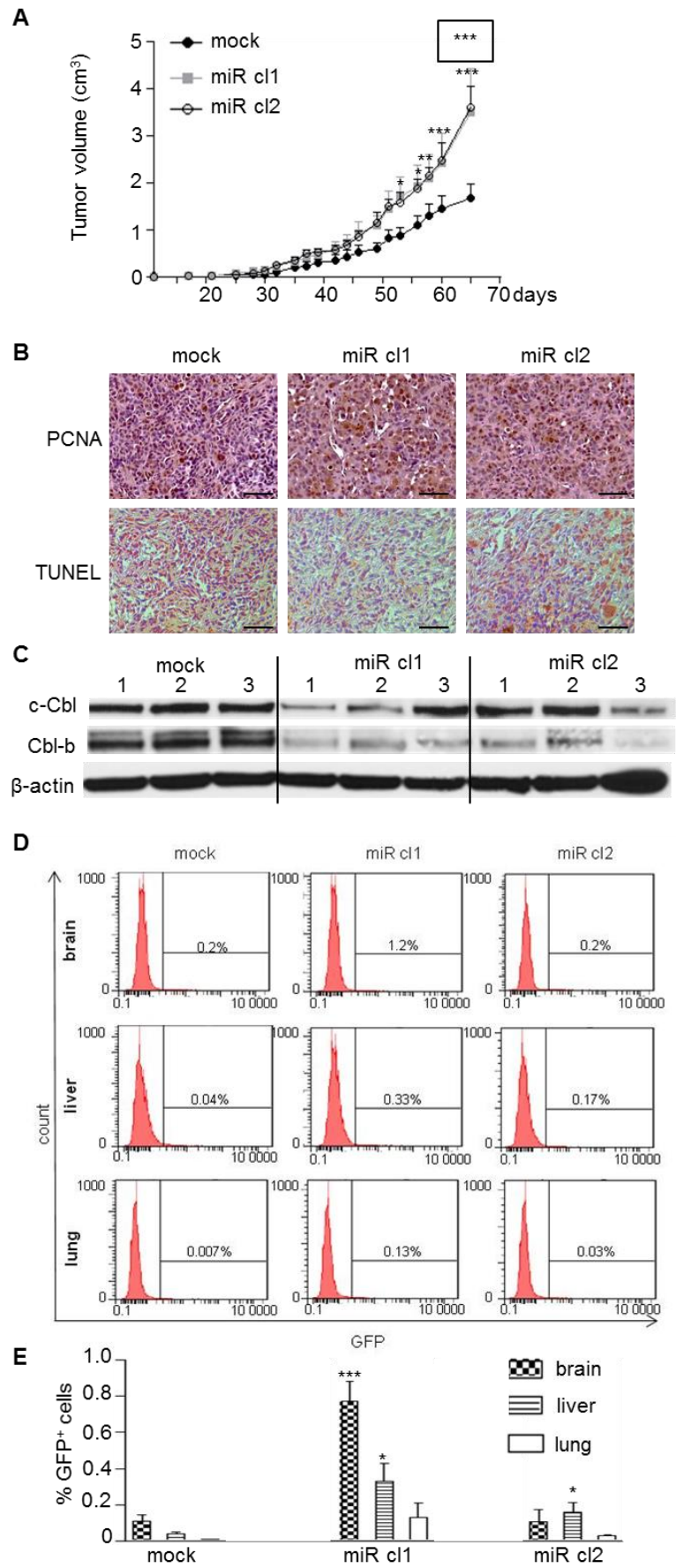
C

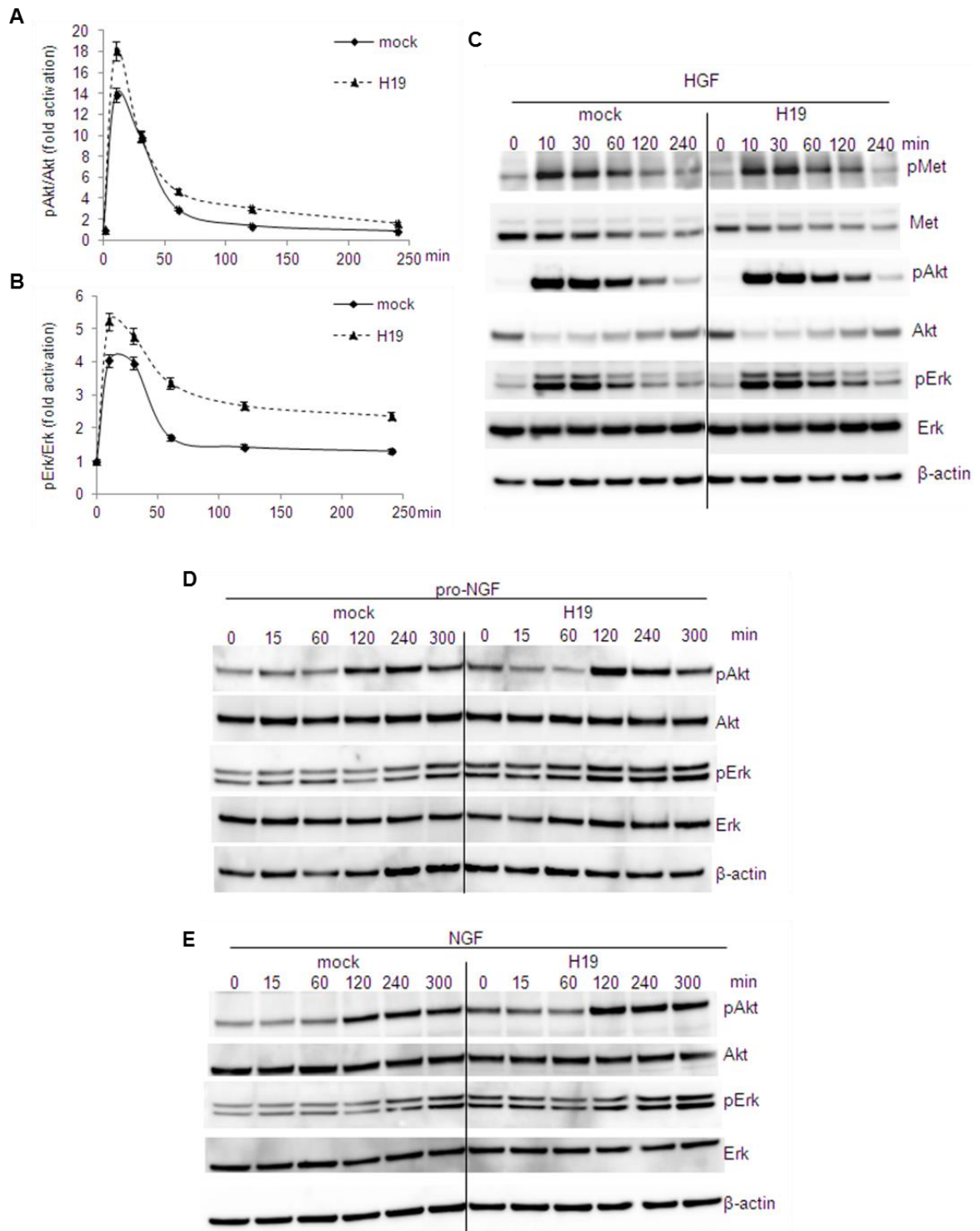


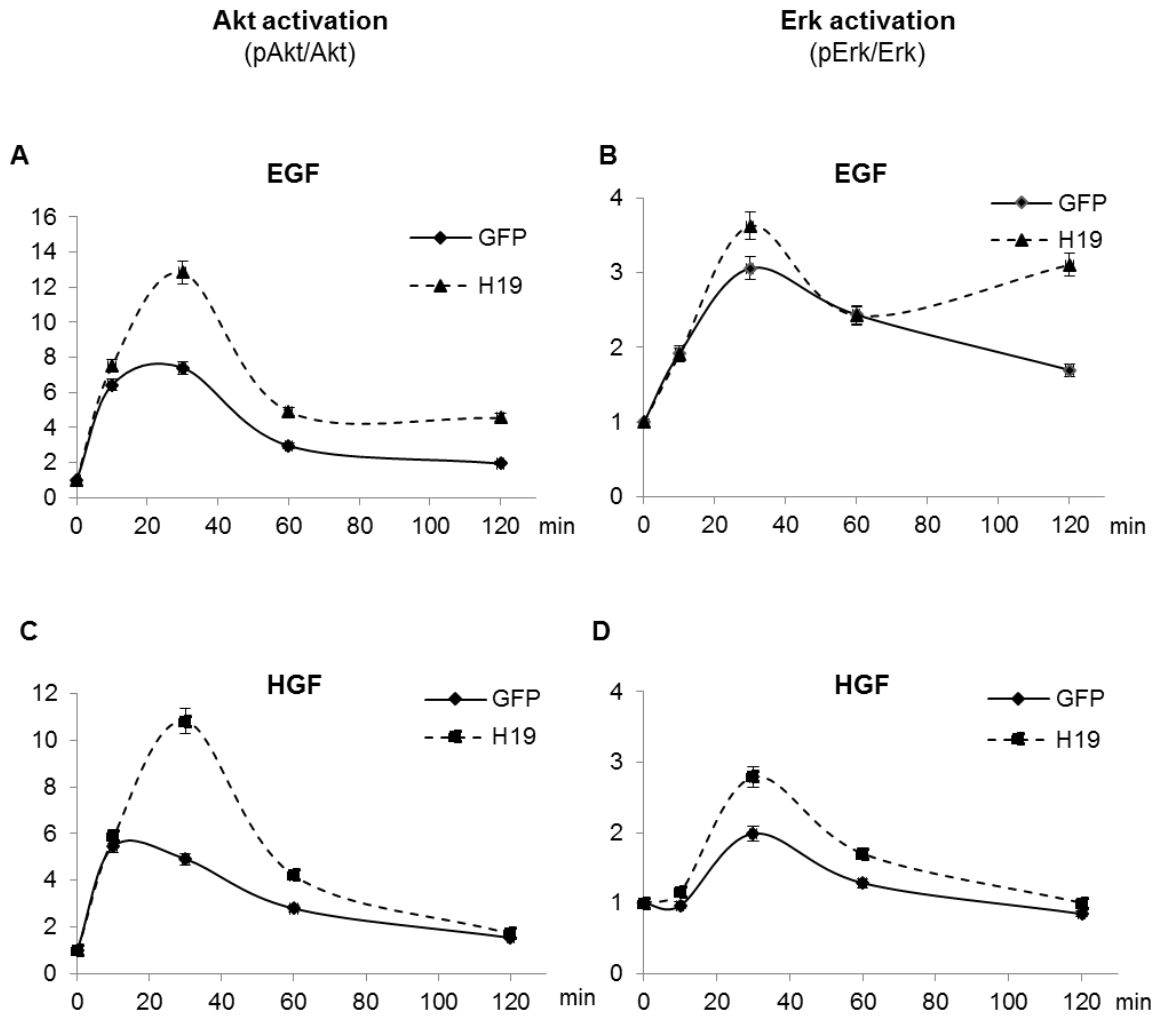


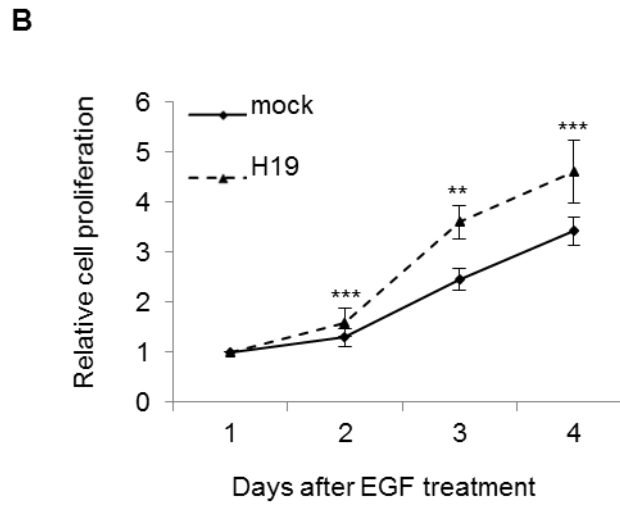
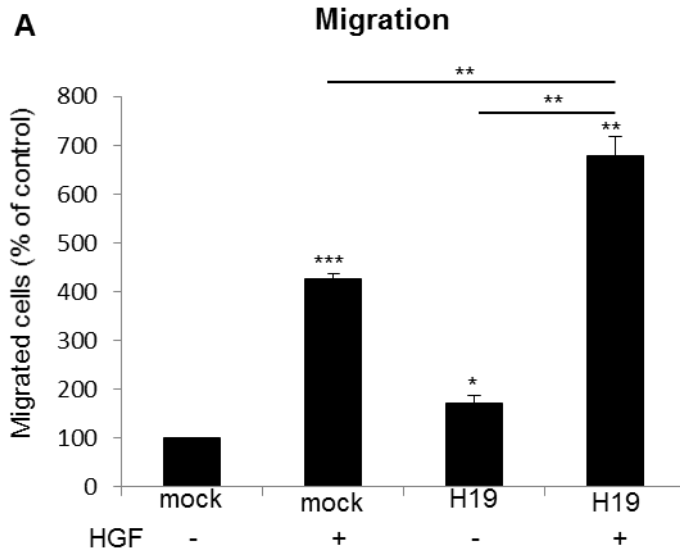


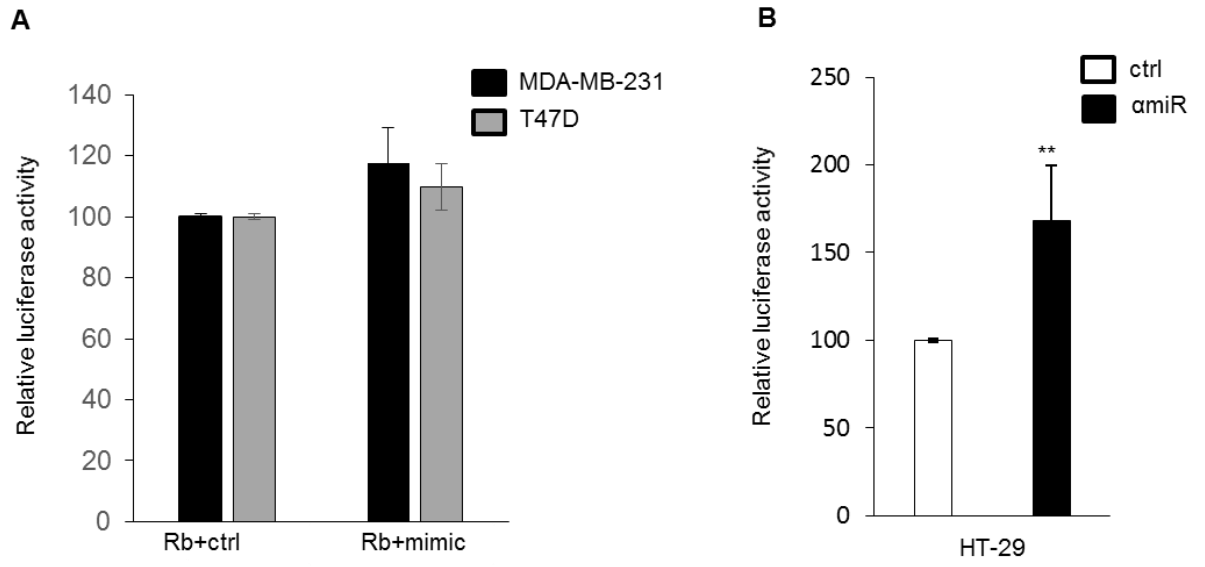






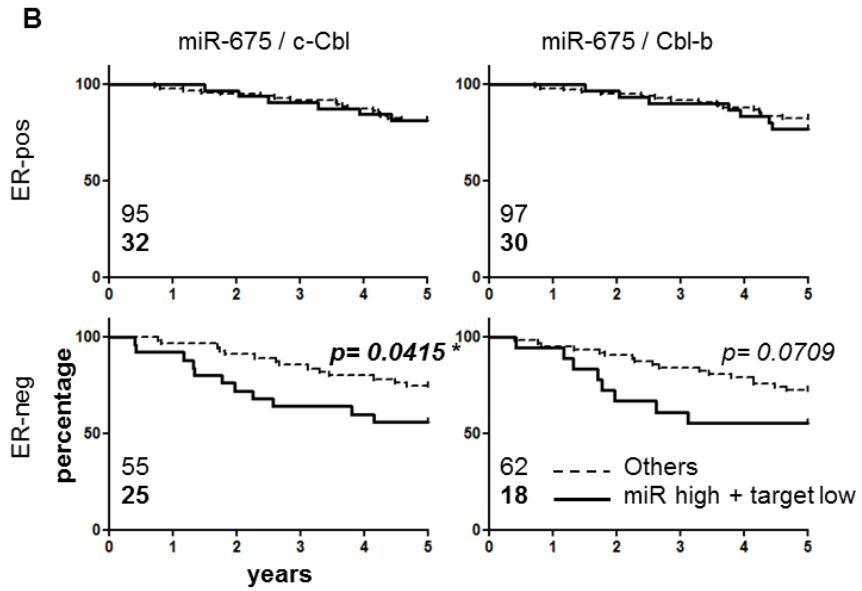
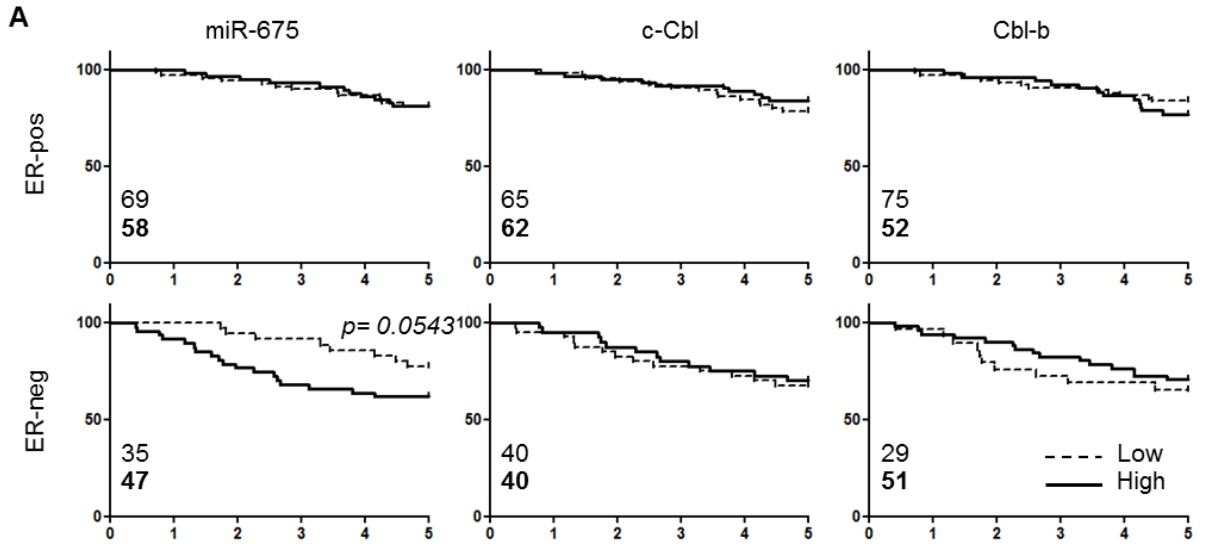


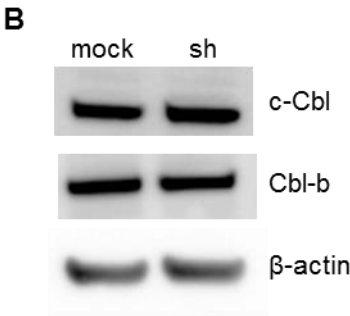
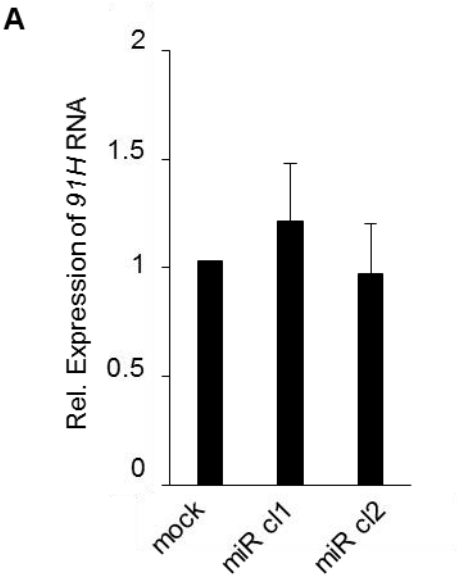




Vennin et al.

Supplementary Figure 4.





Antibody	Reference	Origin
p-Met	#3135	Cell signaling
Met	37-0100	Invitrogen
p-EGFR	#1138-1	Epitomics
EGFR	AHR5062	Invitrogen
p-Akt	4060S	Cell signaling
Akt	2920	Cell signaling
p-Erk	#9106	Cell signaling
Erk	C14/sc-154	Santa Cruz
c-Cbl	Sc-170	Santa Cruz
Cbl-b	Sc-8006	Santa Cruz
β actin	Sc-47778	Santa Cruz

Supplementary Table 1. Antibodies used for western blot analysis.

Primer	5' to 3'
	cloning
c-Cbl CDS1 SpeI	AACTAGTGGCCGGCAACGTGAAGAAGAGCT
c-Cbl CDS1 PmeI	CGTTTAAACGCTGCCACTCCCTCTAGGATC
c-Cbl CDS2 SpeI	GACTAGTGAGGGAGTGGCAGCCTGTTGAGG
c-Cbl CDS2 MluI	ACGCGTCTAGGTAGCTACTAGGGCAGG
c-Cbl 3'UT SpeI	TACTAGTCACACCATCTCCCTGCTGCAGGT
c-Cbl 3'UT HindIII	AAGCTTCACCTTCAAATGCACTCAAGA
Mut c-Cbl antisense	TGGACGTTAGGCTGGAAGGCGTCCTTCATG
Mut c-Cbl sense	CATGAAGGACGCCTTCCAGCCTAACGTCCA
Cbl-b CDS1 SpeI	CACTAGTGCAAACCTCAATGAATGGCAGAAA
Cbl-b CDS1 HindIII	TAAGCTTCGTCCAAGTCTAGCATCGGCATG
Cbl-b CDS2 SpeI	CACTAGTATGCTAGACTTGGACGACGATGA
Cbl-b CDS2 HindIII	CAAGCTTCTATAGATTTAGACGTGGGGATA
Cbl-b 3'UTR SpeI	TACTAGTCAGCCAGAACTGTAGACACCAAA
Cbl-b 3'UTR HindIII	TAAGCTTATGGAAAACCCCTTACAAAAGG
Mut Cbl-b antisense	AGGACGTTATGGGAGTGGTTTTATCTTGTTT
Mut Cbl-b sense	AAACAAGATAAACCACTCCCATAACGTCCT
H19 FL 5p	AGCAGGGTGAGGGAGGGGGT
H19 FL 3p	GTAACAGTGTTTATTGATG
miR-675 HindIII	AAAGCTTAACGAGGCACTGCGGCCAGGGT
miR-675 BamHI	AGGATCCTCTGCCAAGCCAGCCCCAGGGGC
	RT-qPCR
RPLP0 sense	GTGATGTGCAGCTGATCAAGACT
RPLP0 antisense	GATGACCAGCCCAAAGGAGA
H19 sense	TACAACCACTGCACTACCTG
H19 antisense	TGGCCATGAAGATGGAGTCG
c-Cbl sense	GTGATCCCTGGACAGGAAGA
c-Cbl antisense	GGTGGTCACACTCTGGACCT
Cbl-b sense	GAAGAAAGCCACAGCCTGAC
Cbl-b antisense	GGATTGGTGGAGGTCTTTCA
91H sense	GCTTGTAGTAGAGTGCGCC
91H antisense	CATCCAGTTGACCGAGCTTG
	siRNA
siRNA H19 1 sense	UAAGGUGUUCAGGAAGGCC
siRNA H19 1 antisense	GGCCUUCUGAACACCUUA
siRNA H19 2 sense	AGCUUCACCUUCCAGAGCC
siRNA H19 2 antisense	GGCUCUGGAAGGUGAAGCU

Supplementary Table 2. Oligonucleotides used for plasmid construction. Primer used for qRT-PCR. Sequence of *H19* siRNA.

Discussion

Dans cette étude, nous avons mis en exergue la fonction du miR-675 dans les cellules cancéreuses mammaires. Grâce à une analyse *in silico* et des tests fonctionnels, nous avons identifié deux nouveaux gènes cibles du miR-675 : les ARNm *c-Cbl* et *Cbl-b*. Les protéines Cbl sont des ubiquitines ligases impliquées dans la désensibilisation des récepteurs à activité tyrosine kinase. Après stimulation avec différents facteurs de croissance (EGF, HGF, pro-NGF et NGF), nous montrons que l'activation des récepteurs à activité tyrosine kinase ainsi que leurs voies de signalisation sous-jacentes sont augmentées lorsque *H19* ou son microARN sont surexprimés. En effet, l'internalisation des récepteurs à activité tyrosine kinase est diminuée en présence d'*H19*. Les facteurs de croissance sécrétés par la cellule tumorale et son microenvironnement sont nécessaires à la prolifération, la survie, la migration ou encore l'angiogenèse de la tumeur. Notre équipe a mis en évidence les fonctions pro-tumorales des neurotrophines (NGF, BDNF, NT4/5) dans le développement du cancer du sein. Par ailleurs, nous avons déjà mis en évidence que les facteurs de croissance, principalement l'HGF et l'EGF, favorisent l'expression d'*H19* dans les cellules cancéreuses mammaires (Adriaenssens et al., 2002). Dans cette étude, nous montrons donc l'existence d'une boucle positive de rétrocontrôle entre *H19* et les facteurs de croissance.

Par ailleurs, de nombreuses thérapies induisent la dégradation des récepteurs aux facteurs de croissance en recrutant les protéines Cbl. Par exemple, l'icotinib, qui est administré dans les traitements des cancers du poumon, favorise la dégradation de l'EGFR en recrutant les protéines Cbl sur ce récepteur (Mu et al., 2013). L'ensemble de nos données suggère que *H19*, par l'intermédiaire de son microARN, peut être responsable de certaines résistances thérapeutiques en favorisant les voies autocrine et paracrine.

Dans cette étude, nous montrons également, pour la première fois, que le miR-675 augmente les capacités tumorigènes et invasives des cellules cancéreuses mammaires *in vitro* et *in vivo*. Ainsi, le microARN, généré par *H19*, participe aux effets oncogènes déjà décrits pour cet ARN.

Pour la première fois, nous avons mis en évidence un mode d'action de l'ARN H19 dans le développement du cancer de la glande mammaire. Nous révélons un lien entre H19, le miR-675, les protéines CBL et les récepteurs à activité tyrosine kinase favorable au développement du cancer du sein et à son agressivité.

2. Identification d'une nouvelle cible (l'ARNm FADD) et d'un nouveau mode de régulation des microARN

Introduction

Lors de ma recherche *in silico* de cibles potentielles du miR-675, d'autres gènes potentiellement régulés par ce microARN m'ont interpellée. Un gène, impliqué dans l'apoptose cellulaire, a plus particulièrement attiré mon attention. Après avoir vérifié la modification d'expression de celui-ci dans les cellules surexprimant le gène *H19*, j'ai confirmé l'action du miR-675 sur l'UTR3' de ce gène. Nous avons mis en évidence une régulation particulière de ce gène par notre microARN. J'ai donc cherché à identifier le mécanisme responsable de cette nouvelle régulation.

Matériels et méthodes

Culture cellulaire

Les lignées cellulaires utilisées (MDA-MB-231, MCF-7 et T47D) sont entretenues en RPMI 1640 supplémenté de Zell Shield et de sérum de veau fœtal 10%, comme décrit dans le Matériels et Méthodes de l'article 1.

Prolifération, survie cellulaire

Pour les courbes de croissance, les cellules contrôles, celles surexprimant *H19* et celles surexprimant *H19* et transfectées stablement avec les TALEN, sontensemencées en plaque 6 puits à raison de 50 000 cellules/puits. Chaque jour, le nombre de cellules par puits est compté à la cellule de Malassez.

Pour les tests de viabilité, 15×10^4 cellules/puits sontensemencées en plaque 6 puits pendant 24 hrs. Les cellules sont traitées au Trail à 10 μ g/ml pendant 6 hrs. La viabilité cellulaire est déterminée par un test MTT (décrit dans l'article 1).

Etude de l'expression

Les ARNm sont extraits et quantifiés par qRT-PCT selon le protocole décrit dans le matériels et méthodes de l'article 1.

Les amorces de qRT-PCR sont regroupés dans la table 7.

Le protocole d'extraction et de quantification des microARN est décrit dans le matériels et méthodes de l'article 1.

Le protocole permettant l'étude de l'expression des gènes par western blot est décrit dans le matériels et méthodes de l'article 1. Les anticorps anti-FADD (sc-5559, Santa cruz) et anti-pFADD (sc-12439, Santa cruz) sont commercialisés par Santa Cruz.

Résultats

Stabilité des ARNm

Les cellules sont ensemencées en plaque 6 puits à raison de 15×10^4 cellules/puits. Après adhésion, elles sont traitées à l'Actinomycine D (Sigma-Aldrich) à $5 \mu\text{g/ml}$ pendant 4hrs. Les ARNm sont extraits aux différents temps indiqués et quantifiés par qRT-PCR.

Clonage

L'UTR3' de FADD ainsi que l'UTR muté dans le site de fixation du microARN sont clonés dans le pMIR-REPORT luciférase en *SpeI*. L'UTR3' est amplifié par PCR à partir de cDNA MDA-MB-231 avec les amorces suivantes : sense 5'-ATACTAGTGGGCCGCTGCTTT-3', antisense : 5'-CGTTCGAATTAGGAATAGAAAAT-3'. La mutation du site de fixation du microARN est générée par PCR à partir du pMIR-REPORT luciférase contenant l'UTR3' FADD et avec les amorces suivantes pMirXhoI, mut FADD5p, pMirNaeI, mut FADD 3p regroupés dans la table 7. Le clonage est réalisé comme décrit dans le Matériels et Méthodes de l'article 1.

La mutation du site de fixation de FXR1 sur l'UTR3' de FADD est obtenu par PCR avec les amorces FADD RH FXR reverse, FADD RH FXR forward, mut FADD FXR 5p, mut FADD FXR 3p regroupés dans la table 7. Ensuite le fragment de PCR porteur de la mutation est inséré dans le pMIR-REPORT luciférase par recombinaison homologue entre les sites de restriction *SpeI* et *HindIII*. La recombinaison homologue est réalisée avec le kit « GeneArt Seamless cloning and assembly kit » (A13288, Invitrogen). Le plasmide obtenu est amplifié comme décrit dans le Matériel et Méthodes de l'article 1.

Transfection et test d'activité luciférase

Les protocoles de transfection des miRNA et des antimirs ainsi que le dosage d'activité luciférase sont décrits dans le Matériels et Méthodes de l'article 1.

Le siRNA FXR1 est transfecté comme décrit dans le Matériel et Méthodes de l'article 1. Sa séquence est indiquée dans la table 7.

Immunoprécipitation des ARNm

L'immunoprécipitation des ARNm associés aux protéines AGO2 et FXR1 est réalisée avec le kit « Magna RIP RNA-binding protein immunoprecipitation kit » (17-700, Millipore). Les anticorps anti-AGO2 (03-110) et anti-FXR1 (03-176) optimisés pour le RIP sont commercialisés par Millipore.

Extraction ADN génomique

Les cellules sont lysées dans le tampon d'extraction constitué de Tris-HCl 10mM pH 7,4 ; NaCl 50mM ; EDTA 10mM ; SDS 0,2%. Après digestion du lysat à la protéinase K (Roche) une nuit à 55°C , l'ADN génomique est précipité par l'ajout de NaCl 1,6M et éthanol 100% froid. Il est ensuite resuspendu dans 1 ml de Tris-HCl 10 mM pH 8 ; EDTA 0,1mM.

Test statistique

Les données sont analysées avec le t-Test de Student. Une valeur $p < 0,05$ est considérée comme significative.

Résultats

Identification d'une nouvelle cible

Lors de ma recherche *in silico* de cibles potentielles du miR-675, il est apparu que ce microARN pouvait se fixer sur l'UTR3' de l'ARNm *FADD* (*Fas-associated Death Domain*) (fig 25A). Dans un premier temps, j'ai étudié l'expression de ce gène dans nos cellules cancéreuses mammaires surexprimant stablement le gène *H19* (fig 25B, C). De manière surprenante, nous avons observé que l'expression de *FADD* est augmentée lorsque *H19* est surexprimé. Cet effet est retrouvé au niveau ARNm mais aussi au niveau protéique.

J'ai ensuite vérifié si la forte expression du gène *FADD* est associée à une augmentation de la transcription du gène ou à une stabilisation de l'ARNm dans la cellule. Pour cela, j'ai traité les cellules contrôles (mock) et les cellules surexprimant le gène *H19* (H19) avec un inhibiteur de la transcription pendant 4 heures (Actinomycine D) et déterminé la quantité d'ARNm *FADD* par qRT-PCR dans les cellules au cours du traitement (fig 26A). Dans les cellules surexprimant stablement le gène *H19*, nous observons un retard dans la dégradation de l'ARNm *FADD*. En effet, dans ces cellules, la demi-vie de l'ARNm *FADD* est supérieure à 1 heure alors qu'elle est inférieure à 10 minutes dans les cellules contrôles. Dans ces deux lignées cellulaires (mock et H19), la stabilité de l'ARNm *c-myc* (un ARN contrôle) est inchangée (fig 26B). Ainsi, la hausse d'expression de *FADD* dans les cellules *H19*⁺ peut être expliquée par une meilleure stabilité de l'ARNm.

Validation de la cible

Cette régulation étant inattendue, j'ai vérifié si cette stabilisation est due à l'action du miR-675 sur le site prédit. Pour cela, j'ai cloné l'UTR3' de *FADD* en aval de la luciférase dans le pMIR-REPORT luciférase et étudié l'activité luciférase du plasmide obtenu en présence d'*H19*, du miR-675 synthétique (mimic) ou de son inhibiteur spécifique (α miR). J'ai également cloné l'UTR3' de *FADD* possédant une mutation sur le site de fixation du microARN (fig 27A).

Une augmentation de l'activité luciférase de 50% du plasmide 3'UTR est observée en présence du gène *H19* (fig 27B). Cet effet positif est abrogé lorsque l'UTR3' est muté sur le site de fixation du microARN. De façon similaire, une hausse de l'activité luciférase de 50% du plasmide 3'UTR est observée en présence du miR-675 synthétique (mimic) (fig 27C). Cet effet est également aboli lorsque l'UTR3' est muté. Ces tests ont également été réalisés dans les cellules hormono-dépendantes T47D. J'ai obtenu des résultats similaires à ceux obtenus dans les MDA-MB-231 (des cellules hormono-indépendantes). En effet, une augmentation d'activité

Résultats

Luciférase de 50% est obtenue en présence d'*H19* et de son microARN (fig 27B, C). J'ai également transfecté un inhibiteur spécifique du miR-675 dans ces deux lignées cellulaires et étudié l'activité luciférase des différents plasmides. Nous observons une diminution de l'activité luciférase du plasmide 3'UTR dans ces deux lignées en présence de l'inhibiteur (fig 27D). Une mutation sur le site de fixation du microARN abolit cette action. En conclusion, nous montrons que le miR-675 se fixe sur l'UTR3' de *FADD* via la séquence prédite *in silico*. Nous montrons également que le miR-675 régule positivement l'expression de cet ARNm puisque une augmentation d'activité luciférase est observée en sa présence. Nous avons donc identifié une nouvelle régulation, inattendue, de l'action des microARN sur leurs ARNm cibles.

Identification d'un nouveau mécanisme de régulation des microARN

En 2007, Vasudevan et collaborateurs ont montré que le miR-369 exerce un effet positif sur l'ARNm du *TNF α* lorsque les cellules sont privées en sérum (Vasudevan et al., 2007). Ils montrent également l'existence de complexes RISC différents en fonction du statut prolifératif des cellules. En effet, lorsque les cellules sont arrêtées dans le cycle, ils ont montré que la protéine FXR1 est associée au complexe RISC. Cependant, dans notre cas, les cellules étaient en phase de prolifération. Néanmoins, nous avons testé l'hypothèse d'une modulation de l'action des microARN par la protéine FXR1. *In silico*, j'ai identifié un site potentiel de fixation pour cette protéine situé en aval et à proximité du site de fixation pour le miR-675. J'ai alors réalisé une immunoprécipitation des ARNm associés aux protéines AGO2 et FXR1. Nous observons que l'ARNm *FADD* est bien associé à la protéine AGO2 (fig 28). Ce résultat, qui indique que le complexe RISC est associé à l'ARNm *FADD*, est attendu puisque nous montrons que le miR-675 régule l'expression de l'ARNm. Nous montrons également une association de l'ARNm *FADD* avec la protéine FXR1 (fig 28). Ainsi, l'effet positif exercé par le miR-675 sur cet ARNm peut être expliqué par la présence de FXR1 sur le site prédit.

Afin de confirmer cette hypothèse, nous avons utilisé la technologie des TALENs (*Transcription Activator Like-Effector Nuclease*) pour muter le site de fixation de FXR1 sur l'ARNm *FADD*. Les TALENs sont des dérivés de facteurs de transcription, synthétisés pour permettre la reconnaissance de séquence spécifique et couplés à l'endonucléase FokI. Ils permettent alors la formation de coupures doubles brins sur la séquence ciblée (fig 29A). Ces coupures seront ensuite réparées par recombinaison non homologue, ce qui permet l'apparition de mutation type insertion ou délétion de nucléotides ou par recombinaison homologue si l'on transfecte un oligonucléotide. Après transfection des plasmides permettant l'expression des TALENs dans des cellules H19⁺ et sélection des cellules transfectées, nous avons séquencé l'ADN génomique des clones obtenus (fig 29B). En accord avec nos précédents résultats, nous avons choisi trois clones possédant une mutation sur le site de

fixation FXR1, correspondant à une délétion de 9bp, que nous avons nommé TAL1, TAL2 et TAL3 (fig 29B).

Par western blot, j'ai ensuite étudié l'effet de la mutation FXR1 sur l'expression du gène *FADD* (fig 29C). Nous observons une diminution de l'expression de FADD lorsque les cellules expriment les TALENs. Ainsi, la fixation de la protéine FXR1 sur l'ARNm *FADD* à proximité du site de fixation du microARN permet la régulation positive par le miR-675. Afin de conforter cette nouvelle donnée, j'ai également invalidé l'expression de la protéine FXR1 par siRNA et étudié l'expression de la protéine FADD. Nous observons qu'une diminution transitoire de l'expression de FXR1 dans des cellules surexprimant *H19* permet de diminuer l'expression de la protéine FADD (fig 29D). De plus, j'ai également étudié l'activité luciférase du plasmide pMIR-REPORT cloné avec l'UTR3' *FADD* muté sur le site de fixation FXR1 dans nos cellules cancéreuses mammaires en présence du miR-675 et de son inhibiteur spécifique. Nous observons que le miR-675 ou son inhibiteur spécifique n'exerce aucun effet sur l'activité luciférase du plasmide lorsque l'UTR3' est muté sur le site de fixation FXR1 (Fig 29E). Ainsi, l'ensemble des résultats obtenus nous montre que la présence de FXR1 sur l'ARNm *FADD* permet le recrutement du miR-675 associé au complexe RISC et permet également une régulation positive de l'expression de cet ARNm.

Par ailleurs, le gène *H19* possède un long ARN non codant antisens, *91H* (Berteaux et al., 2008). Cet ARN pouvant être une éponge à microARN, nous avons étudié l'expression de la protéine FADD dans des cellules invalidées pour l'ARN *91H*. Aucune variation de l'expression de FADD n'a été détectée dans ces cellules (Fig 29F). Ainsi, l'ARN non codant *91H*, n'est pas une éponge au miR-675 dans les cellules cancéreuses mammaires.

Phénotypes associés à la surexpression de FADD

FADD étant une protéine impliquée dans l'apoptose, nous avons ensuite étudié la résistance des cellules surexprimant *H19* à différents apoptogènes, notamment le Trail. Après traitement, le taux de viabilité cellulaire est déterminé par un test MTT. De manière surprenante, après traitement, les cellules contrôles ont une viabilité cellulaire de 60% tandis que les cellules surexprimant stablement *H19* résistent à l'apoptose à 90% (fig 30A). Des résultats similaires ont été obtenus lorsque les cellules sont traitées avec d'autres drogues (Anisomycin, staurosporine) et dans plusieurs lignées cellulaires (MDA-MB-231 et MCF-7). Par ailleurs, une augmentation du taux de phosphorylation de FADD est observée lorsque *H19* est surexprimé (fig 30B). Cette modification post-traductionnelle empêche la formation du complexe d'apoptose (DISC) et favorise la progression du cycle cellulaire (Zhang et al., 2004b). J'ai alors étudié la croissance des cellules contrôles, surexprimant *H19* et des clones TALENs. Une hausse de la croissance cellulaire est observée lorsque *H19* est surexprimé (*H19*) par rapport aux cellules contrôles (mock) (fig 30C). A l'inverse, les cellules TALENs (surexprimant *H19* et

Résultats

mutées pour FXR1) ont une croissance cellulaire significativement ralentie par rapport aux cellules contrôles et aux cellules *H19*⁺. Ainsi, une surexpression d'*H19* permet une forte expression de FADD et p-FADD dans les cellules cancéreuses mammaires. Cette régulation est associée à une résistance des cellules aux apoptogènes mais aussi à une meilleure croissance cellulaire.

Afin de confirmer ces résultats, les tests de viabilité cellulaire en présence d'apoptogènes mais aussi l'étude de la phosphorylation de FADD doivent être réalisés dans les cellules TALENs (données à venir).

Discussion

Dans cette étude, nous avons identifié une troisième cible du microARN généré par *H19* dans les cellules cancéreuses mammaires : il s'agit de l'ARNm *FADD*. Une meilleure stabilité de l'ARNm *FADD* est obtenue lorsque les cellules surexpriment *H19*. Par ailleurs, le miR-675 régule positivement l'expression de cet ARNm en se fixant sur l'UTR3' *via* le site prédit *in silico*. De plus, nous montrons que l'association de FXR1 à proximité du site de fixation du microARN permet la fixation du miR-675 et permet la régulation positive par ce miR. Ainsi, nous avons mis en lumière un nouveau mode de régulation de l'action des microARN dans la cellule. Il serait intéressant de déterminer si ce nouveau mode d'action s'applique à l'ensemble des microARN dans tous les types cellulaires. En effet, au regard de l'implication des microARN dans de nombreux processus biologiques ou pathologiques, la compréhension de tous les mécanismes d'action s'avère important.

FXR1 est un membre de la famille FXR (*Fragile-X related protein*), composée de 3 protéines : FMRP, FXR1 et FXR2. Elles sont principalement impliquées dans le développement neuronal, musculaire ou encore cardiaque. Un défaut de leur expression est à l'origine du syndrome de l'X fragile. Ces protéines possèdent des domaines de fixation aux ARNm, nommés RGG box. Lorsqu'elles sont associées aux ARNm, elles peuvent moduler leur stabilité, leur traductibilité ou encore leur localisation cellulaire. La protéine FXR1 se fixe sur les UTR3' des ARNm dans des régions riches en AU (ARE), notamment l'ARNm du TNF α lors de la réponse immune et favorise sa stabilité ainsi que sa traductibilité (Khera et al., 2010). En effet, l'association de la protéine FXR1 aux polyribosomes a été montrée. Ici, une nouvelle fonction pour cette protéine a été identifiée. En effet, nous montrons que l'effet positif du miR-675 sur l'ARNm *FADD* est FXR1-dépendant. Nous montrons également que la fixation de FXR1 sur l'ARNm n'empêche pas la fixation du miR-675 puisqu'une mutation de la *seed sequence* réprime l'action de ce microARN (fig 27). Nous montrons également que la présence de FXR1 sur l'ARNm *FADD* est nécessaire au recrutement du complexe miR-675/RISC. Des résultats similaires ont déjà été obtenus dans des cellules neuronales. En effet, Edbauer et collaborateurs ont montré que la

protéine FMRP est essentielle au recrutement et à la fonction des miR-125b et miR-132 sur l'ARNm NR2A dans les cellules neuronales de souris (Edbauer et al., 2010). Cependant, contrairement à nos résultats, dans cette étude, les miR-125b et miR-132 régulent négativement l'expression de l'ARNm ciblé.

Par ailleurs, il est admis qu'un appariement des bases 10 et 11 du microARN permet le clivage de l'ARNm par la protéine AGO2 présente dans le complexe RISC. Dans cette étude, un appariement de ces bases est observé entre le miR-675 et l'UTR3' *FADD*. Mais, nous montrons que le complexe RISC ne clive pas cet ARNm. Nous pouvons supposer que la présence de la protéine FXR1 à proximité de la *seed sequence*, permet le recrutement d'un complexe RISC spécifique contenant des protéines qui répriment la fonction d'AGO2. Ainsi, FXR1 posséderait un nouveau mode d'action encore inconnu de nos jours.

A ce jour, le mode d'action des microARN est mal compris et mal défini. En effet, il était admis que ces derniers se fixaient sur les UTR3' des ARNm pour y exercer un effet négatif. Cependant, nous montrons qu'un même microARN, dans les mêmes lignées cellulaires, peut se fixer sur l'UTR3' d'un ARNm mais aussi sur les séquences codantes d'autres ARNm. De plus, dans un cas, il réprime l'expression de ses ARNm cibles et dans l'autre cas, en stabilisant l'ARNm, il favorise son expression. Ainsi, comprendre la régulation de l'action des microARN peut être primordial pour identifier de nouvelles cibles thérapeutiques plus efficaces et plus spécifiques des cellules cancéreuses.

FADD est une protéine pro-apoptotique de 27 kDa constituée de deux domaines : un domaine de mort (Death Domain, DD) et un domaine effecteur de mort (Death Effector Domain, DED). Lorsque les récepteurs de mort cellulaire sont activés (TNFR, Fas receptor par exemple), la protéine FADD se fixe sur ces derniers grâce à son domaine DD. Elle sert alors de molécule chaperonne puisqu'elle permet ensuite le recrutement et l'activation des caspases 8 et 10 par l'intermédiaire de son domaine DED. Ces caspases actives formeront alors le complexe DISC (*Death-inducing Signaling Complex*) qui permettra l'apoptose de la cellule. FADD est donc impliquée dans la transduction du signal apoptotique de la voie extrinsèque. Cependant, certaines modifications post-traductionnelles peuvent modifier la fonction de cette protéine. En effet, FADD peut être phosphorylée sur les sérines 194 et 200 (Zhang et al., 2004b ; Vilmont et al., 2015). Ces modifications post-traductionnelles sont catalysées par un grand nombre de kinases, notamment les kinases CK1, CK2 (Bowman et al., 2015 ; Vilmont et al., 2015) et possèdent plusieurs fonctions. En effet, ces résidus sérines sont localisés non loin du domaine DED de la protéine, donc leur phosphorylation empêche le recrutement et l'activation des caspases 8 et 10 et permet la localisation nucléaire de la protéine (Curtin et al., 2004 ; Vilmont et al., 2015). Ainsi, ces modifications post-traductionnelles répriment la fonction pro-apoptotique de FADD dans la cellule. De plus, d'autres études ont montré que la

phosphorylation de FADD sur la sérine 194 favorise la progression du cycle cellulaire, notamment la transition G2/M. En effet, p-FADD est capable de se fixer sur certaines kinases impliqués dans ce point de régulation du cycle cellulaire, PLK1 (Polo-like kinase 1) ou AURKA (Aurora kinase A) (Bowman et al., 2015). La phosphorylation, et par conséquent la fonction, de FADD est également régulée par des phosphatases. En effet, Kim et collaborateurs ont mis en évidence que l'adénylate kinase 2 (AK2) associée à la phosphatase DUSP26 permet la déphosphorylation de FADD sur sa sérine 194 chez l'Homme, et 191 chez la souris (Kim et al., 2014c). Cette modification est associée à une diminution de la croissance tumorale. Nous montrons qu'une forte expression de la protéine FADD dans les cellules surexprimant *H19* est associée à une résistance des cellules aux apoptogènes ainsi qu'à une augmentation de la prolifération cellulaire. Ces effets phénotypiques sont associés à la phosphorylation de FADD qui est plus élevée lorsque *H19* est surexprimé.

Nous avons donc mis en évidence un nouveau mécanisme de recrutement et de régulation de l'action des microARN sur leurs ARNm cibles. Nous montrons que les protéines associées aux ARNm, notamment FXR1, sont nécessaires au recrutement de certains microARN sur leur cible et régulent la fonction du complexe RISC sur cette cible. De plus, grâce à ce nouveau mécanisme, l'ARN H19 favorise la progression tumorale ainsi que la résistance aux drogues apoptotiques.

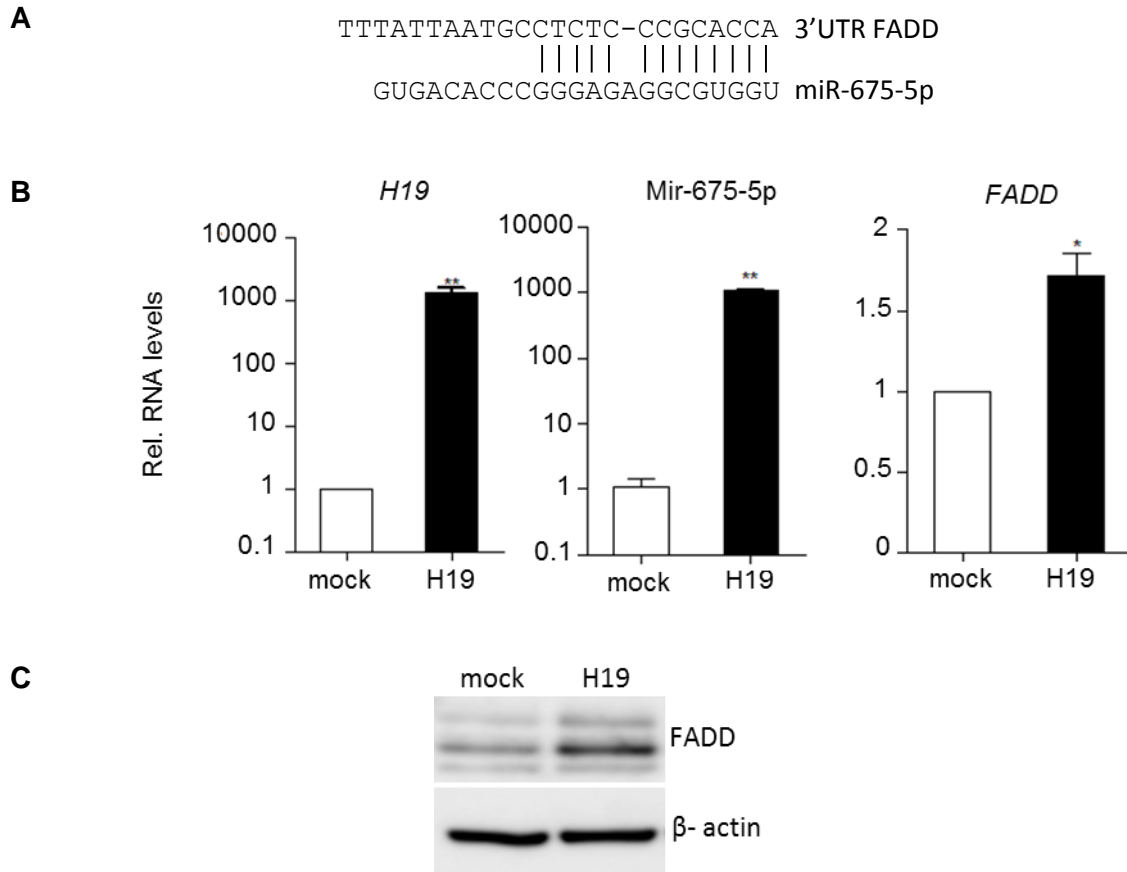


Figure 25. H19/miR675 favorise l'expression de FADD dans les cellules cancéreuses mammaires. (A) Alignement prédictif du miR-675 sur l'ARNm FADD (microRNA). (B) Expression relative d'H19, du miR-675 et de FADD dans les MDA-MB-231 surexprimant H19 (H19) par rapport aux cellules contrôles (mock) déterminée par qRT-PCR. (C) Détection de la protéine FADD par western blot dans les cellules cancéreuses mammaires MDA-MB-231. La β-actin est utilisée comme témoin de charge. *p<0,05 **p<0,005

Résultats

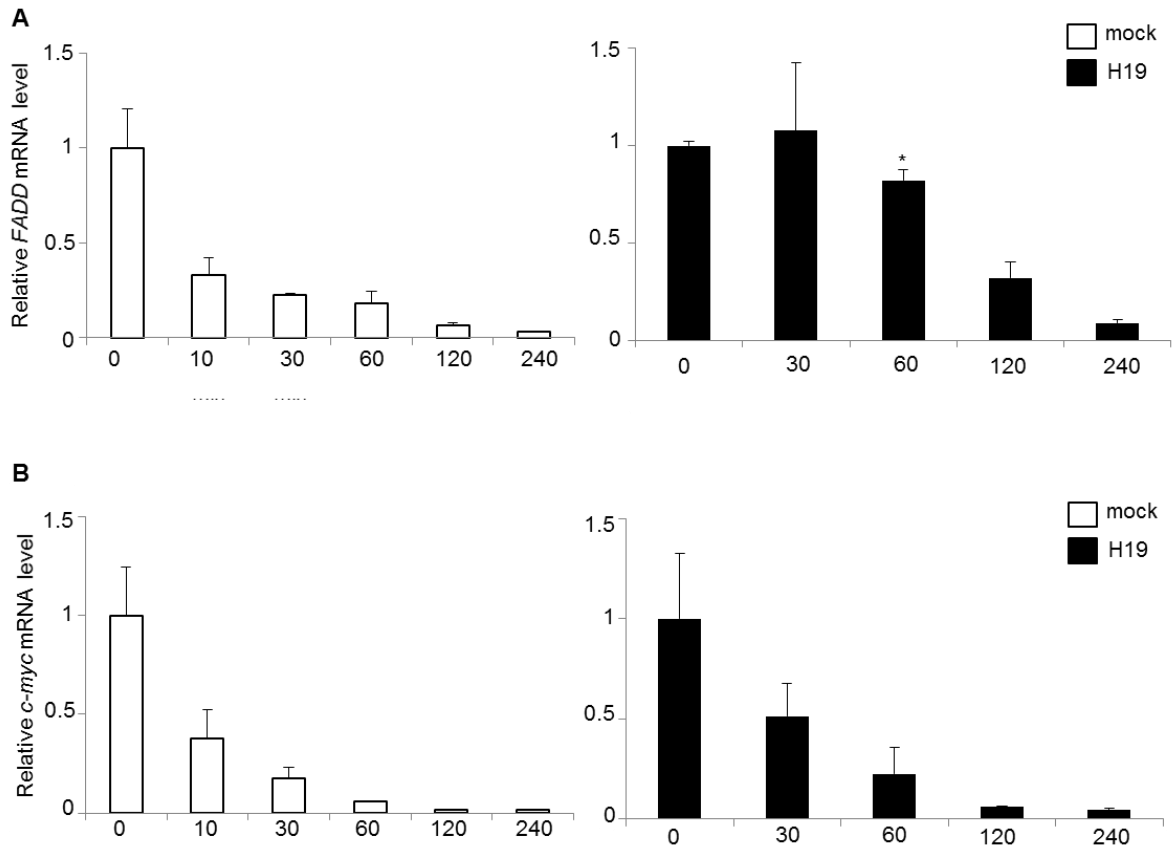


Figure 26. Stabilité des ARNm *FADD* (A) et *c-Myc* (B) dans les MDA-MB-231. Les cellules parentales (mock) et surexprimant *H19* (H19) sont traitées avec 5µg/ml d'Actinomycine D. La quantification des ARNm est réalisée par RT-qPCR aux temps indiqués. *p<0,05

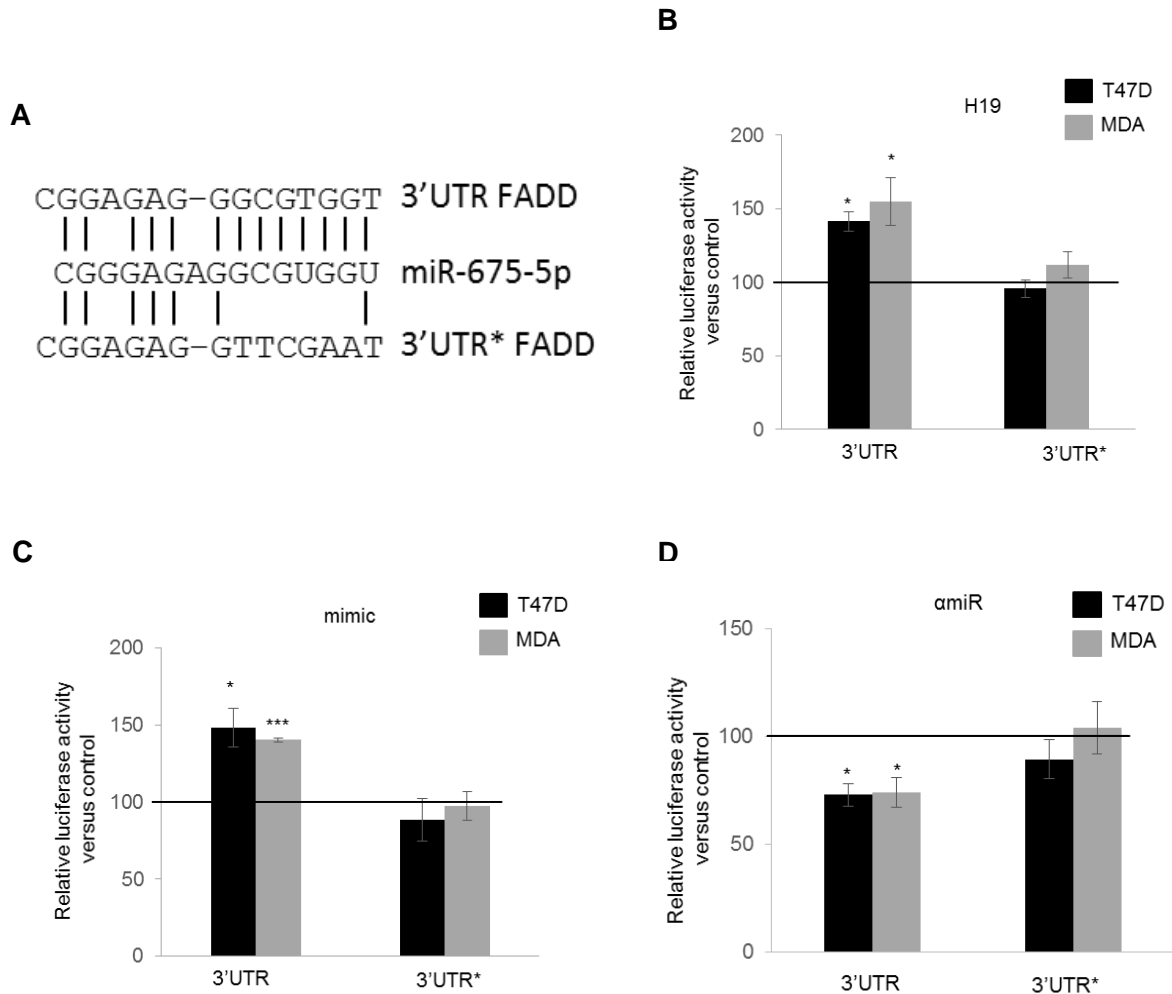


Figure 27. Le miR-675 régule directement *FADD* en se fixant sur l'UTR3'. (A) Interaction prédictive entre le miR-675 et l'UTR3' *FADD*. Les mutations générées dans le site de fixation du miR-675 sont indiquées sur la séquence 3'UTR*. (B, C, D) Activité Firefly luciférase dans les cellules cancéreuses mammaires après cotransfection d'un système rapporteur et d'un vecteur exprimant *H19* (H19) ou du miR-675 (mimic) ou miR-675 inhibiteur (amiR). L'activité Firefly luciférase est mesurée avec le dual-luciférase reporter assay (Promega) et normalisée à l'activité de la Renilla luciférase. Les plasmides sont transfectés avec *H19* ou le mimic ou l'antimir, ou leur contrôle dans des lignées cellulaires cancéreuses mammaires (T47D and MDA-MB-231). Les données représentent la moyenne de 3 expériences indépendantes par rapport à leur contrôle respectif en pourcentage et la barre d'erreur, l'écart à la moyenne. * $p < 0,05$ ** $p < 0,005$ *** $p < 0,001$

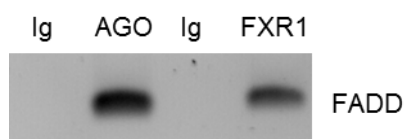


Figure 28. Détection par PCR des ARNm associés à la protéine AGO2 ou FXR1. Les ARN associés aux protéines AGO2 ou FXR1 sont immunoprécipités (RIP) et détectés par PCR.

Résultats

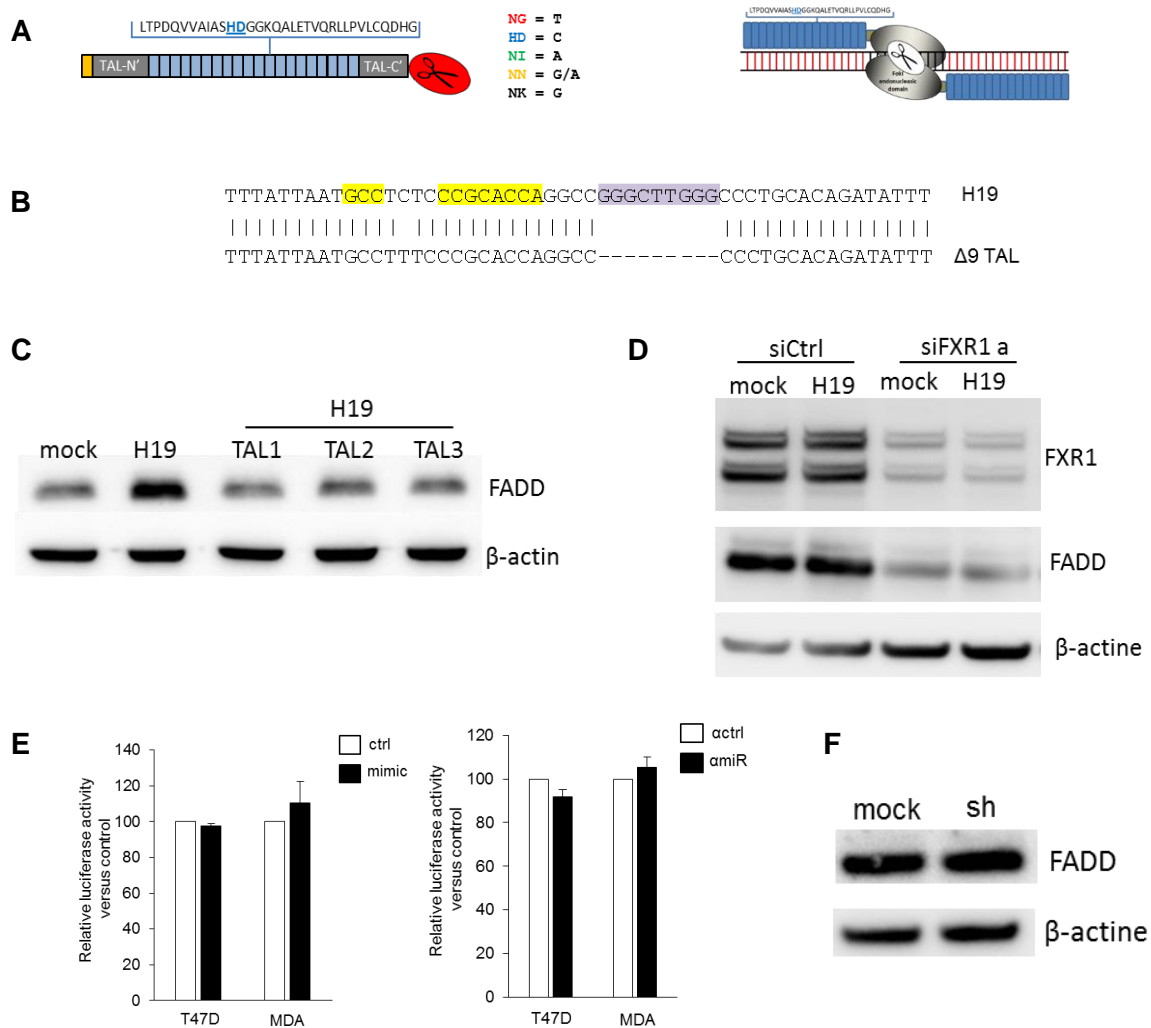


Figure 29. FXR1 recrute le miR-675 et permet son action positive. (A) A gauche : structure d'un TALEN. Les TALENs sont des dérivés de facteurs de transcription associés à l'endonucléase FokI. Ils sont composés de modules interchangeable (rectangle bleu) reconnaissant une base spécifique grâce à deux acides aminés (en gras et en bleu). Il existe 5 modules différents. A droite : association de deux TALENs sur une séquence spécifique permettant de reconstituer l'endonucléase FokI active. (B) Séquencage de l'UTR3' FADD dans des cellules surexprimant *H19* (H19) et dans des cellules surexprimant *H19* transfectées avec les TALENs ($\Delta 9$ TAL). En jaune : site de fixation du miR-675 ; en violet : site de fixation de FXR1 (C) Expression de la protéine FADD détectée par western blot dans des MDA-MB-231 contrôles (mock), surexprimant *H19* (H19) et surexprimant *H19* transfectées avec les TALENs (TAL). (D) Expression des protéines FXR1 et FADD, déterminée par western blot, dans des cellules contrôles (mock) ou surexprimant *H19* (H19) transfectées avec un siRNA ctrl ou un siRNA FXR1. (E) Activité Firefly luciférase dans les cellules cancéreuses mammaires (T47D and MDA-MB-231) après cotransfection du plasmide pMiR-3'UTR FADD muté sur le site de fixation FXR1 et du miR-675 (mimic) ou miR-675 inhibiteur (α mir). Les données représentent la moyenne de 3 expériences indépendantes par rapport à leur contrôle respectif en pourcentage et la barre d'erreur, l'écart à la moyenne. (F) Expression de la protéine FADD détectée par western blot dans des MDA-MB-231 contrôles (mock) ou invalidées pour 91H (sh).

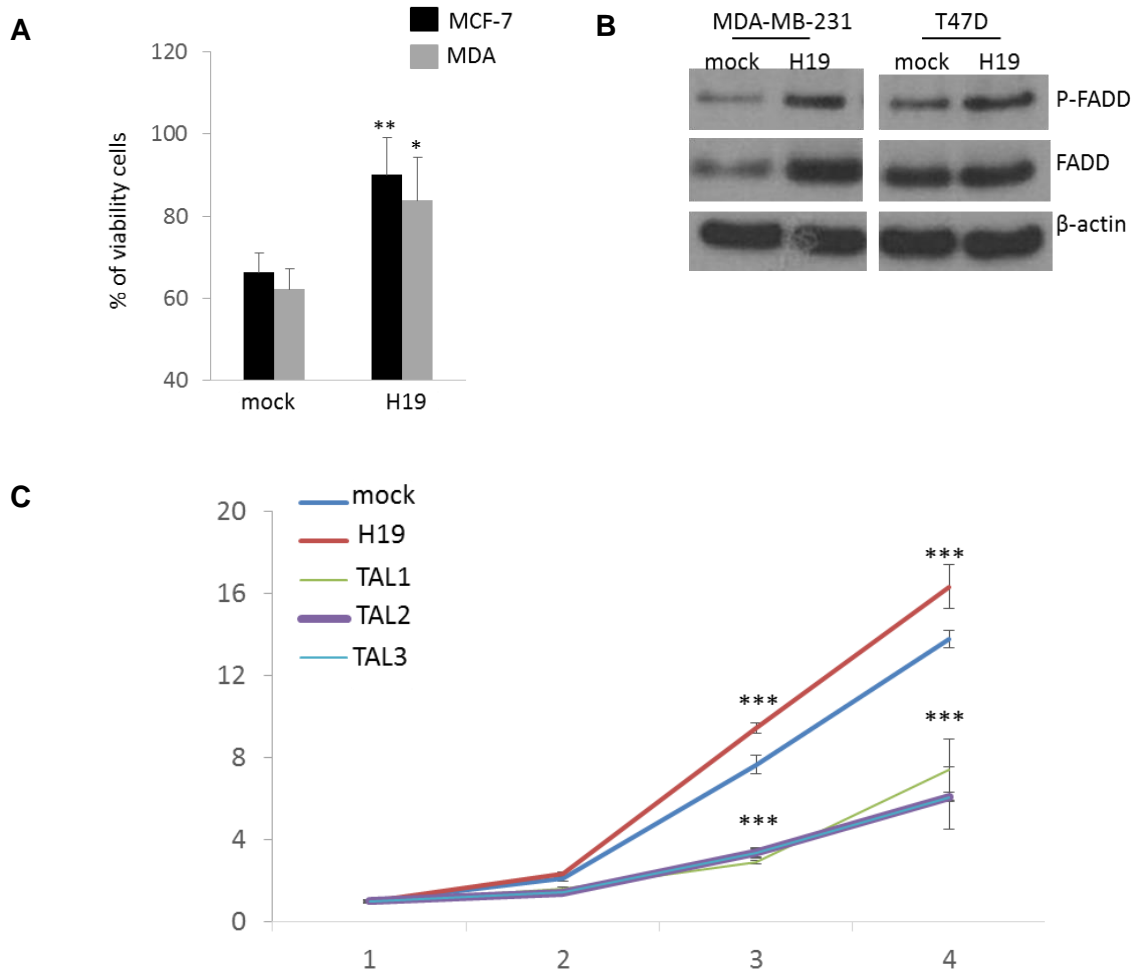


Figure 30. H19 permet la survie et la prolifération cellulaire. (A) Des cellules contrôles (mock) et surexprimant *H19* (H19) sont traitées au Trail (5 μ g/ml) pendant 6hrs et la viabilité cellulaire est déterminée par un test MTT. Deux lignées cellulaires sont utilisées (MDA-MB-231, MCF-7). (B) Détection de p-FADD et FADD par western blot dans des MDA-MB-231 et T47D contrôles (mock) et surexprimant *H19* (H19). (C) Croissance cellulaire relative des MDA-MB-231 contrôles (mock), surexprimant *H19* (H19) et trois clones TAL. * $p < 0,05$ ** $p < 0,01$ *** $p < 0,005$

primer	5'to 3'
qRT-PCR	
RPLP0 sense	GTGATGTGCAGCTGATCAAGACT
RPLP0 antisense	GATGACCAGCCCAAAGGAGA
H19 sense	TACAACCACTGCACTACCTG
H19 antisense	TGGCCATGAAGATGGAGTCG
FADD sense	TGGCTGGGGTGGGGTGGGGAGAC
FADD antisense	CTGCCTTGGCAATTCTGTTATCAG
Myc sense	CAGCAGCTCGAATTTCTTCC
Myc antisense	TTCGGGTAGTGGAAAACCG
cloning	
3'UTR FADD sense	ATACTAGTGGGCCGCTGCTTT
3'UTR FADD antisense	CGTTCGAATTAGGAATAGAAAAT
pmir XhoI	ATTGCTCGAGTGATGAAAGC
pmir NaeI	TCTATGCCGGCCTATGAGCTC
mut FADD 5p	AAGCCCGCCTAAGCTTGGAAGAGGCATTAAT
mut FADD 3p	ATTAATGCCTCTCCAAGCTTAGGCCGGGCTT
FADD RH FXR forward	ATAAAAGATCCTTTATTAAGTCTCCTGCAGTTAAC
FADD RH FXR reverse	AGTGATGAAAGCTGCGCATGGGCCGCTGCTTTGCGCTG
mut FADD FXR 5p	AAATATCTGTGCAGGGGGCCTGGTGC GGGA
mut FADD FXR 3p	TCCCGCACCCAGGCCCTGCACAGATATT
siRNA	
siFXR1 sense	CGAGCUGAGUGAUUGGUCA
siFXR1 antisense	UGACCUACUCAGCUCG
siGFP sense	GCUGACCCUGAAGUUCAUCTT
siGFP antisense	AAGAUGAACUUCAGGGUCAGC

Table 7. Séquence des amorces et siRNA utilisés dans l'étude.

II. Régulation de la protéine P53 par H19

Introduction

Pour exercer leur fonction, les longs ARN non codants peuvent avoir plusieurs modes d'action différents. En effet, ils peuvent réguler l'expression des gènes par modifications épigénétiques, être précurseur de microARN, être des éponges à microARN ou encore interagir avec des protéines pour en moduler leur fonction ou leur localisation cellulaire. L'ARN non codant *H19*, en plus d'être un précurseur et une éponge à microARN, est capable de se fixer sur les protéines EZH2 et P53 dans les cancers gastrique et de la vessie pour en moduler leurs fonctions (Yang et al., 2012 ; Zuo et al., 2013). Nous avons déjà montré que l'expression du gène *H19* est régulée négativement par la protéine P53 dans les cellules cancéreuses mammaires (Dugimont et al., 1998). Ici, nous montrons que l'ARN *H19* se fixe sur la protéine P53 dans les cellules cancéreuses mammaires et empêche sa fonction. Nous montrons que ce nouveau mode d'action de l'ARN *H19* favorise la progression du cycle cellulaire de cellules cancéreuses porteuses de mutations. Ainsi, *H19* est également impliqué dans la transformation d'une cellule cancéreuse en une cellule cancéreuse agressive.

Matériels et Méthodes

Culture cellulaire

Les cellules cancéreuses mammaires MCF-7 et T47D sont cultivées en RPMI1640 (Gibco) supplémenté de sérum de veau fœtal 10% et de zell shield (Biolabs). Les cellules cancéreuses de poumon Calu-6, n'exprimant pas la protéine P53, sont entretenues en MEM (Gibco) supplémenté de sérum de veau fœtal 10% et de Zell Shield (Biolabs). Elles sont maintenues sous atmosphère humide à 37°C, 5% CO₂.

Transfection

Les plasmides permettant l'expression de *P53* et d'*H19* sont transfectés dans les Calu-6 préalablementensemencées à 20x10⁴ cellules/puits de plaque 6 puits. La quantité désirée de plasmides est ajoutée à 4µl d'Exgen 500 (Euromedex) préalablement dilué dans 200µl d'OptiMEM selon les recommandations du fournisseur. Afin de transfecter la même quantité d'ADN dans les cellules, nous transfectons également un plasmide vide pcDNA3.1(-) (Invitrogen). La quantité totale d'ADN ainsi transfectée est de 1,6µg de plasmide. Après 6 heures de transfection à 37°C, le milieu des cellules est changé. Les protéines sont extraites 48 heures après transfection.

Pour les tests d'activité luciférase, 1µg de plasmide P53-pGL4 et 12ng de plasmide pRL-null sont transfectés avec 4µl d'Exgen 500 (Euromedex) selon les recommandations du fournisseur.

Résultats

Clonage

La construction des plasmides permettant l'expression des gènes *P53* et *H19* est décrite, respectivement, par Dugimont et al., 1998 et dans l'article 2.

Les amorces possédant des éléments de réponse à P53 (P53 sense 5'-TCAGACATGCCTAGACATGCCTAACCGGTTAGACATGCCTAGACATGCCTTAAG-3' ; P53 antisense 5'-CTTAACCGATGTCTAGGCATGTCTAACCGGTTAGGCATGTCTAGGCATGTCTGA-3') sont hybridées dans un gradient de 95°C à température ambiante. Les dimères sont ensuite clonés dans le plasmide pGL4.10 [luc2] (Invitrogen) entre les sites de restriction Kpn/XhoI.

Les différentes constructions sont transformées dans des bactéries compétentes TOP 10 (Invitrogen). Après culture liquide bactérienne, les plasmides sont extraits à l'aide du kit Nucleobond PC100 (Macherey-Nagel) selon les recommandations du fournisseur.

qRT-PCR

Les ARNm sont extraits et quantifiés par qRT-PCR selon le protocole décrit dans le matériels et méthodes de l'article 1.

Les amorces de qRT-PCR sont : RPLP0 sense 5'-GTGATGTGCAGCTGATCAAGACT-3', RPLP0 antisense 5'-GATGACCGCCCAAAGGAGA-3', H19 sense 5'-TACAACCACTGCACTACCTG-3', H19 antisense 5'-TGGCCATGAAGATGGAGTCG-3', P21 sense 5'-CACTCCAAACGCCGGCTGATCC-3', P21 antisense 5'-TGTAGAGCGGGCCTTTGAGGCCCTC-3', Bax sense 5'-TTTGCTTCAGGGTTTCATCC-3', Bax antisense 5'-CAGTTGAAGTTGCCGTCAGA-3', P53 sense 5'-GTGGAAGGAAATTTGCGTGT-3', P53 antisense 5'-CCAGTGTGATGATGGTGAGG-3'

Western blot

Les protéines sont extraites dans du RIPA supplémenté d'inhibiteurs de phosphatases et de protéases, comme décrit dans l'article 1. Les protéines MDM2, P53, P21, Bax et β -actin sont révélées avec les anticorps suivant : anti-MDM2 (sc-965, Santa Cruz), anti-P53 (Do-1, sc-126, Santa Cruz), anti-P21 (2947P, Cell Signaling), anti-Bax (sc-493, Santa Cruz) et anti- β -actin (sc-47778, Santa Cruz). Le β -actin est utilisée comme témoin de charge.

Test luciférase

20x10⁴ cellules/puits sont ensemencées en plaques 6 puits pendant 24 heures. Après 24 heures de transfection à l'Exgen 500 de 1 μ g de plasmide P53-pGL4 et 12ng de plasmide pRL-null selon les recommandations du fournisseur, les cellules sont lysées dans 200 μ l de passive lysis buffer (Promega). L'activité luciférase est ensuite dosée avec le kit « Dual-Go luciferase assay » (Promega), comme décrit dans l'article 1.

Immunoprécipitation des ARNm

L'immunoprécipitation des ARNm associés à la protéine P53 est réalisée avec le kit « Magna RIP RNA-binding protein immunoprecipitation kit » (17-700, Millipore) et l'anticorps anti-P53 selon les recommandations du fournisseur (17-613, Millipore).

Co-immunoprécipitation

Après traitement au MG132 à 10 μ M pendant 6 heures (#474790, Calbiochem), les cellules sont lysées avec le tampon Hepes (HEPES 40mM pH 7,5 ; EDTA 1mM pH 8,0 ; NaCl 120mM ; NaPPi 10mM ; Triton X-100 1% ; SDS 0,1% ; et glycérol 10%) supplémenté d'inhibiteurs de protéases et phosphatases (NaF 50mM ; orthovanadate de sodium 1,5mM ; PMSF 1mM ; protéine cocktail inhibitor 1/100). Après centrifugation à 14000rpm, 4°C, 10 minutes, les protéines sont récupérées et dosées au BCA (Pierce). Les billes de protéine G sont saturées avec 1mg de protéine 1 heure à 4°C sur roue. Le surnageant est récupéré et mis en contact avec 5 μ G d'anticorps anti-P53 (Do-1, sc-126, Santa Cruz), anti-MDM2 (sc-965, Santa Cruz) pendant 2 heures sur roue à 4°C. Les billes saturées sont ensuite ajoutées au complexe anticorps-protéine pendant 2 heures sur roue à 4°C. Après trois rinçages des billes avec du tampon de lyse complet, les protéines immunoprécipitées sont éluées et dénaturées par l'ajout de 80 μ l de tampon Laemmli (Tris 12,5mM ; glycérol 2% ; SDS 0,4% ; β -mercaptoéthanol 1% ; bleu de bromophénol) et chauffage à 95°C pendant 7minutes.

Les protéines sont séparées par électrophorèse sur gel d'acrylamide 10% à 150V pendant 1h30. Elles sont transférées sur une membrane de PVDF (polyvinylidène de fluorure) dans un tampon Tris-Gly 25mM, 15% méthanol à 105V pendant 1h30 à 4°C. La membrane est saturée avec de la caséine 0,2% pendant 1 heure à température ambiante. Les anticorps primaires (anti-MDM2, anti-P53) dilués dans du PBS/BSA 5% sont hybridés sur la membrane toute la nuit à 4°C sur agitation lente. Après rinçages au TBS/Tween 0,1%, les anticorps secondaires couplés à la peroxidase sont hybridés sur la membrane pendant 30 minutes à température ambiante. Les membranes sont analysées à l'aide d'une solution de chimiluminescence (SuperSignal west Pico enhancer solution, Pierce) et révélées avec un appareil d'imagerie de chimiluminescence (LAS4000, Fujifilm).

Immunocytochimie

Les cellules sontensemencées sur lamelles pré-traitées au collagène à raison de 30.10⁴ cellules/par lamelle pendant 24 heures. Elles sont ensuite traitées à la doxorubicine (D1515, Sigma-Aldrich) à 0,2 μ g/ml pendant 6 ou 24 heures. Après fixation au PAF 4% pendant 10 minutes, les cellules sont perméabilisées au méthanol pendant 10 minutes à -20°C puis saturées avec l'Ultravision Protein Block (TA-125-PBQ, Thermo Scientific) pendant 10 minutes à température ambiante. L'anticorps primaire, dilué au 1/500, est ajouté sur les lamelles toute

Résultats

la nuit à 4°C. Le lendemain, après rinçages des lamelles au PBS tween, l'anticorps secondaire couplé à l'Alexa Fluor 594 (A211020, Invitrogen) dilué au 1/500 y est déposé pendant 1 heure à température ambiante. Le noyau est ensuite marqué au DAPI (DI306, Invitrogen) pendant 5 minutes. Les cellules sont ensuite analysées à l'Axioplan2.

Cycle cellulaire

30x10⁴ cellules sontensemencées en boîte 60mm pendant 24 heures. Après traitement à la doxorubicine à 0,1µg/ml pendant 24 heures, les cellules sont perméabilisées à l'éthanol 70% froid toute la nuit. Après rinçages au PBS, elles sont marquées à l'Iodure de propidium (1µg/ml) en présence de RNase A (1mg/ml) (R6513, Sigma-Aldrich) pendant 30 minutes à 37°C. Les cellules marquées sont analysées au cytomètre Beckman Coulter. Les données sont traitées avec le logiciel WinCycle.

Test statistique

Les données sont analysées avec le t-Test de Student. Une valeur $p < 0,05$ est considérée comme significative.

Résultats

H19 réprime l'expression de P53

Dans les cancers gastriques, l'association *H19/P53* a déjà été identifiée (Yang et al., 2012). Dans nos cellules cancéreuses mammaires, nous avons détecté cette association par immunoprécipitation (RIP) des ARNm associés à la protéine P53 (fig 31A). Par qRT-PCR des ARNm immunoprécipités, nous remarquons que *H19* est associé à la protéine P53 dans les cellules cancéreuses mammaires MCF-7. Nous avons ensuite déterminé l'effet de cette interaction dans les cellules. Dans un premier temps, j'ai analysé l'expression du gène *P53* lorsque *H19* est surexprimé. De manière surprenante, une diminution d'expression de P53 est observée aussi bien au niveau protéique qu'au niveau ARNm (fig 31B, C). Nous avons également mis en évidence une diminution d'expression de certains gènes cibles de P53 : *P21* et *Bax*, dans ces cellules par qRT-PCR (fig 31C). Ces résultats doivent être confirmés dans une seconde lignée cellulaire cancéreuse mammaire (T47D).

Pour déterminer si *H19* régule transcriptionnellement l'expression de P53, j'ai transfecté dans des cellules cancéreuses de poumon dépourvues de P53 (Calu-6), un plasmide permettant l'expression de P53 sous un promoteur fort ainsi que le plasmide permettant l'expression d'*H19* (fig 31D). Par western blot, nous observons qu'en présence d'*H19*, l'expression de P53 est diminuée de manière dose-dépendante (fig 31D). De plus, une forte expression d'*H19* est associée à une très faible expression de P53. Ces données suggèrent que *H19* régule négativement l'expression de P53 de manière post-transcriptionnelle en se fixant sur cette protéine.

P53 étant impliquée dans les mécanismes de réparation des dommages à l'ADN, elle est exprimée de façon continue dans les cellules. Son association avec la protéine MDM2 permet sa dégradation continue par le protéasome. Par ailleurs, lorsque *H19* est surexprimé, une forte expression de MDM2 a été détectée par western blot (fig 32A). Nous avons alors déterminé si l'association MDM2/P53 est favorisée dans les cellules cancéreuses mammaires T47D surexprimant *H19*, par co-immunoprécipitation (fig 32B). Afin d'augmenter la quantité de protéines immunoprécipitées, les cellules ont été traitées au préalable avec du MG132. Nous observons que l'interaction MDM2/P53 est augmentée lorsque *H19* est surexprimé (fig 32B). Ces données doivent être confirmées dans une seconde lignée cellulaire cancéreuse mammaire (MCF-7). Ainsi, *H19* réprime l'expression de P53 en favorisant les interactions MDM2/P53.

H19 empêche la fonction de P53

Nous avons également déterminé les effets de *H19* sur les fonctions de P53 dans la cellule. Un plasmide contenant des éléments de réponse à P53 en amont de la luciférase a été transfecté dans des cellules cancéreuses mammaires (MCF-7) surexprimant stablement *H19* et l'activité luciférase étudiée (fig 33A). Nous remarquons qu'en présence du gène *H19*, l'activité luciférase du plasmide est diminuée de 50% par rapport aux cellules contrôles (fig 33A). Ces tests ont également été réalisés lors d'un traitement des cellules à la doxorubicine. Les résultats préliminaires ont montré une diminution d'activité luciférase du plasmide dans les cellules surexprimant *H19* après traitement (données non montrées). Ces tests doivent être reproduits pour être soumis à des analyses statistiques et réalisés dans une autre lignée cellulaire.

Par ailleurs, lors d'un traitement à la doxorubicine, une diminution d'expression de quelques gènes cibles de P53 est constatée lorsque *H19* est surexprimé (fig 33B). Ces effets sont retrouvés dans deux lignées cellulaires (MCF-7 et T47D, données non montrées). Nous observons également une diminution de l'activation de plusieurs gènes impliqués dans la régulation du cycle cellulaire (Rb, Chk1, Chk2 ainsi que CDC2) lorsque les cellules surexpriment *H19* (fig 33B). Lors de dommages à l'ADN, P53, en activant P21, induit un arrêt des cellules dans le cycle cellulaire en phase G2/M. Nous avons donc étudié le cycle cellulaire des cellules T47D contrôles et surexprimant *H19* lors d'un traitement à la doxorubicine. Lorsque *H19* est surexprimé, moins de cellules sont arrêtées en phase G2/M (fig 33C). Ces données doivent être confirmées dans une seconde lignée cellulaire, mais *H19* favoriserait la progression du cycle cellulaire même en présence de dommages à l'ADN.

Enfin, P53 est un facteur de transcription qui, lors de dommages à l'ADN, est transloqué dans le noyau pour permettre l'expression de ses gènes cibles (entres autres, *P21* et *Bax*) responsables de ses effets phénotypiques. *H19* est un long ARN non codant principalement

Résultats

cytosolique (Brannan et al., 1990). Lors de dommages à l'ADN, nous avons donc détecté la localisation cellulaire de P53 dans les cellules contrôles et surexprimant *H19* par immunofluorescence. Nous remarquons que le traitement à la doxorubicine permet le transport de P53 dans le noyau dans les cellules contrôles (fig 34). Nous observons également que dans les cellules surexprimant *H19*, la protéine P53 est retenue dans le cytosol. L'ensemble de ces résultats montrent que *H19* est également capable de réprimer les fonctions de P53 dans les cellules en le retenant dans le cytosol.

Discussion

Dans cette étude, nous avons mis en évidence un nouveau mode d'action de l'ARN *H19* dans les cellules cancéreuses mammaires. Nous montrons que *H19* en fixant la protéine P53 régule celle-ci par différents mécanismes. En effet, nous montrons que *H19* favorise l'association MDM2/P53 dans les cellules cancéreuses mammaires. Ce mode d'action peut être responsable de la dégradation de P53 observée lorsque *H19* est surexprimé. De plus, nous montrons que *H19* empêche les fonctions de P53 lorsque les cellules présentent des dommages à l'ADN en le piégeant dans le cytoplasme. Ainsi, par deux mécanismes distincts mais complémentaires, *H19* régule négativement l'expression et les fonctions de P53 dans les cellules cancéreuses mammaires. Cependant, ces deux mécanismes peuvent être inscrits dans une cinétique de temps : *H19* piège P53 dans le cytosol, ce qui favorise ensuite son association avec la protéine MDM2 et sa dégradation. Ce nouveau mode d'action d'*H19* ainsi mis en lumière nous montre que cet ARN favorise l'instabilité génétique cellulaire.

La protéine P53 a été découverte en 1979 dans certains sarcomes chez la souris (DeLeo et al., 1979). Dans un premier temps, cette protéine était associée uniquement au noyau puis sa présence dans le cytosol a été démontrée (Dippold et al., 1981 ; Rotter et al., 1983). P53 est un facteur de transcription, de 53kDa, constitué de plusieurs domaines : un domaine transactivateur, un domaine riche en proline, un domaine de liaison à l'ADN, un domaine de tétramérisation ainsi qu'un domaine Ct. Elle est exprimée de manière constitutive et est continuellement dégradée par le protéasome. En effet, en condition normale, la protéine MDM2, une E3 ubiquitine ligase, se fixe sur le domaine transactivateur de la protéine P53 et est responsable de son ubiquitinylation sur son domaine Ct. Cette modification post-traductionnelle induit la dégradation de P53 par le protéasome de manière constitutive. Lors de stress cellulaire (dommages à l'ADN (cassures simples ou doubles brins), irradiations, stress génotoxiques), les protéines ATM et ATR activées vont phosphoryler les protéines P53 et MDM2 sur leur domaine Nt, le complexe MDM2/P53 sera alors dissocié, la protéine P53 sera active tandis que MDM2 sera dégradée par le protéasome (Toledo et al., 2007). Une troisième protéine, MDMX ou MDM4, de structure homologue à MDM2, peut être associée au complexe

MDM2/P53. Cependant, cette dernière ne possède pas de fonction E3 ubiquitine ligase, elle n'est donc pas responsable de la dégradation de P53 en condition normale (Toledo et al., 2007). Néanmoins, ces deux protéines, MDM2 et MDM4, sont indispensables dans le développement embryonnaire puisque leur *knock-out* induit la mort *in utero* de l'embryon. Ce sont donc des inhibiteurs essentiels et spécifiques de P53 à ce stade de développement (Toledo et al., 2007). En fonction de l'importance des dommages à l'ADN, P53 active l'expression de différents gènes cibles, notamment *P21* ou *Bax* pour induire soit un arrêt du cycle cellulaire soit l'apoptose de la cellule (Freeman & Espinosa, 2013). Il est admis que lors de dommages réparables, un faible taux de P53 est activé, il permettra l'expression de la protéine P21 et donc un arrêt du cycle cellulaire. Lorsque les dommages sont trop importants, une grande quantité de P53 est libérée, il permettra dans ce cas, l'expression de gènes impliqués dans l'apoptose de la cellule, *Bax* par exemple. Ainsi, la protéine P53 est qualifiée de « gardien de génome » puisqu'elle empêche la prolifération de cellules porteuses de dommages à l'ADN pouvant conduire à la transformation d'une cellule normale.

Nous montrons qu'en présence d'*H19* et de dommages à l'ADN, l'activation de P21 est plus faible. Cet ARN empêche donc l'activation de P53 et permet la prolifération de cellules cancéreuses porteuses de nouvelles mutations. Il est donc impliqué dans la progression tumorale, notamment dans la transformation de cellules cancéreuses en cellules cancéreuses agressives. Nous avons ainsi mis en évidence l'un des mécanismes permettant les propriétés oncogéniques associées à *H19*. Nous montrons également un nouveau mécanisme de régulation de l'action de P53 dans les cellules. Par ailleurs, une diminution d'expression de *Bax* est constatée lorsque *H19* est surexprimé. Il serait donc intéressant de déterminer les effets de cette régulation dans les cellules cancéreuses.

Nous montrons également que *H19* favorise la dégradation de P53 en augmentant les associations MDM2/P53. Pour compléter ce travail, il serait intéressant de déterminer si *H19* peut également se fixer sur la protéine MDM2 pour favoriser l'association P53/MDM2 et former ainsi un complexe ternaire. Nous devons également déterminer si l'expression des autres gènes cibles de P53 (*PUMA*, *14-3-3 ζ*) est également diminuée en présence d'*H19* et étudier les effets phénotypiques associés à leur expression.

Au regard de ce nouveau mode d'action, il serait intéressant de déterminer les effets de l'expression d'*H19* dans une cellule normale. En effet, nous pouvons supposer que *H19* soit impliqué dans la transformation d'une cellule normale en une cellule cancéreuse en favorisant l'instabilité génétique.

Dans de nombreux cancers chez l'Homme (50%), la protéine P53 est mutée, elle n'est donc pas fonctionnelle (Chao et al., 2015). Néanmoins, plusieurs études cliniques ont déterminé que

Résultats

l'utilisation d'inhibiteurs spécifiques de MDM2 (peptides, anticorps, séquence oligonucléotidique antisens) permet de réduire la croissance tumorale lorsque P53 est fonctionnelle (Toledo et al., 2007). Cependant, ses inhibiteurs ont un taux de toxicité élevé pour les tissus environnant. Grâce au nouveau mécanisme de régulation de P53 mis ici en lumière, de nouvelles stratégies thérapeutiques pourront être envisagées, dans les cancers où P53 est fonctionnelle.

Dans cette étude, nous avons mis en évidence un nouveau mode d'action de l'ARN H19 ainsi qu'un nouveau mécanisme de régulation de P53 dans les cellules cancéreuses mammaires permettant la prolifération des cellules tumorales porteuses de nouvelles mutations.

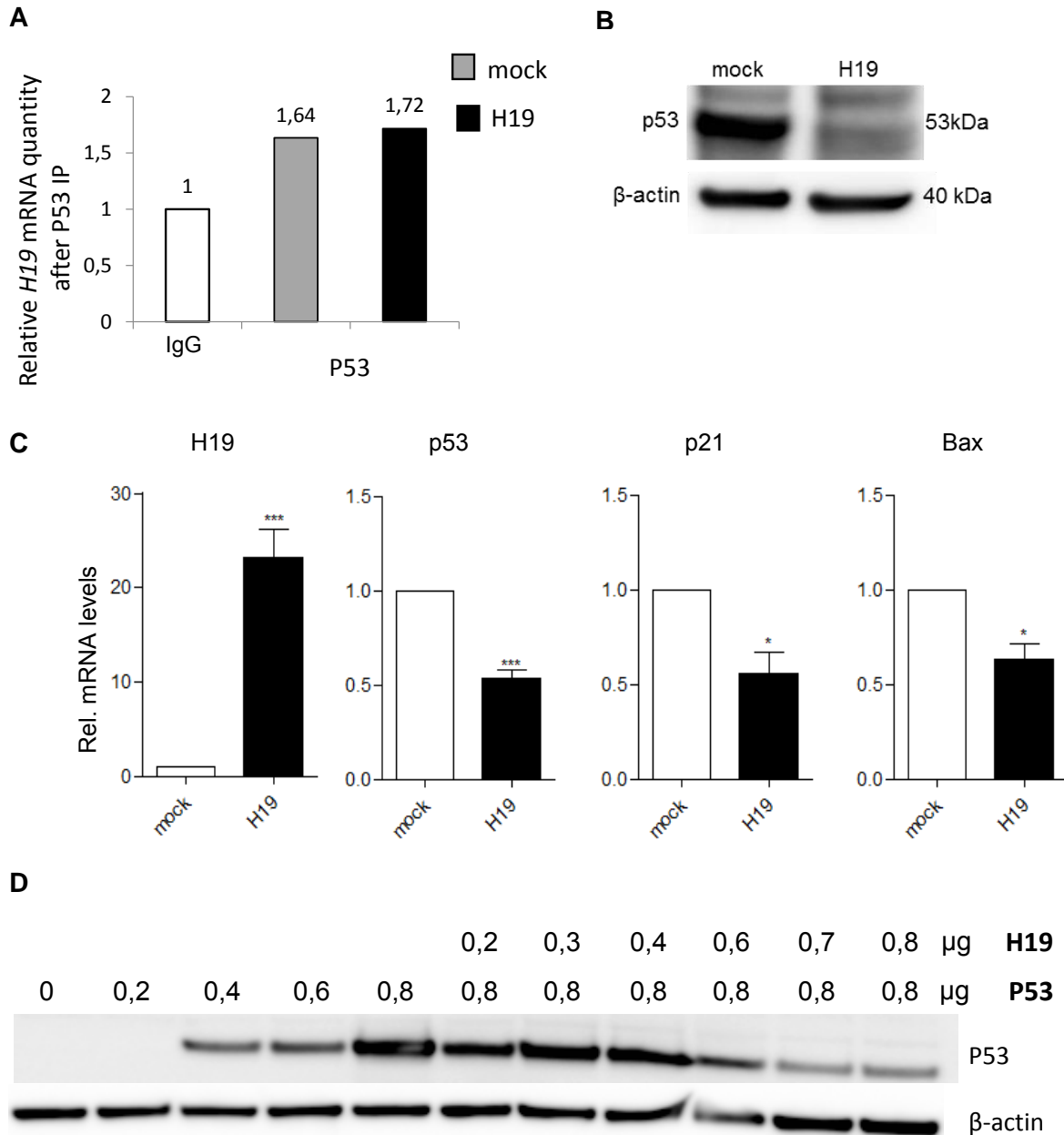


Figure 31. H19 fixe P53 et induit sa dégradation. (A) Immunoprécipitation des ARNm associés à la protéine P53 dans les cellules cancéreuses mammaires (MCF-7) contrôles (mock) ou surexprimant *H19* (H19). Les ARNm associés à cette protéine sont détectés par qRT-PCR. **(B)** Expression de P53 dans les cellules cancéreuses mammaires (MCF-7) contrôles (mock) ou surexprimant *H19* (H19) déterminée par western blot. **(C)** Expression relative des gènes *H19*, *P53*, *P21* et *Bax* déterminée par qRT-PCR dans les cellules cancéreuses mammaires (MCF-7) contrôles (mock) ou surexprimant *H19* (H19). **(D)** Expression de P53 déterminée par western blot dans les cellules P53 null (Calu-6) après transfection d'une dose croissante d'un plasmide exprimant P53 et d'une dose croissante d'un plasmide exprimant *H19*. La quantité d'ADN transfectée est constante dans toutes les conditions. Elle est donc complétée à 1,6μg d'ADN avec un plasmide vide pcDNA 3.1(-). *p<0,05 **p<0,01 ***p<0,005

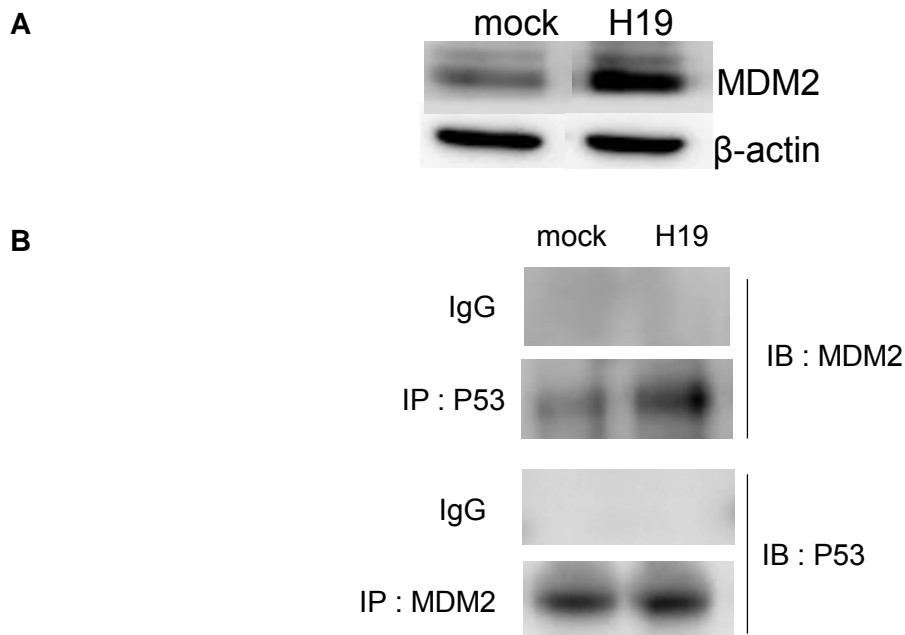


Figure 32. H19 favorise la dégradation de P53. (A) Expression de MDM2 dans les cellules cancéreuses mammaires (MCF-7) contrôles (mock) ou surexprimant *H19* (H19) déterminée par western blot. **(B)** Co-immunoprécipitation des protéines P53 et MDM2 dans les cellules cancéreuses mammaires (T47D) contrôles (mock) ou surexprimant *H19* (H19). Détection des protéines co-immunoprécipitées par western blot.

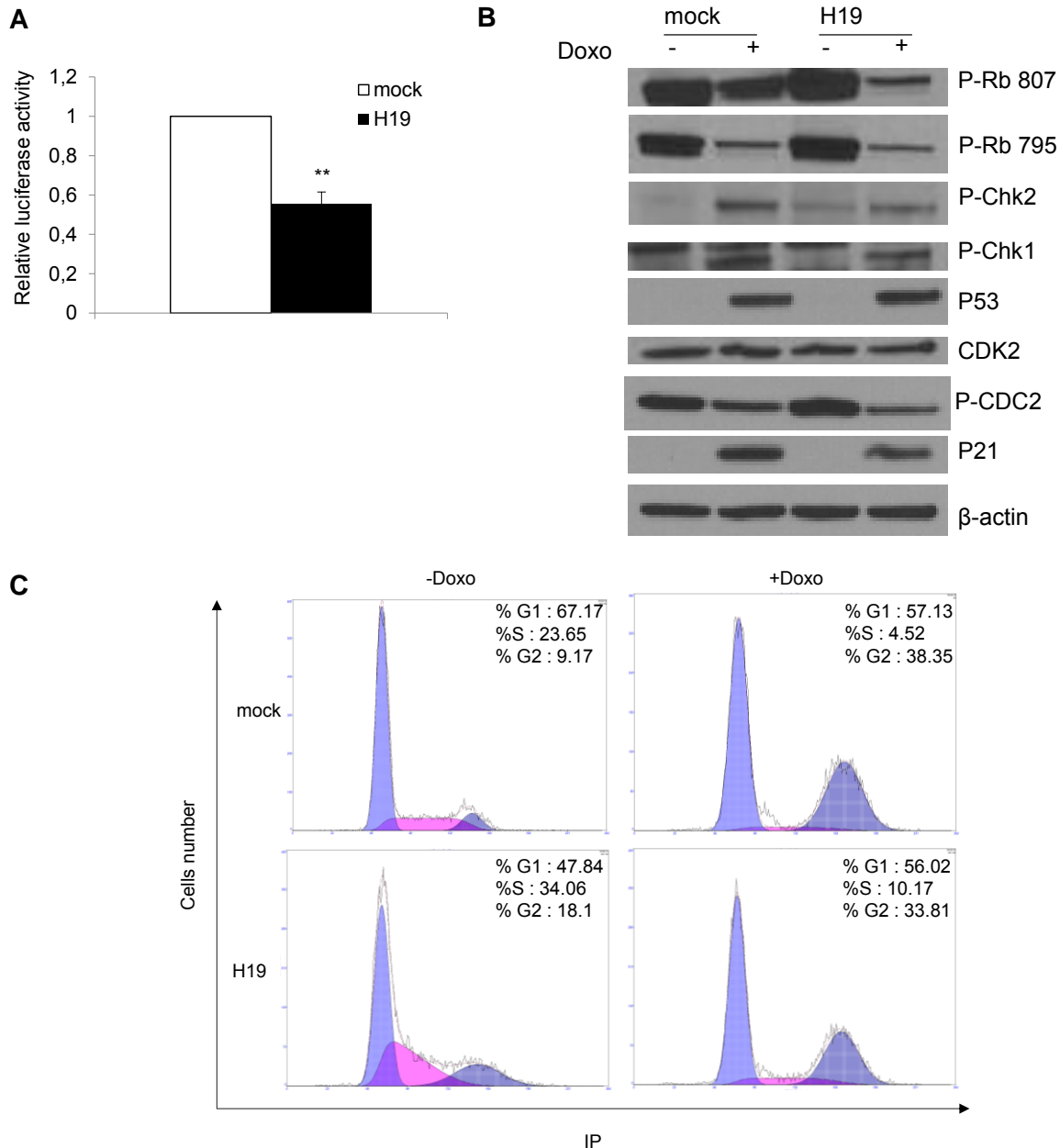


Figure 33. H19 empêche les fonctions de P53. (A) Test d'activité luciférase du plasmide P53-pGL4 luc transfecté dans les cellules cancéreuses mammaires (MCF-7) contrôles (mock) ou surexprimant *H19* (H19). **(B)** Activation de protéines impliquées dans la régulation du cycle cellulaire (Rb, Chk1, Chk2, P53, CDK2, CDC2, P21) déterminée par western blot dans les cellules cancéreuses mammaires (MCF-7) contrôles (mock) ou surexprimant *H19* (H19) et traitées à la doxorubicine 0,2µg/ml pendant 24 heures. **(C)** Analyse du cycle cellulaire des cellules cancéreuses mammaires (T47D) contrôles (mock) ou surexprimant *H19* (H19) et traitées à la doxorubicine 0,1µg/ml pendant 24 heures. **p<0,01

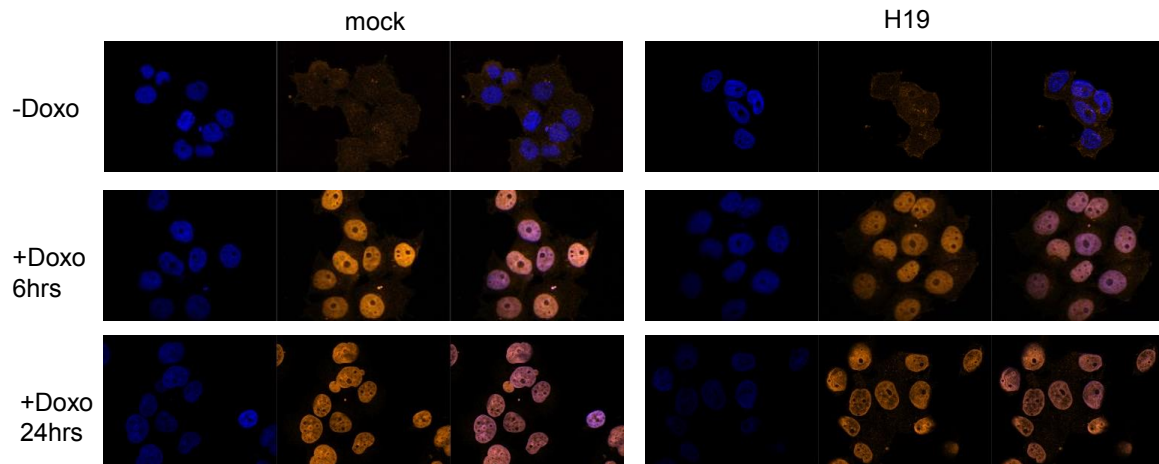


Figure 34. *H19* retient P53 dans le cytoplasme. Détection de la localisation cellulaire de P53 dans des cellules cancéreuses mammaires (MCF-7) contrôles (mock) ou surexprimant *H19* (H19) traitées à la doxorubicine 0,2 μ g/ml pendant 6 ou 24 heures par immunofluorescence.

III. H19 et les cellules souches

Introduction

Chez l'Homme, le cancer de la prostate est le 2^{ème} cancer le plus fréquent et la 6^{ème} cause de mortalité dans le monde. Cette forte mortalité est principalement due aux résistances des tumeurs aux thérapies conventionnelles mais aussi aux récives. Ces dernières sont issues des cellules souches cancéreuses (CSC) qui résistent aux thérapies. En effet, les thérapies conventionnelles ciblent les cellules tumorales en prolifération ; ainsi les CSC, qui sont quiescentes, ne sont pas éliminées avec ces traitements. Identifier et caractériser cette sous-population cellulaire est nécessaire pour mettre en place de nouvelles thérapies plus efficace.

Jusqu'à présent, l'isolement et la caractérisation des CSC nécessitaient l'utilisation d'une combinaison de marqueurs qui sont spécifiques pour chaque tissu cellulaire étudié. Avant de caractériser les CSC, il est important de caractériser les cellules souches dans le tissu normal.

Chez la souris, un marqueur unique de cellules souches prostatiques et mammaires a été identifié : il s'agit du gène *s-SHIP* (*stem-SH2-domain containing 5'-inositol phosphatase*) (Rohrschneider et al., 2005 ; Bai & Rohrschneider, 2010). En utilisant un système rapporteur couplant la GFP sous le contrôle du promoteur *s-SHIP*, l'équipe du Dr Bourette a isolé et caractérisé une population de cellules souches prostatiques normales chez l'Homme. En collaboration avec cette équipe, nous avons mis en évidence l'implication de l'ARN non codant *H19* dans ces cellules souches prostatiques normales humaines.

Article 2.

Hélène Bauderlique-Le Roy, **Constance Vennin**, Guillaume Brocqueville, Nathalie Spruyt, Eric Adriaenssens and Roland P. Bourette (2015). Enrichment of human stem-like prostate cells with s-SHIP promoter activity uncovers a role in stemness for the long non coding RNA H19. *Stem cells and development*, 24, 1252-1262.

Enrichment of human stem-like prostate cells with s-SHIP promoter activity uncovers a role in stemness for the long non-coding RNA *H19*

Hélène Bauderlique-Le Roy¹, Constance Vennin², Guillaume Brocqueville¹, Nathalie Spruyt¹,

Eric Adriaenssens^{2*#}, Roland P. Bourette^{1*#}

¹ Institut de Biologie de Lille, UMR 8161 CNRS, SIRIC ONCOLille, Institut Pasteur de Lille,

²INSERM U908, Université de Lille

* E.A. and R.P.B. contributed equally to this study

Address correspondence to:

Roland P. Bourette, Ph.D.

Institut de Biologie de Lille, UMR 8161 CNRS

1 rue du Professeur Calmette, CS5 0447, 59021 Lille Cedex, France

phone: (33) 3 20 87 12 96 / Fax: (33) 3 20 87 10 19

E-mail: roland.bourette@ibl.cnrs.fr

Eric Adriaenssens, Ph.D.

INSERM U908, Université de Lille

Cité Scientifique, Bat SN3, 59655 Villeneuve d'Ascq Cedex

phone: (33) 3 20 33 64 06 / Fax: (33) 3 20 43 40 38

E-mail: eric.adriaenssens@univ-lille1.fr

Running title : s-SHIP⁺ stem-like prostate cells express lnc RNA *H19*

Abstract

Understanding normal and cancer stem cells should provide insight into the origin of prostate cancer and their mechanisms of resistance to current treatment strategies. In this study, we isolated and characterized stem-like cells present in the immortalized human prostate cell line, RWPE-1. We used a reporter system with green fluorescent protein (GFP) driven by the promoter of s-SHIP (for stem-SH2-domain-containing 5'-inositol phosphatase) which stem cell-specific expression has been previously shown. We observed that s-SHIP-GFP-expressing RWPE-1 cells showed stem cell characteristics such as increased expression of stem cell surface markers (CD44, CD166, TROP2) and pluripotency transcription factors (*Oct4*, *Sox2*, *Nanog*), and enhanced sphere-forming capacity and resistance to arsenite-induced cell death. Concomitant increased expression of the long non-coding RNA *H19* was observed which prompted us to investigate a putative role in stemness for this oncofetal gene. Targeted suppression of *H19* with siRNA decreased *Oct4* and *Sox2* gene expression and colony-forming potential in RWPE-1 cells. Conversely, overexpression of *H19* significantly increased gene expression of these two transcription factors and the sphere-forming capacity of RWPE-1 cells. Analysis of *H19* expression in various prostate and mammary human cell lines revealed similarities with *Sox2* expression, suggesting that a functional relationship may exist between *H19* and *Sox2*. Collectively, we provide the first evidence that s-SHIP-GFP promoter reporter offers a unique marker for the enrichment of human stem-like cell populations and highlight a role in stemness for the long non-coding RNA *H19*.

Introduction

Prostate cancer is the second most frequently diagnosed cancer and the sixth leading cause of cancer death in males worldwide and the second in developing countries [1]. Characterizing the prostate cells that are more susceptible to transformation and capable of cancer maintenance represents an essential step to a better understanding and more efficient treatment of the disease. Conventional therapeutics target and kill proliferating cells, sparing the putative cancer stem cell fraction and allowing for the recurrence of the disease. Successful therapy must not only kill the proliferating tumor cells, but also eliminate or differentiate the cancer stem cells. Thus, improvements in isolating purified normal or cancer stem cell fractions will then be crucial for advancing therapeutic potential in this field [2].

In vitro cell culture may provide convenient models for studying normal and cancer prostate stem cells since stem cell properties reside naturally among populations of normal or neoplastic cells propagated in culture. Primary cell cultures represent an ideal approach to investigate the nature and functions of normal or cancer stem cells *in vitro*, but they also present inconvenient features, such as a short life span. To circumvent this difficulty, several studies used non-malignant-immortalized or malignant cell lines that contain sub-populations of cells with stem cell properties [3]. Enrichment in stem-like cell subpopulations could be achieved by multiple approaches: by modifying cell culture conditions like cultivating cells in low-calcium serum-free defined medium [4] or under hypoxia [5]; by using combination of cell surface markers, such as CD133 [6], CD44 [7], CD166 [8,9], or TROP2 [10]; by developing functional assays on cytoprotective activity such as the side population assay based on the efflux of Hoechst 33342 fluorescent dye by the ATP-binding cassette (ABC) transporter and the ALDEFLUOR assay based on activity of aldehyde dehydrogenases (ALDH) detoxifying enzyme activity [11-13]; or finally by using reporter vectors containing fluorescent protein driven by stemness gene promoters [14,15].

In the present study, we applied this latter approach using the green fluorescent protein (GFP) cDNA under the control of the promoter of s-SHIP. *Ship1* gene (SH2-containing Inositol 5'-Phosphatase-1) encodes a 145-kDa signaling protein with 5' phosphatase activity. From this *Ship1* gene, a second protein (~104-kDa) is encoded but lacking the amino-terminal SH2 domain compared to the SHIP1 protein and is expressed in embryonic stem cells and bone marrow cells enriched for the stem cell population [16,17]. This protein was termed s-SHIP, suggesting its potential for expression in stem cells. The SHIP1 protein is produced from a full-length mRNA, whereas, s-SHIP expression is produced from an internal promoter within intron 5/6 of the full-length *ship1* gene [18]. Stem cell-specific expression of s-SHIP promoter was determined by generating a transgenic mouse containing the 11.5kb s-SHIP promoter driving

the expression of GFP [18]. In these mice, s-SHIP promoter expression marks activated stem cells in the developing mammary tissue at puberty and during pregnancy [19]. Expression of the transgene was also observed in embryonic prostatic buds, suggesting that s-SHIP promoter expression may mark also prostate stem/progenitor cells [18]. To test this hypothesis, we used as a model the non-tumorigenic human prostate cell line RWPE-1 that was derived from normal human prostate epithelium immortalized by human papillomavirus 18 [20]. RWPE-1 cells and its derivatives contain stem, intermediate, and differentiated cell types and offer valuable models for studies of adult prostate stem cells [21-22].

In this report, we show that s-SHIP-GFP promoter reporter tracks subsets of RWPE-1 cells enriched in stem cell characteristics such as enhanced stem cell markers expression. In this subset population, higher expression of the long non-coding RNA (LncRNA) *H19* [23] was observed and further investigations strongly suggested that *H19* may play a role in prostate stemness through the expression of key pluripotency transcription factors, especially *Sox2*. Altogether, these data provide new insights into the genetic network controlling stem cell identity, uncovering a role for the long non-coding RNA *H19* as a potential stemness regulator.

Materials and Methods

Prostate and mammary cell lines and cell culture

RWPE-1 cells (a gift of Dr. BS Kundsen, Fred Hutchinson Cancer Research Center, Seattle, USA) were maintained in Keratinocyte Serum Free Medium (Gibco, Life Technologies) supplemented with 5 ng/ml epidermal growth factor (EGF, PeproTech, France), bovine pituitary extract (BPE, Gibco, Life Technologies) and Zell Shied (Minerva Biolabs, Biovalley, France). Normal human prostate epithelial cells (PrEC) were obtained from Lonza (Allendale, NJ) and cultured in PrEC basal media containing PrEGM SingleQuot Kit supplements and growth factors (Lonza). Human androgen-dependent (LNCaP), and androgen-independent (PC-3 and DU145) prostate cancer epithelial cells were obtained from American Type Culture Collection (ATCC, Manassas, VA), and were maintained in RPMI 1640 Medium (Gibco, Life Technologies) supplemented with 10% fetal bovine serum (FBS, Gibco, Life Technologies) and Zell Shied. The highly metastatic M12 subline (a gift of Dr. BS Kundsen) was cultured in RPMI 1640 medium supplemented with 10 ng/ml EGF, 0.1 μ M dexamethasone (Sigma Aldrich, France), 5 μ g/ml insulin, 5 μ g/ml transferrin, and 5 ng/ml selenium (ITS medium, Sigma) and Zell Shied. The estrogen-sensitive MCF7 and T47D and the estrogen-insensitive MDA-MB-231 human cancerous mammary epithelial cell lines were obtained from the ATCC and maintained routinely in RPMI 1640 medium containing 10% of FBS and Zell Shield. Normal mammary epithelial cells (hTERT, hMEC) were obtained from ATCC and maintained in MEGM

(Lonza) supplemented with gentamycin and 1% penicillin/streptomycin. All cells were maintained at 37°C, in a humidified incubator of 5% CO₂.

Plasmids and siRNA transfections

The 11.5-kb s-SHIP promoter GFP construct (a gift of the late Dr. LR Rohrschneider, FHCRC, WA) has been described previously [18]. RWPE-1 cells were grown to 70%-80% confluence and transfected with 1 µg of DNA using PEI/ExGen 500 (Euromedex, France) according to the manufacturer's instructions; twenty-four hours after transfection, the culture medium was replaced with fresh medium. Between five and seven days after transfection, green fluorescent protein (GFP)-positive RWPE-1 cells were sorted by using a FACSAria cell sorter (Becton Dickinson, BD), then grown for one week and sorted to a purity of ≥ 95%, and analyzed. Human long non coding RNA *H19* cDNA was amplified by standard PCR with Go Taq Hot Start DNA Polymerase (Promega) and cDNA obtained from MDA-MB-231 cells as the template. The primers were Forward, 5'-AGCAGGGTGAGGGAGGGGGT -3' and Reverse, 5'-GTAACAGTGTATTGATG -3'. The amplicon was cloned into pcDNA3.1(-) vector using *Not*I and *Bam*H1 sites. RWPE-1 cells were transfected with pcDNA3-H19 or empty pcDNA3 as a control, using PEI/ExGen 500, and transfected cells were allowed to recover for 48 hours prior to selection with G418 (0.5 mg/ml, Sigma) for four weeks and beyond this time when all the control (non-transfected) cells had expired. For siRNA, transfections were performed using DharmaFect according to the manufacturer's transfection protocol (Dharmacon, GE Healthcare) and total RNA was extracted 48h after transfection for analysis. The siRNA sequences are presented in Supplementary Table S1.

Antibodies and flow cytometry

All primary antibodies were from eBiosciences: anti-CD44 (APC) (cat. #17-0441), rat IgG2b κ isotype control (APC) (cat. #17-4031), anti-Human CD166 (ALCAM) PerCP-eFluor[®]710 (cat. #46-1668), mouse IgG1 κ isotype control PerCP-eFluor[®]710 (Cat #46-4714), anti-Human TROP2 (EGP-1) (cat. #14-6024), mouse IgG2α isotype control purified (cat#14-472), anti-CD49f APC (cat. #17-0495), anti-Human CD29 (APC) (cat. #17-0299), rat IgG2a κ isotype control (APC) (cat. #17-4321). APC-goat anti-mouse Ig multiple adsorption (cat. #550826) secondary antibody was from BD Pharmingen. For immunofluorescence staining, 2.5 x 10⁵ cells were incubated in the dark with primary antibody at recommended concentration for 30 min on ice, cells were washed with complete medium, and resuspended with secondary antibody at recommended concentration for 30 min if necessary, or directly analyzed on flow cytometer (Canto II, BD).

Résultats

RNA isolation and Quantitative RT-PCR

5 x 10⁵ cells were plated into a 100-mm-diameter dish, a cell density that allow the cells to continue exponential growth for an additional 24 hours. Total RNA was then isolated using the RNeasy Plus Extraction kit (Qiagen). Reverse transcription was performed using QuantiTect Reverse Transcription kit (Qiagen) on 1 µg of RNA, according to the manufacturer's protocol. Real time PCR was achieved on Stratagene Mx3005P (Agilent Technologies) using KAPA SYBR® Fast Universal qPCR kit (Kapa Biosystems, CliniSciences, France): 2 µl of DNase-treated cDNA (from a 50 µl RT reaction) was added to 10 µl Master Mix, 2 µl of oligo mix (1 µM). Experiments were performed in triplicate and the comparative threshold cycle method was used for the calculation of amplification fold. The expression level of each gene was normalized by the expression level of the "housekeeping" RPLP0 gene (Δ Ct method). Primer sequences are presented in Supplementary Table S1.

Growth studies

Cell proliferation was evaluated by plating 2 x 10⁴ cells per plate in 6-well plate in 2 ml of culture medium. After 24, 48, 72 and 96 hours, cells were isolated after trypsinisation using trypsin-versene (EDTA) solution and cell counting was done under the microscope using a Malassez counting chamber. For colony-forming unit (CFU) assay, single cells (50, 100, and 200 cells/ well) were plated in 6-well plates and cultured for 2 weeks in K-SFM medium. Cells were fixed with 4% paraformaldehyde made in phosphate buffered saline (PBS) for 10 min at room temperature, rinsed with PBS and stained with Giemsa solution (Sigma). The plates were gently washed with PBS and dried. Macroscopic colonies visible by naked eyes were enumerated. Colony formation efficiency was evaluated as follows: CFU (%) = colonies/input cells x 100.

Sphere formation assay

Cells (100, 250, 500 cells) were plated in 0.5 ml of PrEC basal media containing PrEGM SingleQuot Kit supplements and growth factors, in ultra-low attachment 24-well plates (Costar) and grown for two weeks. To not disturb sphere formation and to avoid sphere aggregation, fresh medium was gently added to each well after five and ten days and no medium was aspirated. Total number of spheres was counted under microscope and photographed. Sphere formation efficiency was evaluated as follows: sphere forming unit SFU (%) = spheres/input cells x 100.

Chemoresistance

RWPE-1 and RW-GFP cells (10⁵ cells per well in 6-well plates) were plated and allowed to attach for 24 hours; then they were treated with doxorubicin (D1515, Sigma) at 0.5 µg/ml for

48 hours and photographed. P53 activation was determined by western blot: cells were lysed in RIPA buffer containing protease and phosphatase inhibitors (protease inhibitor P8340; phosphatase inhibitor cocktail 2 P5726, Sigma) and proteins were quantified with BCA protein assay (Pierce). Proteins were reduced in NuPAGE LDS Sample buffer with NuPAGE Reducing Agent (Invitrogen, Life Technologies) at 70°C for 10 min. Proteins were separated on SDS-PAGE 4-12% (Invitrogen, Life Technologies) and transferred onto PolyVinylidene Fluoride (PVDF) membrane (Millipore). After saturation in PBS 0.2% of casein, membranes were incubated with primary antibodies over night at 4°C (p53 Do-1 sc-126, Santa Cruz; β -actin sc-47778, Santa Cruz). Membranes were washed with PBS 0.5% Tween for 30 min and incubated with secondary antibodies conjugated with Horse Radish Peroxidase (HRP) for 30 min at room temperature. Membranes are analysed with SuperSignal west Dura Chemiluminescence Substrate (Pierce). For arsenite-induced cytolethality, cells (10^5 cells per well in 6-well plates) were plated and allowed to attach for 24 hours; fresh medium was added containing sodium arsenite (NaAsO_2 , #35000, Sigma) at different concentrations (between 5 μM and 40 μM); after 48h, viable cells were photographed and counted with a Malassez hemocytometer (trypan blue method).

Statistical analysis

Data are expressed as mean values \pm standard error of the mean of at least 3 independent experiments. The statistical analysis was done by using Student's *t*-test and *p* value < 0.05 was considered significant.

Results

Isolation of RWPE-1 cells with s-SHIP promoter activity

Previous studies showed that the intron 5/6 promoter region of the *Ship1* gene regulates expression of a shorter isoform of the protein, called s-SHIP, in various stem/progenitor cell populations [18,19]. We tested whether high s-SHIP promoter expression could enrich stem-like cells within a human cell line. In this attempt, we used as a model the non-tumorigenic prostate cell line RWPE-1 in which s-SHIP transcript expression could be detected (Fig. 1A). Since s-SHIP promoter region shows high conservation between human and mouse (Supplementary Figure 1), we hypothesized that a DNA construct incorporating the 11.5kb s-SHIP promoter from the mouse upstream of the green fluorescent protein (GFP) substituted at the normal translation start site for s-SHIP [18] could be used to detect the live human cells with higher s-SHIP promoter activity. Among RWPE-1 cells which were transfected with 11.5kb-GFP, a small number of GFP-positive (GFP⁺) cells were observed (Fig. 1B) and then isolated by FACS. After two rounds of cell sorting, a population of GFP⁺ cells were obtained (Fig. 1C) and then tested for native s-SHIP transcript expression. We observed

Résultats

a four-fold increased expression as compared to the parental population (Fig. 1D), thus demonstrating that 11.5kb murine s-SHIP promoter region is functional in human prostate cells and allows enrichment in cells expressing endogenous s-SHIP mRNA. This population of RWPE-1 cells expressing higher level of s-SHIP transcript was named RW-GFP cells and was then analyzed for its stemness characteristics.

s-SHIP-GFP promoter reporter tracks subsets of RWPE-1 cells enriched in clonogenic properties and chemoresistance

When progenitor/stem cells are cultivated in defined culture conditions, they have the ability to grow in an anchorage-independent manner and to form suspension spheres [24,25]. We showed that cells with higher s-SHIP expression (RW-GFP cells) were able to form more spheres compared to the parental RWPE-1 cells when cultivated in defined PrEGM serum-free medium and onto ultra-low attachment plates (Fig. 2A). RW-GFP-derived spheres appeared to be smaller with a homogeneous distribution of size as compared with RWPE-1-derived spheres (Supplementary Figure 2). Similarly, the colony-forming efficiency of RW-GFP cells was significantly higher, compared to parental RWPE-1 cells (Fig. 2B). In contrast, RW-GFP cells grew slower than parental RWPE-1 cells (Fig. 2C).

Stemness is associated to chemoresistance [26]; both RWPE-1 and RW-GFP cells responded to chemotherapeutic drug doxorubicin by inducing p53 protein expression and cytolethality (Fig. 3A,B); interestingly, p53 induction was lower in RW-GFP cells together with better cell survival as compared to parental RWPE-1 cells (Fig. 3 A,B). This prompted us to investigate innate resistance to arsenite-induced cytolethality since this was a trait associated with stemness in RWPE-1 cells as previously described [22]. Following short-term (*i.e.* 48 hours) arsenite exposure (Fig. 3 C,D), clear difference in arsenite-induced cytotoxicity occurred between RW-GFP cells and the parental RWPE-1 cells at all concentrations tested, especially at 20 μ M with a 2-fold increased in cell survival for the RW-GFP cells as compared to parental RWPE-1 cells (24.3% \pm 5.4 versus 11.8% \pm 3,2).

Taken together, these results suggested that higher s-SHIP expression was associated to lower proliferation and higher clonogenicity, a property of stem-like cells. We then investigated if this phenotype was also associated to the expression of stemness markers.

RW-GFP cells expressed higher levels of stem cell markers and pluripotency transcription factors

RW-GFP cell stemness characteristics were further analyzed for cell surface markers expression and mRNA profile of stemness-related genes. Although CD44 expression level was already high in parental RWPE-1 cells, RW-GFP subpopulation shifted significantly to a higher

CD44 expression level (Fig. 4A and Supplementary Figure 3A). Next we examined whether cell surface expressions of CD166 and TROP2, two other prostate stem cell markers [9,10] were modified in RW-GFP cells. As expected, CD166 and TROP2 expression levels were also significantly increased as compared to parental RWPE-1 cells (Fig. 4A and Supplementary Figure 3A). Similarly, integrin $\alpha 6$ (CD49f) and $\beta 1$ (CD29) expression showed a homogeneous expression of both markers with a slight increased expression as compared to parental RWPE-1 cells (Supplementary Figure 3B).

Transcription regulators such as *octamer-binding transcription factor 4 (oct4)*, *sex determining region Y-box 2 (sox2)*, *nanog*, *kruppel-like factor 4 (klf4)*, and *c-myc* are associated with induction or maintenance of stemness in different tissues [27]. We observed a significant higher expression of *oct4* and *sox2*, but not *nanog*, *klf4* or *c-myc*, in cells with high s-SHIP promoter activity (RW-GFP) as compared to parental RWPE-1 cells (Fig. 4B). Examination of mRNA expression profile was extended to different stemness-related genes, like the ATP-binding cassette transporter G2 (*abcg2*), or genes implicated in signaling pathways regulating stem cells (sonic hedgehog (*shh*) and *notch1*), or in prostate development (*p63* transcription factor) [28]. We observed significant increased expressions in RW-GFP cells for all of them, *p63* expression being the most important (~18-fold). The long non coding (lnc) RNA *H19* is involved in prostate carcinogenesis [29] and plays a role in early development [23]. These results led us to examine the expression of the non coding RNA *H19* in RW-GFP cells. As shown in Fig. 4C, a marked increased expression (~5-fold) was observed in RW-GFP cells as compared to parental cells; this prompted us to further investigate the relationship between *H19* expression and stemness characteristics in RWPE-1 cells.

Expression of the long non coding RNA H19 favors stemness phenotype of RWPE-1 cells

To evaluate the role of *H19* in RWPE-1 stemness, we analyzed whether the expression level of *H19* correlated with stemness markers expression and spheres formation. First, we delivered siRNA to knock down *H19* expression in RWPE-1 cells. As shown in Figure 5A, two days after transfection of RWPE-1 cells with *H19* siRNA, *H19* RNA level was repressed up to 80% of negative control. Expression of *oct4*, *sox2*, *notch1* and *s-SHIP* were significantly reduced, with *sox2* showing the most dramatic reduction (Fig. 5A). Furthermore, colony-forming efficiency of these transfected cells was examined, and we observed a significant decrease of CFU potential for cells transfected *H19* siRNA, compared to GFP siRNA-transfected control cells (Fig. 5B). Next, we determined the impact of *H19* overexpression on stem cell characteristics in RWPE-1 cells. Cells were transfected and selected by G418 and *H19* RNA level was analyzed by qRT-PCR. *H19* RNA level was significantly increased ~6-fold in RW-pcDNA-H19 cells relative to that in RW-pcDNA3 control cells (Fig. 6A). Then, expression of

several stemness genes was examined. A significant higher expression of *oct4*, *sox2*, *klf4*, *c-myc* and *abcg2* but not *nanog* or *p63* was observed in RW-pcDNA-H19 cells as compared to RW-pcDNA3 control cells (Fig. 6A). Thus, RWPE-1 cells showed a global increase in stemness gene expression when *H19* gene is overexpressed, but with some difference with RWPE-1 cells expressing higher level of *s-SHIP* transcript (Fig. 4B). Since *H19* gene overexpression increased the expression of stemness factors, we next investigated whether lnc RNA *H19* influenced the clonogenicity and the proliferation of RWPE-1 cells. When RW-pcDNA-H19 and RW-pcDNA3 control cells were analyzed for their sphere formation capabilities in ultra-low attachment plates, we observed that control RW-pcDNA3 exhibited a lower sphere-forming potential than the parental RWPE-1 cells, suggesting that long-term culture in the presence of G418 affected their clonogenicity (Fig. 6B). Nevertheless, RW-pcDNA-H19 cells that have been cultivated in identical culture conditions exhibited a significant increase of SFU potential (Fig. 6B), suggesting that *H19* overexpression promotes a stem-like cell phenotype similar to the RW-GFP cells. However, at the difference of RW-GFP cells, when proliferation in 2D culture conditions was analyzed, RW-pcDNA-H19 grew faster than RW-pcDNA3 control cells (Fig. 6C).

Correlation between H19 and sox2 expressions in various prostate and breast cell lines

H19 loss- or gain-of-function experiments in RWPE-1 cells (Fig. 5 & 6) suggested a positive correlation between *H19* expression and *sox2*, *oct4* and *abcg2* stemness gene expression. In order to assess this correlation, we determined the level of expression of these four genes in another population of normal human prostate epithelial cells (PrEC) and in three different prostate cancer cell lines: two androgen-independent (PC3 and M12) and one androgen-dependent (LNCaP) cell lines. In the PrEC normal cell line, a high *H19* RNA expression was detected as compared to RWPE-1 cells (Fig. 7A). Similarly, PrEC cells exhibited a high increased expression for *sox2* (54-fold), *oct4* (19-fold) and *abcg2* (8.4-fold) as compared to RWPE-1 cells (Fig. 7A). In cancerous cells, the M12 cell line expressed the highest *H19*, *sox2* and *abcg2* expression levels (Fig. 7A). To extend these observations and further investigate a putative co-expression between *H19* and stemness genes in epithelial cells, we used a set of mammary cell populations with two non-malignant cell lines (hTERT and hMEC) and three mammary cancer cell lines (MDA-MB-231, MCF7, T47D). Again, a strong positive correlation was observed between *H19*, *sox2* and *abcg2* expressions and to lesser extent with *oct4* expression (Fig. 7B).

Discussion

Understanding of the prostate stem cell biology is essential since there is a growing body of evidence suggesting that benign prostate hyperplasia and prostate cancer may arise from the stem or stem-like cell compartments [30]. In this study, we took advantage of the s-SHIP promoter-driven reporter to detect and isolate s-SHIP-expressing human prostate cells. First, we demonstrated that s-SHIP promoter activity was well correlated with expression of different stemness markers and sphere-formation capabilities in human prostate RWPE-1 cell line. These results are in good agreement with studies performed in mice models and identifying s-SHIP as a stem cell-specific isoform of SHIP1, which is expressed in both pluripotent ES cells and adult tissue-specific multipotent cells such as hematopoietic stem cells [16,17]. Rohrschneider et al. confirmed that s-SHIP promoter, located within the intron 5/6 of the SHIP1 gene, promotes stem cell-specific transcription of the GFP reporter gene in a transgenic mouse model (Tg11.5kb-GFP) [18]. In these mice, GFP expression from the s-SHIP promoter marks activated mammary stem cells that specifically localize to the cap cell region of terminal end buds in developing mammary glands and alveolar units [19,31,32]. Regarding the prostate tissue, s-SHIP promoter-GFP reporter is expressed in a specific subset of cells within the embryonic prostate [18] and we have recently extended this result to prostate postnatal development and demonstrated that in developing prostatic buds of newborn mice, GFP-expressing cells correspond to a population of basal cells with stem cell characteristics (manuscript in preparation).

All of these experiments were previously performed in mouse tissues. In this work, we provide the first evidence that s-SHIP promoter expression could also be used for the enrichment of human stem-like cells. SIP-110, the human homolog of s-SHIP, is expressed in various human cell lines including hematopoietic [33], mammary and melanoma cell lines (our unpublished data), suggesting that the s-SHIP-GFP promoter reporter should be functional in these cells. More experiments are now needed to determine if our approach could be useful in other cell types, including cancer cell lines with the ultimate goal to isolate and characterize human cancer stem cells. The exact role of s-SHIP as a signaling molecule remains to be determined in these stem cell populations. Even though this protein lacks the SH2 domain as compared to SHIP1, s-SHIP contains cell signaling domains, including the 5'-inositol phosphatase domain [16], that might play a key role in controlling phosphatidylinositol-3-kinase (PI3K)/AKT activation. In this point of view, an interesting study has recently demonstrated the importance of the enzymatic activity of SHIP2, another 5'-inositol phosphatase, for breast cancer stem cells [34]. Altogether, these studies support the idea that s-SHIP promoter expression offers a valuable and unique marker of mammary and prostate stem cell populations that could lead to the identification of molecules involved in stem cell biology.

Thus, characterization of our stem-like enriched RWPE-1 cell population uncovered a potential role for lncRNA *H19* in stemness. The *H19* gene, located in human in 11p15.5 locus, is submitted to genomic imprinting, being expressed only from the maternal allele [35]. No protein associated to this transcript has been discovered and it has been identified initially as a riboregulator [36], and more recently as a miRNA precursor [37]. A status of oncofetal gene has been ascribed to *H19*: during embryogenesis, *H19* is highly expressed both in extraembryonic tissues and in the embryo; after birth, its expression is repressed to a basal activity that subsists in several tissues [38-41]. In cancer, *H19* gene is reexpressed and acts as an oncogene [42-44].

In recent years, lncRNAs have been emerging as important components of gene regulation that may play key roles in regulating quiescence, survival and self-renewal of pluripotent stem cells [45]. Few mechanisms of action of lncRNAs in stem cell biology have been investigated until now. Among them, *lncRNA-RoR* (Regulator Of Reprogramming or *lincRNA-ST8SIA3*) is a suppressor of p53 response [46] whereas p53 suppression promotes stem cells expansion by activation of self-renewing divisions, symmetric division and reprogramming of somatic/progenitor cells in stem cells [47,48]. *LncRNA-RoR* can also act as endogenous miRNA sponge (impairing miR-145 repression) to positively regulate *Oct4*, *Nanog*, and *Sox2* and promote embryonic stem cell self-renewal [49]. Similarly, *H19* interacts with p53 and impairs its signaling [50], and acts as miRNA sponge to antagonize *let-7* miRNAs [49]. In breast cancer stem cells, *Let-7* miRNAs down-regulate self-renewal and tumorigenicity [52], and in prostate cancer cells, loss of *Let-7* up-regulates EZH2 (a component of PRC2, polycomb repressive complex) and induces a stem cell signature [53]. Interestingly, *H19* by squelching *Let-7* miRNAs could enhance EZH2 expression and activity to promote epithelial-mesenchymal transition in bladder cancer [54]. In RWPE-1 cell line, *H19* could confer stemness phenotype by these multiple pathways. In addition, knockdown of *H19* in parthenogenetic embryonic stem cells (pES) promotes differentiation of pES to epidermis, showing that *H19* RNA contributes to stem cell integrity [55].

Finally, *H19* loss- or gain-of-function experiments showed coregulation of *H19*, *sox2*, *abcg2* and in a lesser extend *oct4*. Even if *H19* could act by modifying chromatin conformation at these loci, our results suggest a direct role of *H19* in the regulation of these genes (Fig. 5A & 6A). Conversely, *oct4* and *sox2* positively control *H19* expression by impairing the methylation of imprinting control region (ICR) and promoter [56], and a defect of *oct4* and *sox2* binding sites is associated with loss of expression of *H19* in Beckwith-Wiedemann syndrome [57] suggesting an amplification loop of *H19*/Sox2 regulation.

In conclusion, our study is the first to demonstrate that s-SHIP-GFP promoter reporter offers a unique marker for the enrichment of human stem-like cell populations and this strategy enabled us to unveil lncRNA *H19* as a potential factor for inducing and maintaining the biological nature of stem cells. Further analysis should be performed to determine the exact role and the mechanisms of action of *H19* RNA.

Acknowledgments

This study was supported by grants from the Centre National de la Recherche Scientifique (CNRS), the Région Nord-Pas de Calais (plan de lutte contre le cancer), the SIRIC ONCOLille (INCa-DGOS-Inserm 6041). R.P.B. was supported by the Philippe Foundation. E.A. was supported by INCA (PLBio 2010-180). H.B.L.R. and G.B. were supported by postdoctoral fellowships from the Région Nord-Pas de Calais and the Institut Pasteur de Lille. We thank the flow cytometry platform of the BioImaging Center of Lille (BICeL) and Lucie Deschamps for their supports, Cateline Guerardel for technical assistance, and Dr. Jérôme Vicogne for revision of the manuscript. Roland P. Bourette dedicates this paper to the memory of his mentor and friend, the late Dr. Larry R. Rohrschneider who was the initiator of the s-SHIP project.

Author Disclosure Statement

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

References

1. Jemal A, Bray F, Murray T, Samuels A, Ma J, Ferlay J, et al. (2011). Global cancer statistics. *CA Cancer J Clin* 61:69-90.
2. Goldstein AS, Stoyanova T, Witte ON. (2010). Primitive origins of prostate cancer: *in vivo* evidence for prostate-regenerating cells and prostate cancer-initiating cells. *Mol Oncol* 4:385-396.
3. Miki J, Rhim JS. (2008). Prostate cell cultures as *in vitro* models for the study of normal stem cells and cancer stem cells. *Prostate Cancer Prostatic Dis* 11:32-39.
4. Litvinov IV, Vander Griend AJ, Xu Y, Antony L, Dalrymple SL, Isaacs JT. (2006). Low-calcium serum-free defined medium selects for growth of normal prostatic epithelial stem cells. *Cancer Res* 66:8598-8607.
5. Ma Y, Liang D, Liu J, Axcróna K, Kvalheim G, Stokke T, Nesland JM, Suo Z. (2011). Prostate cancer cell lines under hypoxia exhibit greater stem-like properties. *PLoS One* 6:e29170.

6. Richardson GD, CN Robson, SH Lang, DE Neal, NJ Maitland and AT Collins. (2004). CD133, a novel marker for human prostatic epithelial stem cells. *J Cell Sci* 117(Pt16):3539-3545.
7. Guo C, H Liu, BH Zhang, RM Cadaneanu, AM Mayle and IP Garraway. (2012). Epcam, CD44, and CD49f distinguish sphere-forming human prostate basal cells from a subpopulation with predominant tubule initiation capability. *PloS One* 7:e34219.
8. Rajasekhar VK, L Studer, W Gerald, ND Socci and HI Scher. (2011). Tumor-initiating stem-like cells in human prostate cancer exhibit increased NF- κ B signalling. *Nat Commun* 2:162.
9. Jiao J, A Hindoyan, S Wang, LM Tran, AS Goldstein, D Lawson, D Chen, Y Li, C Guo, B Zhang, L Fazli, M Gleave, ON Witte, IP Garraway and H Wu. (2012). Identification of CD166 as a surface marker for enriching prostate stem / progenitor and cancer initiating cells. *PloS One* 7:e42564.
10. Goldstein AS, DA Lawson, D Cheng, W Sun, IP Garraway and ON Witte. (2008). Trop2 identifies a subpopulation of murine and human prostate basal cells with stem cell characteristics. *Proc Natl Acad Sci USA* 105:20882-20887.
11. van der Hoogen C, G van der Horst, H Cheung, JT Buijs, JM Lippitt, N Guzman-Ramirez, FC Hamdy, CL Eaton, GN Thalmann, MG Cecchini, RC Pelger and G van der Pluijm. (2010). High aldehyde dehydrogenase activity identifies tumor-initiating and metastasis-initiating cells in human prostate cancer. *Cancer Res* 70:5163-5173.
12. Foster BA, KJ Gangavarapu, G Mathew, G Azabdaftari, CD Morrison, A Miller and WJ Huss. (2013). Human prostate side population cells demonstrate stem cell properties in recombination with urogenital sinus mesenchyme. *PLoS One* 8:e55062.
13. Gangavarapu KJ, G Azabdaftari, CD Morrison, A Miller, BA Foster and WJ Huss. (2013). Aldehyde dehydrogenase and ATP binding cassette transporter G2 (ABCG2) functional assays isolate different populations of prostate stem cells where ABCG2 function selects for cells with increased stem cell activity. *Stem Cell Res Ther* 4:132.
14. Jeter CR, B Liu, X Liu, X Chen, C Liu, T Calhoun-Davis, J Repass, H Zaehres, JJ Shen and DG Tang. (2011). NANOG promotes cancer stem cell characteristics and prostate cancer resistance to androgen deprivation. *Oncogene* 30:3833-3845.
15. Liang S, M Furuhashi, R Nakane, S Nakazawa, H Goudarzi, J Hamada and H Iizasa. (2013). Isolation and characterization of human breast cancer cells with SOX2 promoter activity. *Biochem Biophys Res Commun* 437:205-211.

16. Tu Z, JM Ninos, Z Ma, JW Wang, MP Lemos, C Desponts, T Ghansah, JM Howson and WG Kerr. (2001). Embryonic and hematopoietic stem cells express a novel SH2-containing inositol 5'-phosphatase isoform that partners with the Grb2 adapter protein. *Blood* 98:2028-2038.
17. Desponts C, JM Ninos and WG Kerr. (2006). s-SHIP associates with receptor complexes essential for pluripotent stem cell growth and survival. *Stem Cells Dev* 15:641-646.
18. Rohrschneider LR, JM Custodio, TA Anderson, CP Miller CP and H Gu. (2005). The intron-5/6 promoter region of the *ship1* gene regulates expression in stem/progenitor cells of the mouse embryo. *Dev Biol* 283:503-521.
19. Bai L and LR Rohrschneider. (2010). s-SHIP promoter expression marks activated stem cells in developing mouse mammary tissue. *Gene Dev* 24:1882-1892.
20. Bello DB, MM Webber, HK Kleinman, DD Wartinger and JS Rhim. (1997). Androgen responsive adult human prostatic epithelial cell lines immortalised by human papillomavirus 18. *Carcinogenesis* 18:1215-1223.
21. Tokar EJ, BB Ancrile, GR Cunha and MM Webber. (2005). Stem/progenitor and intermediate cells and the origin of human prostate cancer. *Differentiation* 73:463-473.
22. Tokar EJ, W Qu, J Liu, W Liu, MM Webber, JM Phang and MP Waalkes. (2010). Arsenic-specific stem cell selection during malignant transformation. *J Natl Cancer Inst* 102:638-649.
23. Vennin C, F Dahmani, N Spruyt and E Adriaenssens. (2013). Role of non-coding RNA in cells: example of the H19/IGF2 locus. *Adv Biosci Biotechnol* 4:34-44.
24. Shi X, J Gipp and W Bushman. (2007). Anchorage-independent culture maintains prostate stem cells. *Dev Biol* 312:396-406.
25. Xin L, RU Lukacs, DA Lawson, D Cheng and ON Witte. (2007). Self-renewal and multilineage differentiation *in vitro* from murine prostate stem cells. *Stem cells* 25:2760-2769.
26. Ni J, P Cozzi, J Hao, W Duan, P Graham, J Kearsley and Y Li (2014). Cancer stem cells in prostate cancer chemoresistance. *Curr Cancer Drug Tar* 14:225-240.
27. Calloni R, EA Cordero, JA Henriques and D Bonatto. (2013). Reviewing and updating the major molecular markers for stem cells. *Stem Cells Dev* 22:1455-1476.

28. Pignon JC, C Grisanzio, Y Geng, J Song, RA Shivdasani and S Signoretti. (2013). P63-expressing cells are the stem cells of developing prostate, bladder, and colorectal epithelia. *Proc Natl Acad Sci USA* 110:8105-8110.
29. Berteaux N, S Lottin, E Adriaenssens, F Van Coppenolle, X Leroy, J Coll, T Dugimont and JJ Curgy. (2004). Hormonal regulation of *H19* gene expression in prostate epithelial cells. *J Endocrinol* 183:69-78.
30. Prajapati A, S Gupta, B Mistry and S Gupta. (2013). Prostate stem cells in the development of benign prostate hyperplasia and prostate cancer: emerging role and concepts. *Biomed Res Int* 2013:107954.
31. Huo Y and IG Macara. (2014). The Par3-like polarity protein Par3L is essential for mammary stem cell maintenance. *Nature Cell Biol* 16:526-534.
32. Kogata N, E Oliemuller, O Wansbury and BA Howard. (2014). Neuregulin-3 regulates epithelial progenitor cell positioning and specifies mammary phenotype. *Stem Cells Dev* 23:2758-2770.
33. Kavanaugh WM, DA Pot, SM Chin, M Deuter-Reinhart, AB Jefferson, FA Norris, FR Masiarz, LS Cousens, PW Majerus and LT Williams. (1996). Multiple forms of an inositol polyphosphate 5-phosphatase form signaling complexes with Shc and Grb2. *Curr Biol* 6:438-445.
34. Fu C-H, R-J Lin, J Yu, W-W Chang, G-S Liao, W-Y Chang, L-M Tseng, Y-F Tsai, J-C Yu and AL Yu. (2014). A novel oncogenic role of inositol phosphatase SHIP2 in ER-negative breast cancer stem cells: involvement of JNK/vimentin activation. *Stem Cells* 32:2048-2060.
35. Bartolomei MS, S Zemel and SM Tilghman. (1991). Parental imprinting of the mouse *H19* gene. *Nature* 351:153-155.
36. Brannan CI, EC Dees, RS Ingram and SM Tilghman. (1990). The product of the *H19* gene may function as an RNA. *Mol Cell Biol* 10:28-36.
37. Cai X and BR Cullen. (2007). The imprinted *H19* noncoding RNA is a primary microRNA precursor. *RNA* 13:313-316.
38. Liu J, AI Kahri, P Heikkilä, V Ilvesmäki, R Voutilainen. (1995). *H19* and insulin-like growth factor-II gene expression in adrenal tumors and cultured adrenal cells. *J Clin Endocrinol Metab* 80:492-496.

39. Ariel I, D Weinstein, R Voutilainen, T Schneider, O Lustig-Yariv, N de Groot and A Hochberg. (1997). Genomic imprinting and the endometrial cycle. The expression of the imprinted gene H19 in the human female reproductive organs. *Diagn Mol Pathol* 6:17-25.
40. Adriaenssens E, L Dumont, S Lottin, D Bolle, A Leprêtre, A Delobelle, F Bouali, T Dugimont, J Coll and JJ Curgy. (1998). H19 overexpression in breast adenocarcinoma stromal cells is associated with tumor values and steroid receptor status but independent of p53 and Ki-67 expression. *Am J Pathol* 153:1597-1607.
41. Adriaenssens E, S Lottin, T Dugimont, W Fauquette, J Coll, JP Dupouy, B Boilly and JJ Curgy. (1999). Steroid hormones modulate H19 gene expression in both mammary gland and uterus. *Oncogene* 18:4460-4473.
42. Lottin S, E Adriaenssens, T Dupressoir, N Berteaux, C Montpellier, J Coll, T Dugimont and JJ Curgy. (2002). Overexpression of an ectopic H19 gene enhances the tumorigenic properties of breast cancer cells. *Carcinogenesis* 23:1885-1895.
43. Berteaux N, S Lottin, D Monté, S Pinte, B Quatannens, J Coll, H Hondermarck, JJ Curgy, T Dugimont and E Adriaenssens. (2005). H19 mRNA-like noncoding RNA promotes breast cancer cell proliferation through positive control by E2F1. *J Biol Chem* 280:29625-29636.
44. Matouk IJ, E Raveh, R Abu-lail, S Mezan, M Gilon, E Gershtain, T Birman, J Gallula, T Schneider, M Barkali, C Richler, Y Fellig, V Sorin, A Hubert, A Hochberg and A Czerniak. (2014). Oncofetal H19 RNA promotes tumor metastasis. *Biochim Biophys Acta* 1843:1414-1426.
45. Ng JH and HH Ng. (2010). LincRNAs join the pluripotency alliance. *Nat Genet* 42:1035-1036.
46. Zhang A, N Zhou, J Huang, Q Liu, K Fukuda, D Ma, Z Lu, C Bai, K Watabe and YY Mo. (2013). The human long non-coding RNA-RoR is a p53 repressor in response to DNA damage. *Cell Res* 23:340-350.
47. Cicalese A, G Bonizzi, CE Pasi, M Faretta, S Ronzoni, B Giulini, C Brisken, S Minucci, PP Di Fiore and PG Pelicci. (2009). The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells. *Cell* 138:1083-1095.
48. Bonizzi G, A Cicalese, A Insinga and PG Pelicci. (2012). The emerging role of p53 in stem cells. *Trends Mol Med* 18:6-12.

49. Wang Y, Z Xu, J Jiang, C Xu, J Kang, L Xiao, M Wu, J Xiong, X Guo and H Liu. (2013). Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal. *Dev Cell* 25:69-80.
50. Yang F, J Bi, X Xue, L Zheng, K Zhi, J Hua and G Fang. (2012). Up-regulated long non-coding RNA H19 contributes to proliferation of gastric cancer cells. *FEBS J* 279:3159-3165.
51. Kallen AN, XB Zhou, J Xu, C Qiao, J Ma, L Yan, L Lu, C Liu, JS Yi, H Zhang, W Min, AM Bennett, RI Gregory, Y Ding and Y Huang. (2013). The imprinted H19 lncRNA antagonizes let-7 microRNAs. *Mol Cell* 52:101-112.
52. Yu F, H Yao, P Zhu, X Zhang, Q Pan, C Gong, Y Huang, X Hu, F Su, J Lieberman and E Song. (2007). let-7 regulates self renewal and tumorigenicity of breast cancer cells. *Cell* 131:1109-1123.
53. Kong D, E Heath, W Chen, ML Cher, I Powell, L Heilbrun, Y Li, S Ali, S Sethi, O Hassan, C Hwang, N Gupta, D Chitale, WA Sakr, M Menon and FH Sarkar. (2012). Loss of let-7 up-regulates EZH2 in prostate cancer consistent with the acquisition of cancer stem cell signatures that are attenuated by BR-DIM. *PLoS One* 7:e33729.
54. Luo M, Z Li, W Wang, Y Zeng, Z Liu and J Qiu. (2013). Long non-coding RNA H19 increases bladder cancer metastasis by associating with EZH2 and inhibiting E-cadherin expression. *Cancer Lett* 333:213-221.
55. Yin Y, H Wang, K Liu, F Wang, X Ye, M Liu, R Xiang, N Liu and L Liu. (2014). Knockdown of H19 Enhances Differentiation Capacity to Epidermis of Parthenogenetic Embryonic Stem Cells. *Curr Mol Med* 14:
56. Zimmerman DL, CS Boddy and CS Schoenherr. (2013). Oct4/Sox2 binding sites contribute to maintaining hypomethylation of the maternal igf2/h19 imprinting control region. *PLoS One*. 8:e81962.
57. Abi Habib W, S Azzi, F Brioude, V Steunou, N Thibaud, CD Neves, M Le Jule, S Chantot-Bastarud, B Keren, S Lyonnet, C Michot, M Rossi, L Pasquier, C Gicquel, S Rossignol, Y Le Bouc and I Netchine. (2014). Extensive investigation of the IGF2/H19 imprinting control region reveals novel OCT4/SOX2 binding site defects associated with specific methylation patterns in Beckwith-Wiedemann syndrome. *Hum Mol Genet*. pii: ddu290.

Figure Legends

FIG. 1. Isolation of RW-GFP cells, a subset of RWPE-1 cells with higher levels of *s-SHIP* promoter expression. (A) RT-PCR analysis of *s-SHIP* transcript in parental RWPE-1 cell line. (B) RWPE-1 cells were transfected with *s-SHIP*-GFP promoter reporter (11.5kb-GFP) and few GFP⁺ cells were obtained (arrow, left panel), then isolated by FACS and amplified in culture to create RW-GFP cells (right panels). (C) Flow cytometry analysis of GFP expression in RW-GFP cells (filled graph) after two round of cell sorting and cell amplification as compared to parental cells (empty graph). (D) qRT-PCR analysis of endogenous *s-SHIP* transcript expression in RW-GFP cells as compared to parental RWPE-1 cells. Data represent mean values \pm standard error of the mean of three independent experiments, p values was determined by Student's test *** p < 0.001.

FIG. 2. Clonogenicity and growth curve of RW-GFP cells. (A) When cultivated at low density in liquid PrEGM culture medium in ultra-low attachment 24-well plates in non adherent conditions, RW-GFP cells formed more spheres than parental RWPE-1 cells ; lower panels represent spheres formed after 15 days of culture. (B) When 50 to 200 cells/well were plated in 6-well plates in K-SFM medium, RW-GFP cells formed more colonies than parental RWPE-1 parental cells ; lower panels are pictures of 2-weeks-old colonies after fixation and Giemsa staining. (C) RW-GFP cells grew slower than parental RWPE-1 cells ; cell growth curves were obtained after plating 2×10^4 cells/well in 6-well plate and viable cell numbers were determined daily after trypsinisation. (A,B,C) Data represent mean values \pm standard error of the mean of three independent experiments, p values was determined by Student's test *** p < 0.001, ** p < 0.01, * p < 0.05.

FIG. 3. Chemoresistance of RW-GFP cells. (A,B) RWPE-1 and RW-GFP cells were untreated (- Dox) or treated (+ Dox) with doxorubicin (Dox, 0.5 μ g/ml) for 48 hours and (A) immunoblot analysis was then performed on whole cell lysate for p53 expression, using anti- α -actin antibodies as a loading control; (B) cells were photographed (bar = 100 μ M), (C, D) RWPE-1 and RW-GFP cells were untreated or treated with arsenite (5, 10, 20, 30, 40 μ M) for 48 hours and (C) viable cells were counted using the trypan blue dye exclusion method, and (D) cells were photographed (bar = 100 μ M); (C) Data represent mean values \pm standard error of the mean of five independent experiments, p values was determined by Student's test ** p < 0.01, * p < 0.05.

FIG. 4. RW-GFP cells exhibit higher expression level of stem cell markers and a significant increase of long non coding RNA *H19* expression. (A) Fold-enrichment over the parental RWPE-1 cells for CD44, CD166, and Trop2 cell surface marker expression analyzed by flow cytometry; data are expressed as mean fluorescence intensity (MFI) of RW-

GFP cells (black bars) as compared to parental RWPE-1 cells (open bars). **(B)** Expression of stemness genes and **(C)** long non coding RNA *H19* was analyzed by qRT-PCR using total RNA from parental RWPE-1 cells (open bars) or RW-GFP cells (closed bars). Levels were normalized to those of *RPLP0* internal control. Graphs show fold-enrichment in RW-GFP cells over the parental RWPE-1 cells for each gene. **(A, B, C)** Data represent mean values \pm standard error of the mean of three independent experiments, p values was determined by Student's test *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

FIG. 5. Inhibition of long non coding RNA *H19* expression decreases stemness gene expression and clonogenicity. **(A)** Knockdown of *H19* by siRNA reduced expression of several stemness genes ; expression was analyzed by qRT-PCR using total RNA from control (GFP) siRNA-transfected RWPE-1 cells (open bars) or from *H19* siRNA-transfected RWPE-1 cells (closed bars) and levels were normalized to those of *RPLP0* internal control. Graphs show fold-decrease over control siRNA-transfected RWPE-1 cells for each gene. **(B)** When 50 to 200 cells/well were plated in 6-well plates in K-SFM medium, *H19* siRNA-transfected RWPE-1 cells formed fewer colonies than control siRNA-transfected RWPE-1 cells ; lower panels are pictures of 2-weeks-old colonies after fixation and Giemsa staining. **(A,B)** Data represent mean values \pm standard error of the mean of three independent experiments, p values was determined by Student's test *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

FIG. 6. Overexpression of long non coding RNA *H19* favors stemness gene expression and clonogenicity. RWPE-1 cells were transfected either with pcDNA-H19 or pcDNA-control and selected with G418 to obtain stable neo^r cell lines. **(A)** Overexpression of *H19* in RWPE-1 cells increased expression of several stemness genes ; expression was analyzed by qRT-PCR using total RNA from empty vector-expressing G418-resistant RWPE-1 cells (open bars) or from *H19*-expressing G418-resistant RWPE-1 cells (closed bars) and levels were normalized to those of *RPLP0* internal control. Graphs show fold-change over empty vector-expressing G418-resistant RWPE-1 cells for each gene. **(B)** Cells were cultivated at low density in PrEGM culture medium in ultra-low attachment 24-well plates in liquid cultures and RW-pcDNA-H19 formed more spheres than RW-pcDNA3 control cells ; lower panels represent spheres formed after 15 days of culture. **(C)** RW-pcDNA-H19 cells grew faster than RW-pcDNA3 control cells; cell growth curves were obtained after plating 2×10^4 cells/well in 6-well plate and viable cell numbers were determined daily after trypsinisation. **(A,B,C)** Data represent mean values \pm standard error of the mean of three independent experiments, p values was determined by Student's test *** $p < 0.001$, ** $p < 0.01$.

FIG. 7. Long non coding RNA *H19* expression is associated with *sox2* expression in various prostate and mammary cell lines. *H19*, *oct4*, *sox2* and *abcg2* expression was analyzed by qRT-PCR using total RNA from (A) prostate non-malignant RWPE-1 and prostate epithelial cells (PrEC) and from PC3, M12, LNCaP prostate cancer cell lines, and (B) mammary hTERT and hMEC non-malignant immortalized cell lines and breast tumor cell lines (MDA-MB-231, MCF7, T47D). Levels were normalized to those of RPLP0 internal control. Graphs show fold-enrichment over (A) RWPE-1 cells and (B) hTERT cells for each gene. (A,B) Data represent mean values \pm standard error of the mean of three independent experiments, p values was determined by Student's test *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Figure S1. Human and Murine intron 5 s-SHIP promoter regions.

EMBOSS Needle alignment of the human and murine intron 5 of *Ship1* gene ; p53 motif is shown ; Inr : initiator element.

Figure S2. Spheres derived from RWPE-1 and RW-GFP cells.

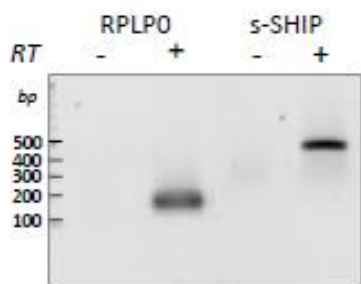
Typical floating spheres formed were photographed (bright field images) after 2 weeks in cultures.

Figure S3. Flow cytometry analysis of stem cell markers expression.

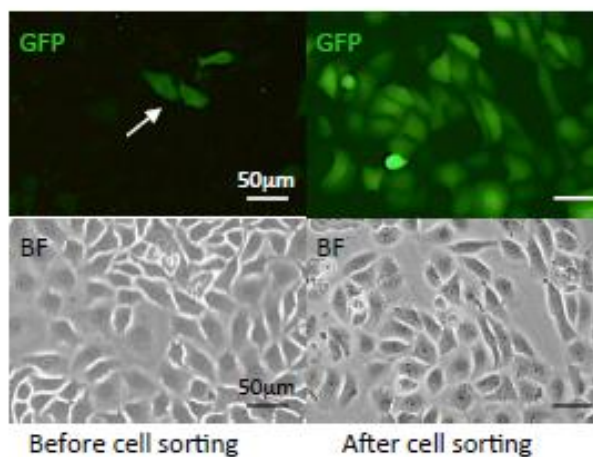
Representative dot plots of parental RWPE-1 and RW-GFP cells labelled with each antibody. (A) CD44, CD166, TROP2 and (B) integrin $\alpha 6$ (CD49f) and integrin $\alpha 1$ (CD29); setting were done using unstained or isotype-labelled cells which gave similar results; FSC : forward side scatter.

Table S1. Primers and siRNA used in this study.

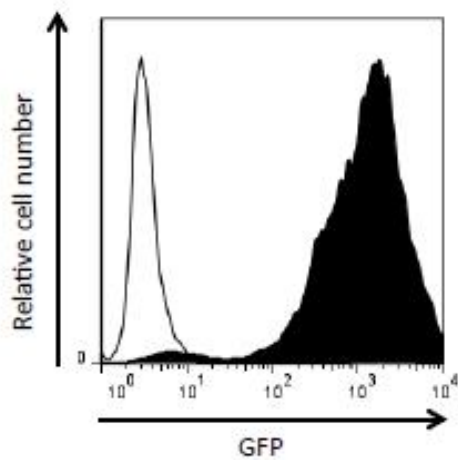
A



B



C



D

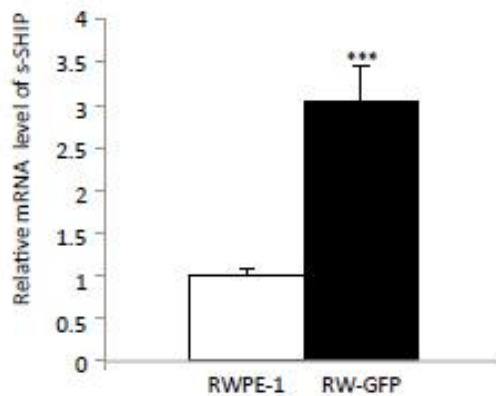


Figure 1.

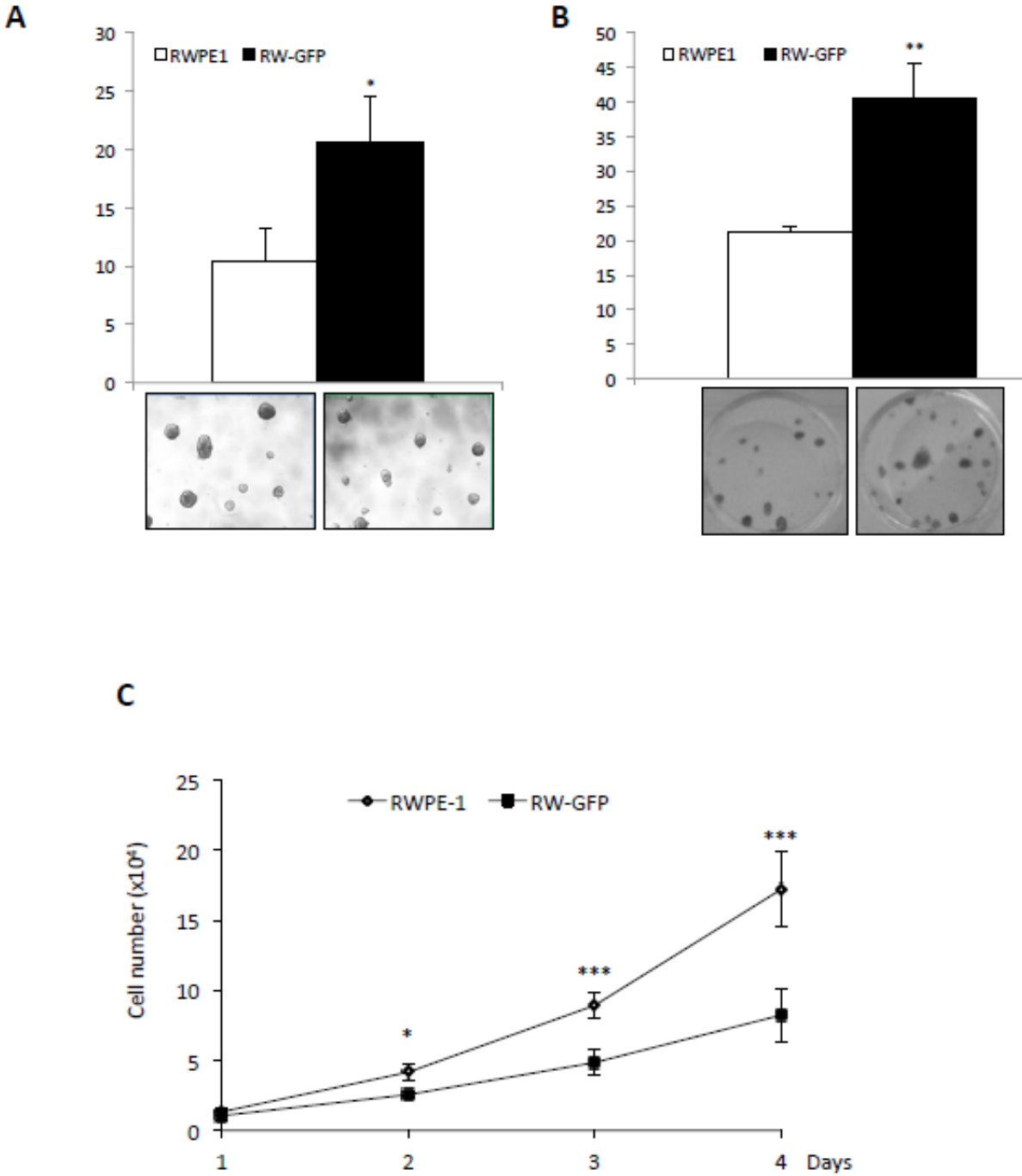


Figure 2.

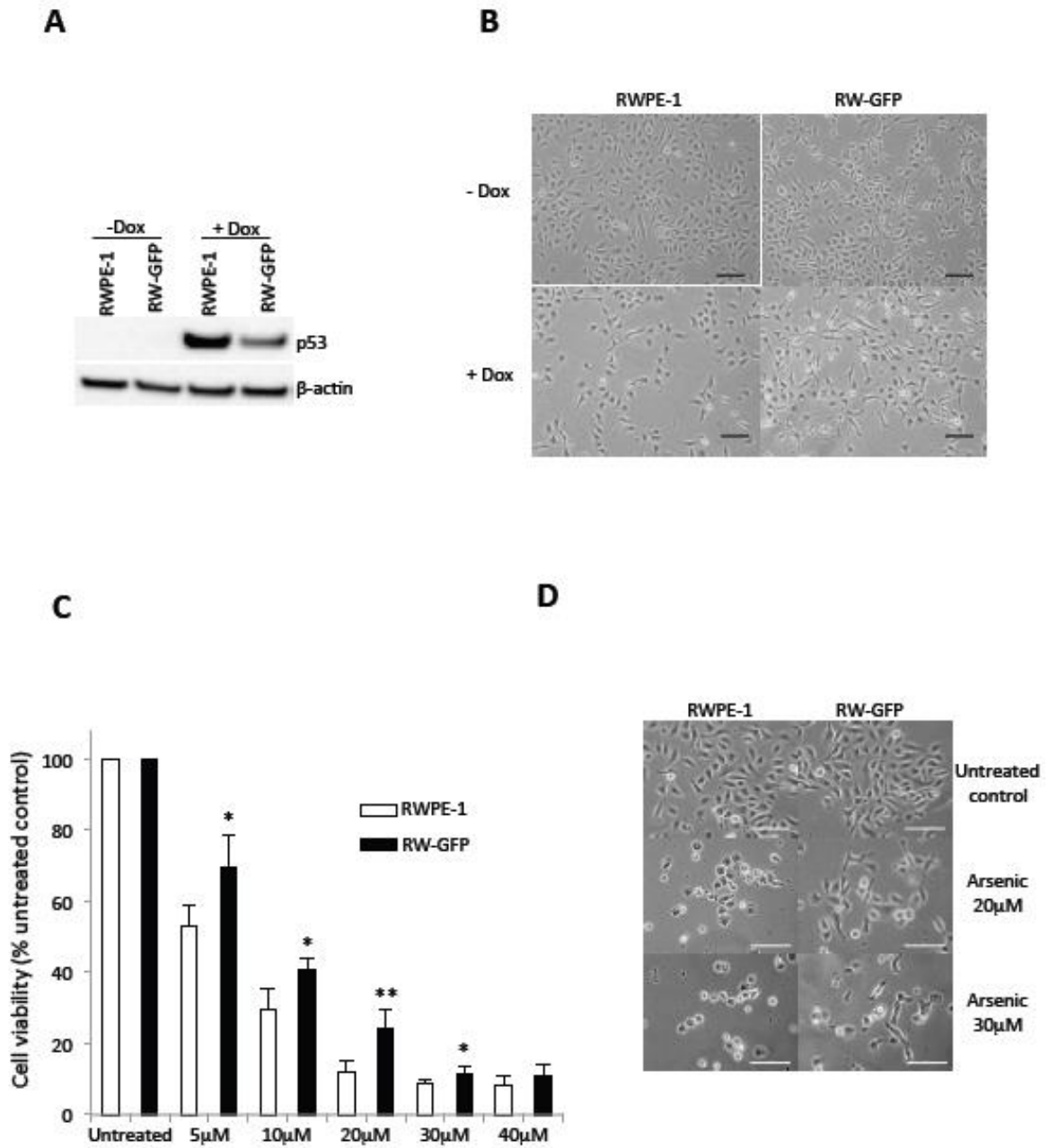
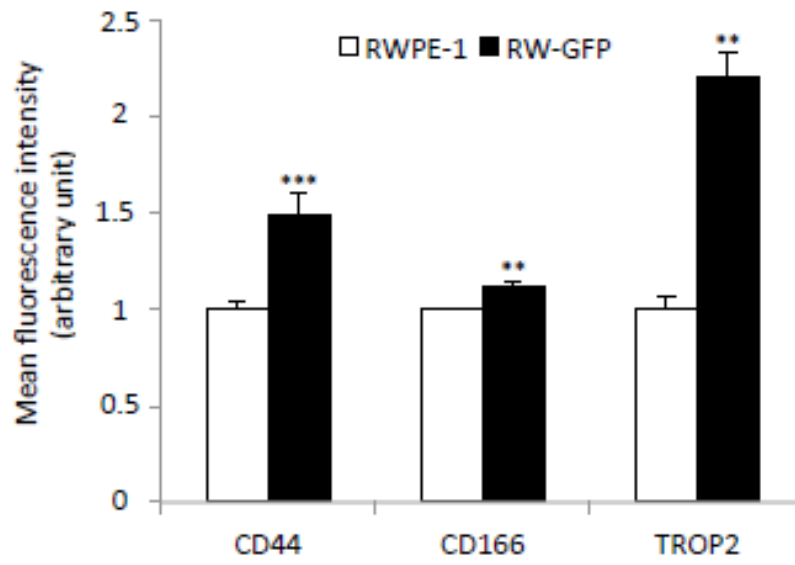
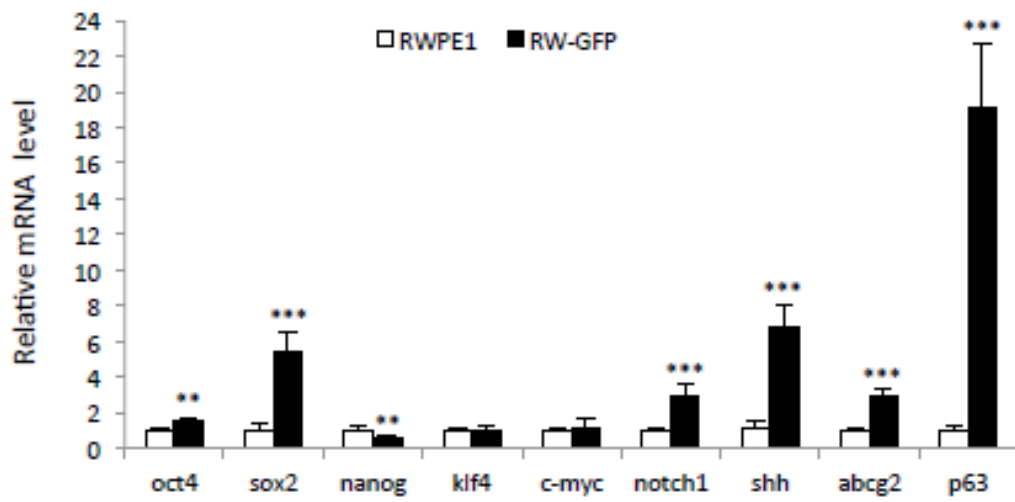


Figure 3.

A



B



C

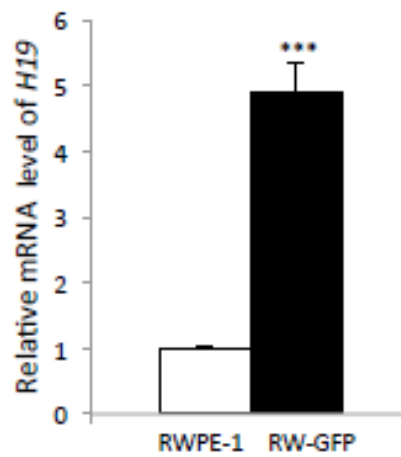


Figure 4.

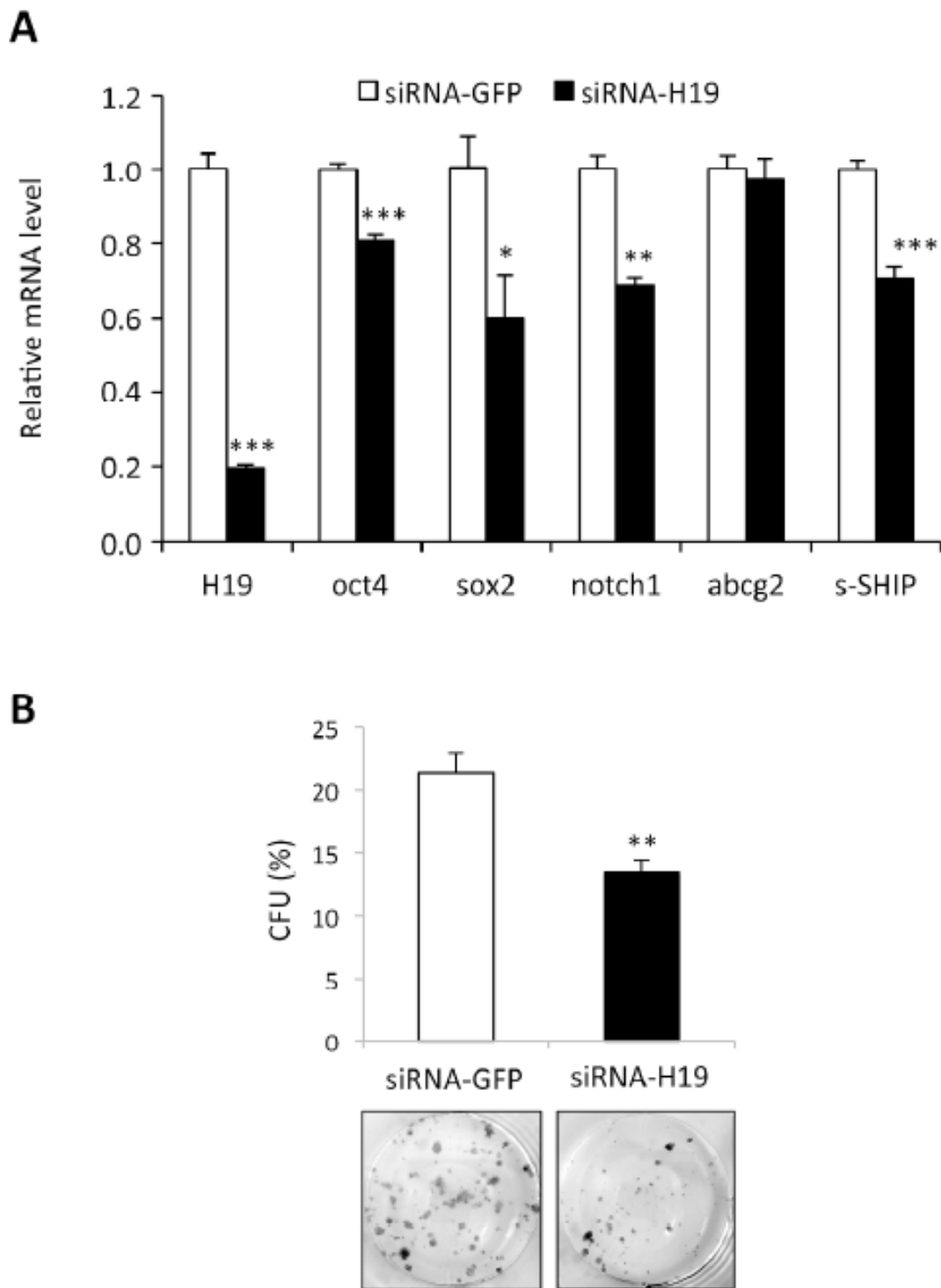
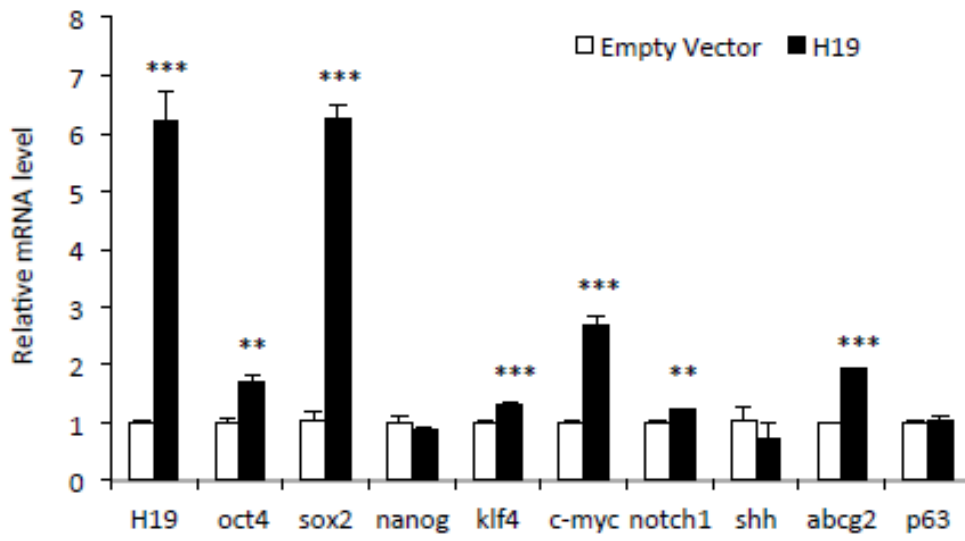
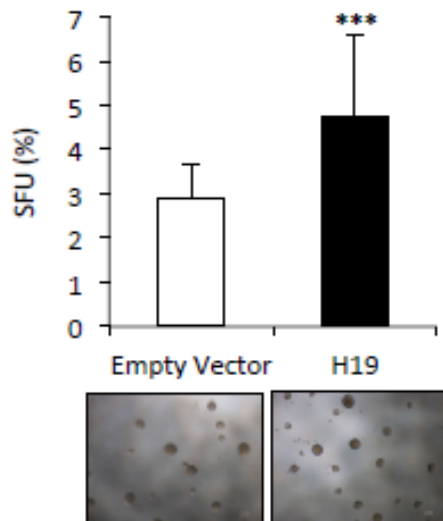


Figure 5.

A



B



C

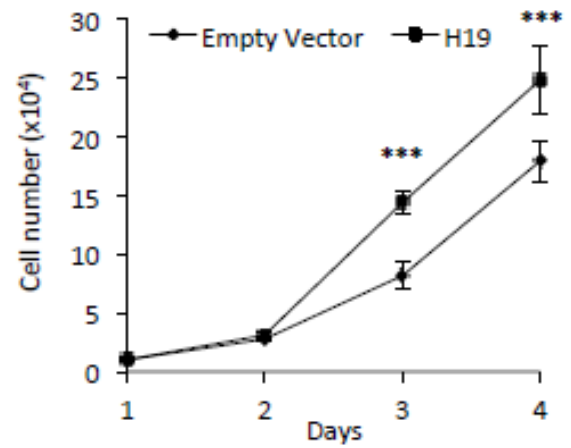


Figure 6.

A

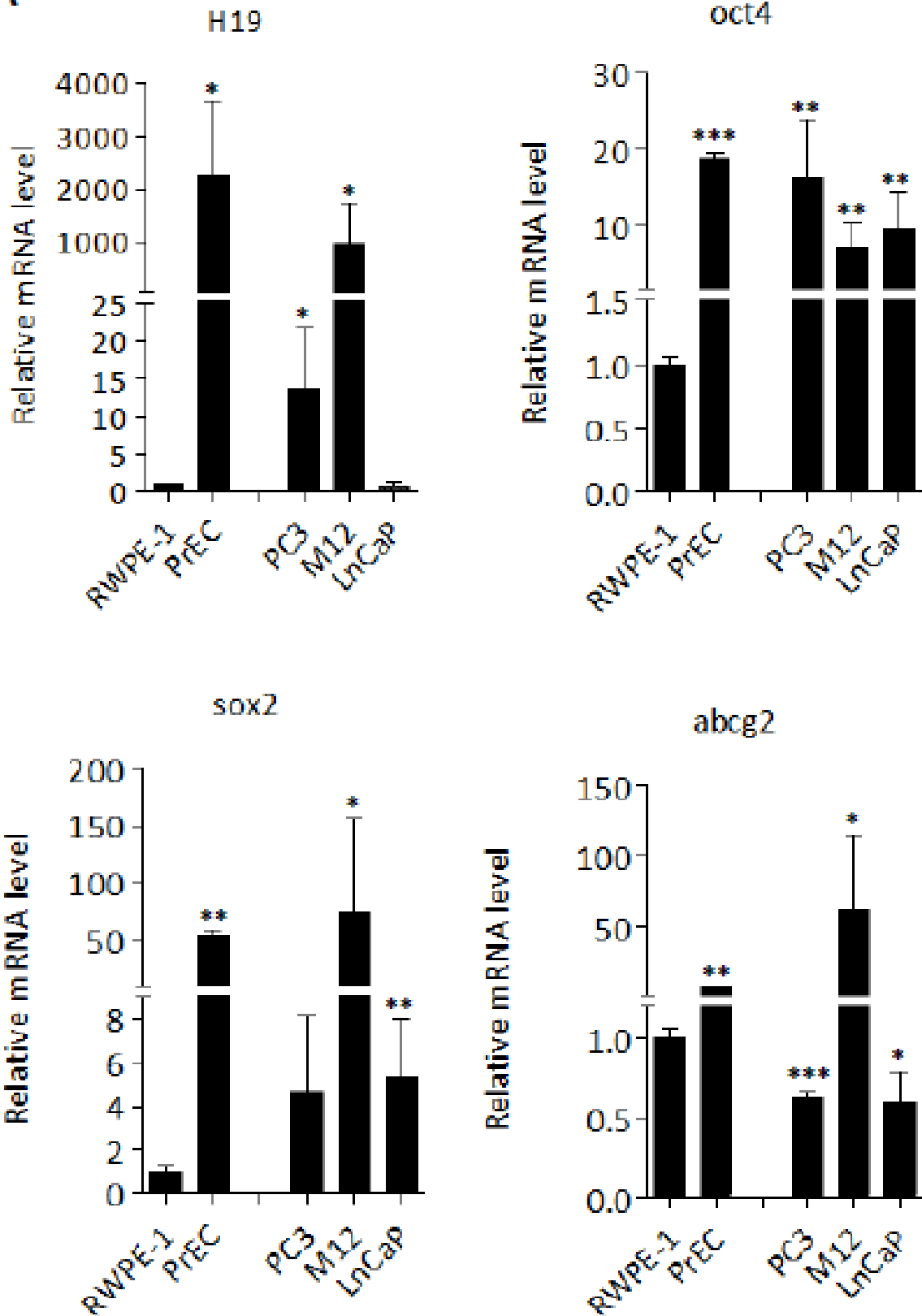


Figure7.

B

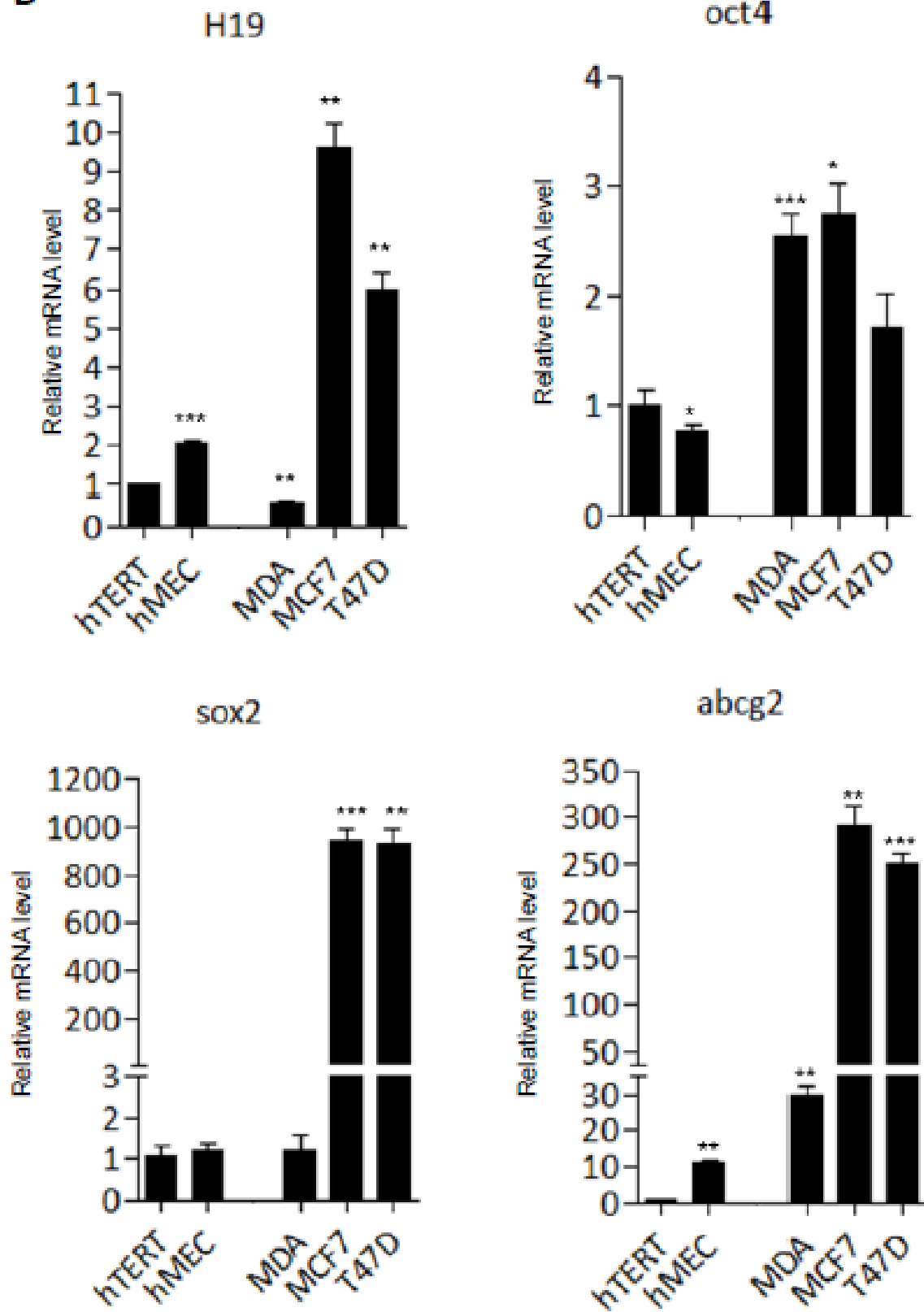


Figure 7.

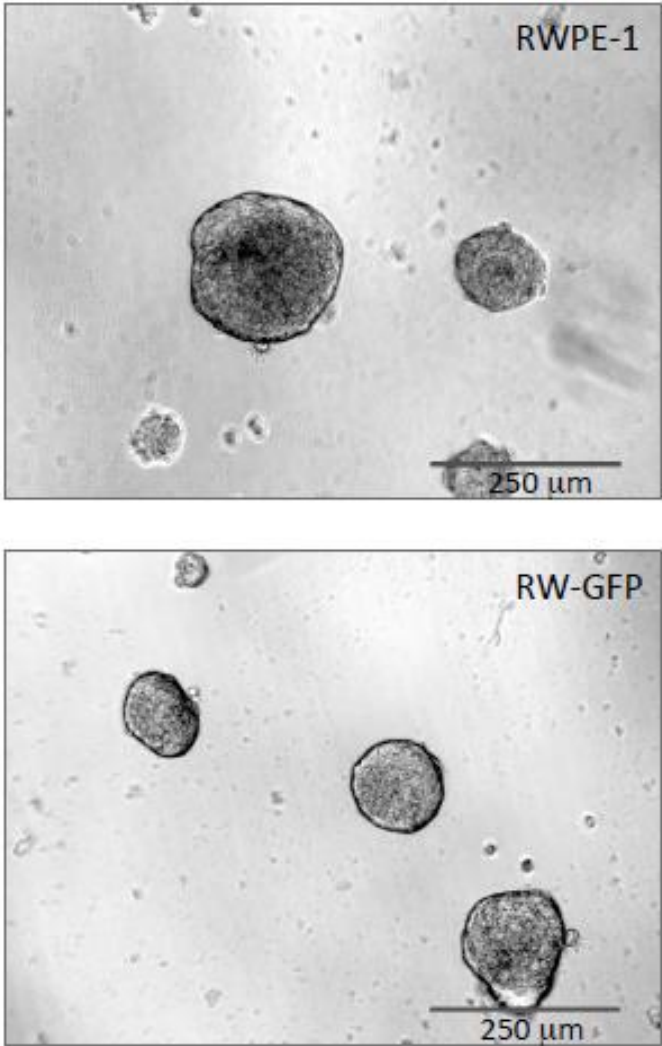


Figure S2.

A

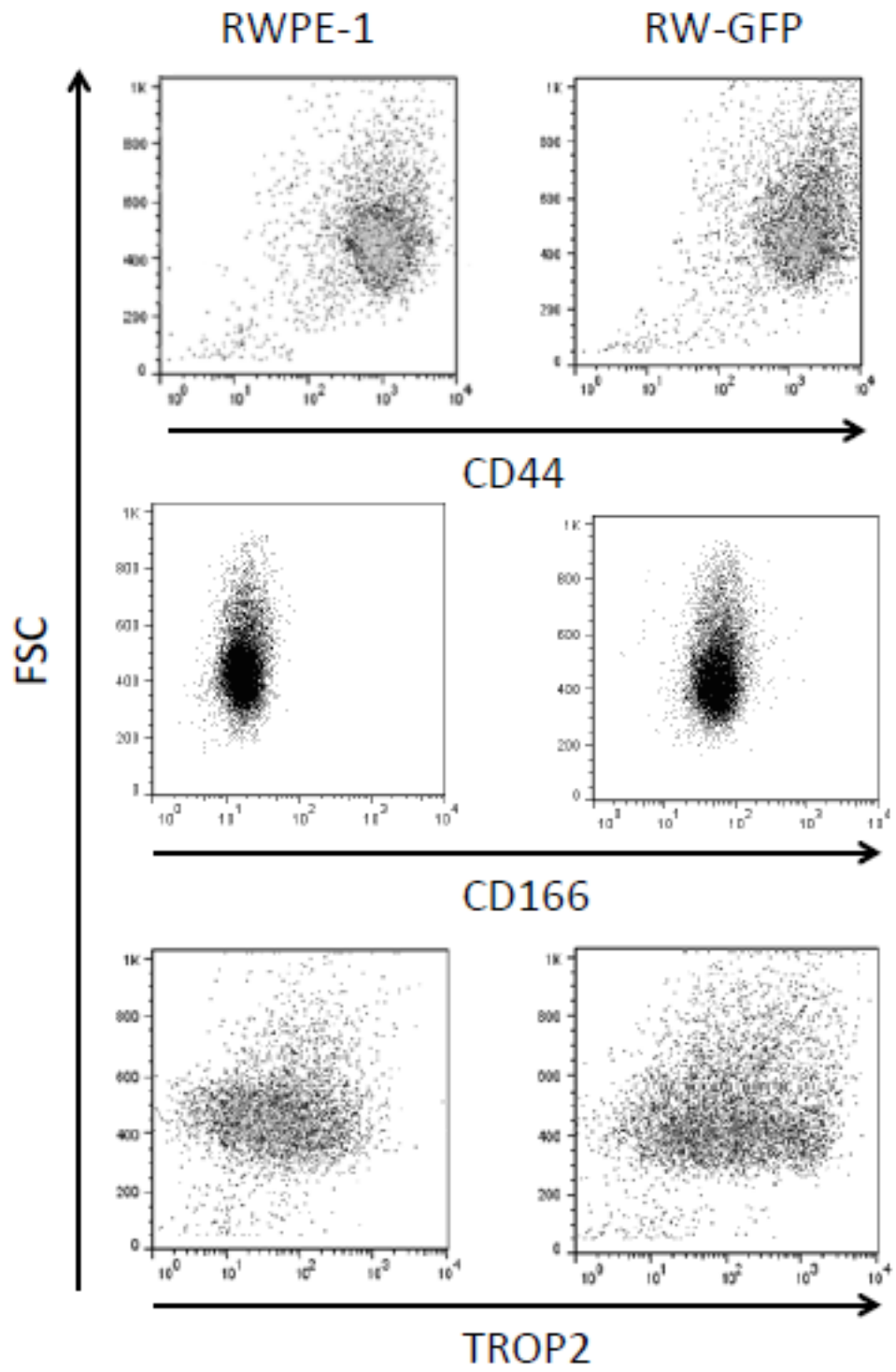


Figure S3.

B

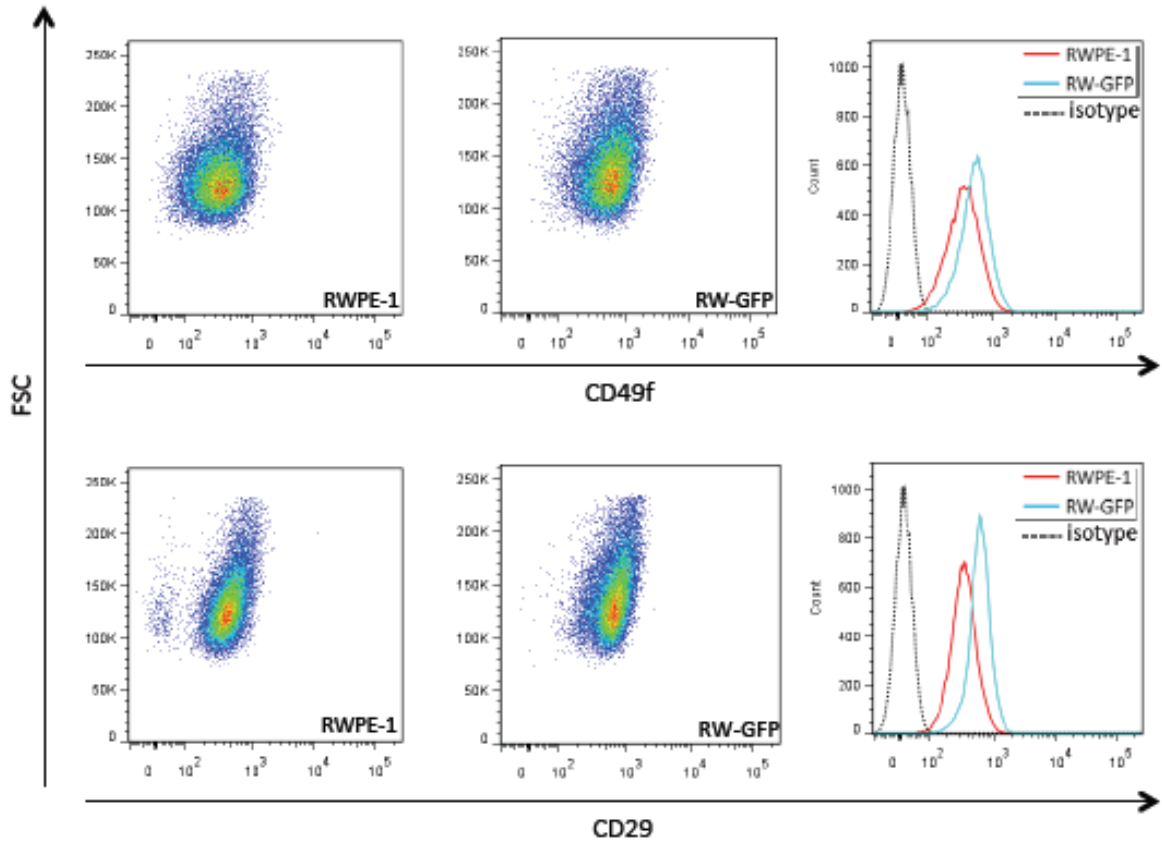


Figure S3.

name	5' to 3'
	RT-qPCR
abcg2 as	GCTGCAAAGCCGTAATCCATATC
abcg2 s	ACAACCATTGCATCTTGGCTGTC
c-myc as	TCTTCTTGTTCCCTCCTCAGAGTCG
c-myc s	TTCTCTCCGTCCTCGGATTCTCTG
H19 as	TGGCCATGAAGATGGAGTCG
H19 s	TACAACCACTGCACTACCTG
klf4 as	GGTGTGCCTTGAGATGGGAA
klf4 s	TACCAAGAGCTCATGCCACC
nanog as	CAGGTCTTCACCTGTTTGTAGCTGAG
nanog s	CCTCCTCCATGGATCTGCTTATTCA
notch as	GTA CTGACCTGTCCACTCT
notch s	CTGTGAGGAAAATATCGAC
oct3/4 as	GTGAAGTGAGGGCTCCATA
oct3/4 s	GAAGGATGTGGTCCGAGTGT
P63 as	AACATATAGGCTGTTCCAGA
P63 s	GAAAGTTATTACCGATCCAC
RLP0 as	GATGACCAGCCCAAAGGAGA
RPLP0 s	GTGATGTGCAGCTGATCAAGACT
shh as	TCAACTTGTCTTACACCTCTGA
shh s	AAGGAACTACCCCCAATTACAA
sox2 as	CGGGGCCGGTATTTATAATC
sox2 s	AACCCCAAGATGCACA ACTC
S-ship as	TCAACGTCGACTTTGAGCTGC
S-ship s	CTTGATGTTACCTTGTCCTCTGC
	siRNA
siRNA H19 1 as	GGCCUUCUGAACACCUUA
siRNA H19 1 s	UAAGGUGUUCAGGAAGGCC
siRNA H19 2 as	GGCUCUGGAAGGUGAAGCU
siRNA H19 2 s	AGCUUCACCUUCCAGAGCC
siRNA GFP as	AAGAUGAACUUCAGGGUCAGC
siRNA GFP s	GCUGACCCUGAAGUUCAUCTT

Table S1.

Discussion

Dans cette étude, nous avons isolé une population cellulaire exprimant la GFP sous le contrôle du promoteur de *s-SHP* et caractérisé cette population cellulaire. Nous montrons que les cellules exprimant le gène *s-SHIP* possèdent des caractéristiques de cellules souches. En effet, une augmentation de l'expression de marqueurs de surface (CD44, CD166, TROP2), de facteurs de transcription spécifique des cellules souches (Sox2, Oct4, Notch, shh, abcg2) a été observée dans cette population cellulaire. Nous montrons également une augmentation de leur capacité à former des sphères ainsi que leur résistance aux drogues, notamment l'arsenic.

Nous avons mis en évidence que le gène *H19* est surexprimé dans la population de cellules souches prostatiques ainsi isolé. Par invalidation ou surexpression de ce gène, nous montrons qu'il est impliqué dans le maintien des cellules souches prostatiques. Dans une étude précédente, nous avons déjà mis en évidence l'implication de l'ARN *H19* dans le développement de la prostate normale et cancéreuse (Berteaux et al., 2004).

Par ailleurs, une corrélation étroite entre l'expression du gène *H19* et du gène *Sox2* a été mise en évidence. En effet, une forte expression d'*H19* est associée à une forte expression de *Sox2*. De plus, *Sox2* régule positivement l'expression d'*H19* en se fixant sur le centre de l'empreinte sur l'allèle maternel (Zimmerman et al., 2013). Cette régulation d'expression a également été retrouvée dans les cellules cancéreuses mammaires (figure 7 de l'article).

Nous avons ainsi mis en évidence l'implication de l'ARN H19 dans la formation et le maintien des cellules souches prostatiques normales. H19 étant principalement impliqué dans la tumorigenèse, il serait intéressant de déterminer s'il est également impliqué dans les cellules souches cancéreuses. Des analyses complémentaires sont nécessaires pour identifier le rôle exact et le(s) mécanisme(s) d'action de l'ARN H19 dans les cellules souches normales mais aussi dans les cellules souches cancéreuses.

IV. L'ARN 91H régule le locus H19/IGF2 et est impliqué dans la tumorigenèse.

Introduction

Le cancer du sein est la première cause de mortalité par cancer chez la Femme dans le monde. Malgré de nombreux progrès sur le diagnostic et les traitements, l'identification de nouvelles cibles thérapeutiques est plus que nécessaire. Comme décrit dans le chapitre III, les longs ARN non codants sont impliqués dans divers processus cellulaires grâce à l'ensemble de leur mode d'action.

Au locus *H19/IGF2*, notre équipe a identifié un nouveau long ARN non codant antisens à *H19*, *91H*, exprimé à partir de l'allèle maternel et conservé entre les espèces (Berteaux et al., 2008 ; Tran et al., 2012). De nombreux loci soumis à l'empreinte génomique parentale sont régulés par ce type d'ARN. Nous pouvons citer par exemple les loci *Igf2r/Air* ou encore *Kcnq1/Kcnq1ot1* dont l'empreinte génomique est régulée par les ARN *Air* et *Kcnq1ot1*, respectivement. Quelques études ont déjà décrit la fonction de cet ARN dans certains cancers (colorectal et gastrique) (Gao et al., 2015 ; Deng et al., 2014). Cependant, son rôle dans les cancers n'est pas clairement établi puisqu'il semblerait qu'il soit un suppresseur de tumeur dans les cancers colorectaux et un oncogène dans les tumeurs gastriques (Gao et al., 2015 ; Deng et al., 2014). De plus, la fonction de l'ARN *91H* au locus *H19/IGF2* n'a pas encore été identifiée chez l'Homme.

Dans cette étude, par invalidation de *91H* à l'aide d'ARN interférant, nous définissons le rôle de l'ARN non codant *91H* au cours de la tumorigenèse de la glande mammaire. Nous montrons également que *91H* est impliqué dans le maintien de l'empreinte génomique au locus *H19/IGF2*.

Article 3.

Constance Vennin, Nathalie Spruyt, Yves-Marie Robin, Xuefen Le Bourhis, Eric Adriaenssens. The long non-coding RNA *91H* increases aggressive phenotype of breast cancer cells and up-regulates *H19/IGF2* expression through epigenetic modifications. (En préparation)

The long non-coding RNA *91H* increases aggressive phenotype of breast cancer cells and up-regulates *H19/IGF2* expression through epigenetic modifications

Constance Vennin^{1,2}, Nathalie Spruyt³, Yves-Marie Robin⁴, Xuefen Le Bourhis^{1,2}, Eric Adriaenssens^{1,2*}

¹ INSERM U908, 59655 Villeneuve d'Ascq, France.

² University of Lille, 59655 Villeneuve d'Ascq, France.

³ CNRS UMR 8161, 59021 Lille, France.

⁴ Centre Oscar Lambret, 59000 Lille, France

* Correspondance: INSERM U908, Cell Plasticity and Cancer, Bâtiment SN3, University of Lille, 59655 Villeneuve d'Ascq, France. eric.adriaenssens@univ-lille1.fr

Tel: 00 33 (0)3 20 43 64 06

Fax: 00 33 (0)3 20 43 40 38

<http://signaling-reprogramming-csc.univ-lille1.fr>

Key Words: long non-coding RNA, *91H*, breast cancer

ABSTRACT

Numerous genomic imprinting loci are regulated by long non-coding RNA (lncRNA). We have previously identified a new lncRNA at the *H19/IGF2* locus transcribed in *H19* antisense orientation and named *91H*. This RNA is conserved between mammals. In mice, *91H* regulates positively *IGF2* expression from a novel promoter. However, in human the function of *91H* at the *H19/IGF2* locus remains undeciphered. Recently, several studies observed a contradictory function of *91H* in colorectal and gastric cancer.

Here, by using *91H*-knockdown cells, we demonstrated that *91H* regulates positively *H19* and *IGF2* expression. Our results suggest that *91H* prevents methylation on the maternal allele.

In addition, we determine the function of *91H* in breast cancer cells. We demonstrate that *91H* have oncogenic properties. Moreover, in human breast tumor, we observed a positive regulation of *91H*, *H19* and *IGF2*. So, we have identified the mechanism of *91H* at the *H19/IGF2* locus in human cells and we decipher the function of *91H* breast cancer cells. As *91H* mechanism increase expression of oncogenic gene, this new mechanism could be targeted in cancer therapies.

INTRODUCTION

Breast cancer is a common cause of cancer death in female in the world with 1.671 million new cancer cases and 522 thousand cancer deaths in 2012 worldwide. Despite numerous progresses in the past decade in diagnosis, characterization of tumors or treatment, the overall 5-year survival for several subsets of breast cancer remains unsatisfactory. Identification of novel molecular markers or actors of tumorigenesis may improve the patient care and may open up new opportunities for innovative therapeutic strategies.

The long non coding RNAs (lncRNAs) are essential in cell biology and dysfunction could be playing a critical role in cancer development or progression. Even if, numerous data indicate that lncRNAs are involved in cellular processes as cell proliferation, apoptosis, differentiation, pluripotency, their mechanisms of action remain unclear.

At the *H19/IGF2* locus, we have previously identified a long non-coding RNA antisense to *H19*, transcript from the maternal allele and conserved among mammalian, *91H* (Berteaux et al., 2008; Tran et al., 2012). We demonstrate that *91H* regulates IGF2 expression. Numerous genomic imprinting regions are regulated by long non-coding RNA antisense. For example, lncRNA *Kcnq1ot1*, *Air*, regulate genomic imprinting of their locus, respectively *Kcnq1/Kcnq1ot1*, *Igf2r/Air*, by epigenetic modifications (Sanli & Feil, 2015).

The *H19/IGF2* locus, located on the chromosome 11p15.5 and submitted to genomic imprinting, contains two major genes *H19* and *IGF2*. Misregulations of their expression promote cancer progression in various tissues. The *H19* gene expresses from the maternal allele, encodes for a long non-coding RNA spliced, polyadenylated and capped.

The *H19* gene is expressed during embryonic development and repressed after birth except in few tissues like mammary gland, adrenal gland and uterus (Adriaenssens, 1998; Adriaenssens, 1999; Liu, 1995; Ariel, 1997). In breast cancer, we have shown that *H19* is frequently overexpressed (70%) (Adriaenssens 1998). Moreover, we have determined that *H19* promotes cell cycle progression, tumorigenesis and invasion of breast cancer cells (Lottin et al., 2002; Berteaux et al., 2005). In addition, several studies show that *H19* acts as an oncogene in other cancers: gastric, lung, esophageal, and bladder (Tsang, 2010; Kondo, 1995; Hibi, 1996; Ariel, 1995). Collectively, these data indicate that, in breast cancer, *H19* acts as an oncogene.

The *IGF2* gene, expressed from the paternal allele, encodes for a protein of 7.5kDa involved in embryonic growth. The *IGF2* gene is expressed from 4 promoters, but after birth, only the promoter P1 allow *IGF2* expression in a few tissues: liver, choroid plexus and leptomeninge (Bergman et al., 2013). In *IGF2* overexpressing mice, the risk to develop cancer increases significantly notably for mammary and pulmonary tumors (Bates et al., 1995; Moorehead et al., 2003). So, major genes at the *H19/IGF2* locus promote cancer development.

To date, the exact role of *91H* in tumorigenesis is poorly described. Recently, two studies have examined the *91H* function in esophageal and colorectal tumors progression (Deng et al., 2014; Gao et al., 2014). However, the *91H* mechanism still unknown. Here, we demonstrate in breast cancer that *91H* increases *H19* and *IGF2* expression. *91H* gene expression allows the maintenance of the methylation difference between the two alleles. We also demonstrate that *91H* increases human breast cancer cells aggressiveness.

MATERIEL AND METHODS

Construction of plasmids

To obtain a plasmid generating shRNA against *91H*, the followed DNA sequence was annealed at 95°C to room temperature, and the duplex was cloned in pSuperRetroNeo+GFP (pSR) (Oligoengine) between *Bgl*II and *Hind*III restriction site.

5' GATCTT**GGCGTCATTCTGATGGGACT**GGAATTCTGT**CCCATCAGAATGACGCCTTTT**TA
3';

5'AGCTTAAAA**GGCGTCATTCTGATGGGAC**AGAATTCAGT**CCCATCAGAATGACGCCA**
A 3'.

In these sequences, siRNA *91H* sense was indicated in bold, siRNA *91H* antisense was indicated in bold grey. Plasmid productions were performed in E. Coli TOP 10 (Invitrogen). Plasmid extraction was performed by using Nucleobond PC1000 (Macherey-Nagel).

Transfection

A total of 5×10^4 cells were plated in 100mm dish. Cells are transfected with 6 µg of plasmid with PEI (Ozyme). After 6h of transfection, the medium was removed. After 24h, the transfected cells were selected with 800µg/ml G418 for at least one month.

RNA extraction, reverse transcription and real-time RT-PCR

RNA extraction and qRT-PCR was performed as previously described (Berteaux et al., 2008). Primers used for qRT-PCR are described in supplemental Table S1.

Genomic DNA extraction

Genomic DNA (gDNA) was extracted in Tris-HCl 10mM pH7.4, NaCl 50mM, EDTA 10mM and SDS 0.2%. After protein digestion with proteinase K (Roche) overnight at 55°C, gDNA was precipitated with NaCl 1.6M and ethanol 100%. gDNA was dissolved in Tris-HCl 10mM pH8, EDTA0.1mM.

Chromatin immunoprecipitation

Chromatin immunoprecipitation was performed on 1×10^6 cells with "Magna ChiP A-chromatin Immunoprecipitation" (17-610, Millipore) and antibodies : anti- c-myc (N-262) (sc-764, Santa Cruz); anti-histone H3 monomethyl Lys9 (#39681, Active motif); anti-CTCF (#07-729, Millipore); anti-histone H3 acetyl (17-610, Millipore). After cell fixation with PAF1% for 10 min, PAF was quenched with glycine for 5 min at room temperature. Cells were harvested in PBS

supplemented with proteases inhibitor (P8340, Sigma Aldrich) and centrifuged at 100 g for 5 min at 4°C. Cells and nucleus are lysed in appropriate lysis buffer supplemented with phosphatases and proteases inhibitors. gDNA was fragmented with Bioruptor UCD200 in 300 to 600bp fragments. Beads were blocked with salmon sperm DNA for 1h at 4°C. Antibodies and chromatin were incubated overnight at 4°C, then, beads were added for 1h at room temperature. After wash, immunoprecipitated DNA were eluted by proteinase K for 2h at 65°C then 95°C for 10 min. Eluted DNA was clean with “nucleospin gel PCR cleanup” (Nucleobond). Immunoprecipitated DNA was quantified by qPCR.

Methylation analysis

gDNA methylation analysis was performed with “Methyl Collector Kit” and “Unmethyl Collector Kit” according to manufacturer’s instructions (#55005, #55004; Active Motif). After gDNA fragmentation in 200 to 500 bp, methylated and unmethylated fragments were extracted and quantified by qPCR.

Cells culture

The T47D estrogen-sensitive and the MDA-MB-231 estrogen-insensitive breast cancer cell lines were obtained from the American Type Culture Collection and maintained routinely in RPMI (Gibco) containing 10% of fetal bovine serum (FBS) and 0.01% of Zell Shield (Minerva Biolabs). Cell lines were cultured at 37°C with 5% CO₂, 95% of air in humidified atmosphere.

Migration assays

Cell migration was determined by wound healing assay. Wound healing test was performed by using Culture-inserts for Live Cell Analysis (Ibidi). A total of 3x10⁴ cells were plated on each compartment of insert and cultured for 5h. The insert was then removed (0h) and cells were further cultured for 24h. Cells were photographed at 0h and after 24h of culture to record the wound width.

Cell invasion was determined by transwell assay. A total of 3x10⁴ cells were seeded on collagen (1/10, Millipore) coated insert (0.8µm, BD Bioscience) of 6 well-plates in RPMI containing 10% FBS. At the end of the experiments, cells migrating to the other side of the filter were stained with 0.5% crystal violet.

Proliferation assays

To obtain cells growth curve, a total of 5x10⁴ cells were plated in 6-well plate. Each day, number of cells in well was counted with hemocytometer.

To measure clonogenicity abilities, cells were plated at low density (5x10² cells/100mm dish) for 14 days. Colonies were stained with 0.5% crystal violet. Coloration was dissolved in acetic acid and absorbance determined at 595 nm.

Tumorigenesis in SCID mice

Female SCID mice (8-weeks-old) were purchased from Pasteur Institute, Lille and kept under pathogen-free condition. Animals were handled in accordance with the European Communities Recommendations for Animals Experimentation. Exponentially growing cells were harvested, resuspended in PBS, and subcutaneously (2×10^6 cells in 150 μ l) injected into the flank fat pad of each mouse. Tumor volume and metastasis were monitored as previously described (Lagadec et al., 2008).

Analysis of cell apoptosis in xenografted tumors

Tumors were fixed, paraffin embedded and cut on 5 μ m tumor sections. Apoptosis was measured using *in situ* cell death detection kit, POD (Roche), according to manufacturer's instructions.

Human breast tumors

Human breast tumors were collected from the CRRC from Lille, France (Centre de Référence Régional en Cancérologie). Tumors were lysed in Trizol by mechanical action. Then, RNA was extracted and precipitated as already described. Relative expression of *H19*, *IGF2* and *91H* was determined by qRT-PCR with primers described in supplemental table S1.

Statistical analysis

Data are expressed as mean values \pm standard error of the mean of at least 3 independent experiments. The statistical analysis was done by using Student's *t*-test and *p* value < 0.05 was considered significant.

RESULTS

91H RNA expression in human breast cancer

The expression of lncRNA *91H* as well as the neighboring genes *H19* and *IGF2* was detected by qRT-PCR in 50 human breast tumors and compared to normal cells (hMEC). We have observed an increase expression of these genes in human breast cancer (Figure 1).

91H RNA increases breast cancer cells tumorigenicity in vitro

To determine the *91H* properties in breast cancer cells, we have generated two breast cancer cell lines (MDA-MB-231 and T47D) invalidated for *91H* by using *91H* shRNA. After G418 selection for at least one month, we have obtained numerous cell clones invalidated for *91H*. All clones displayed same properties and we have chosen randomly one clone per cell line. We have shown a significant decrease of *91H* expression (Figure 2A and sup Figure 1A).

Then, we have studied *91H*-invalidated cells phenotypes. We observed that *91H* invalidation (sh) represses significantly breast cancer cells growth compared to control cells (mock) (Figure 2B and sup Figure 1B). Moreover, when *91H*-sh cells are plated at low density, their capacities to growth decrease significantly compared to control cells (Figure 2C and sup Figure 1C). Similar results were obtained with cells embedded in soft agar and cultured in anchorage independent manner (Figure 2D and sup Figure 1D). By wound healing assay, we observed a decrease of cell migration when cells were invalidated for *91H* (Figure 2E and sup Figure 1E). Indeed, 24h after removing insert for MDA-MB-231, control cells have covered the wound healing unlike sh-cells (Figure 2E). For T47D cells, this time of recovering is 72h due to their low migratory capacities (sup Figure 1E). We also determined the role of *91H* in invasive capacity of breast cancer cells. *91H*-invalidated cells (sh) displayed a decrease invasion (Figure 2F).

91H RNA increases in vivo tumorigenesis of breast cancer cells

We have also researched the effect of *91H* invalidation of cells tumorigenic capacity. We have xenografted MDA-MB-231 cells control (mock) or invalidated for *91H* (sh) in immunodeficient SCID mice. We observed a decrease of tumor volume with *91H*-invalidated cells compared to control cells (Figure 3A). By TUNEL assay, we observed an increase of apoptotic cells in tumor obtained from *91H*-knockdown cells (sh) (Figure 3B). Taken together, we demonstrate that *91H* invalidation repress breast cancer cells proliferation, migration, invasion and tumorigenesis.

91H RNA regulates H19 and IGF2 genes through epigenetic modifications

91H is a long non coding RNA antisense to *H19*. In imprinted loci, numerous studies demonstrate that antisense lncRNA regulate genomic imprinting by inducing epigenetic modifications. The *H19/IGF2* locus is regulated by Imprinting Control Region (ICR), located 2kbp upstream of *H19* gene, and containing CTCF and c-myc binding sites (Hark, 2000; Takai, 2001; Barsyte-Lovejoy, 2006; Zhang, 2015).

By using qRT-PCR, we determined the effect of *91H* knock-down on genes located in *H19/IGF2* locus. The *91H* down-regulation leads to *H19* and *IGF2* repression in MDA-MB-231 (Figure 4A) as well as in T47D (sup Figure 2). These both genes are regulated by binding of CTCF and c-myc factors. To decipher the mechanism of action of *91H* RNA, we performed ChIP of CTCF and c-myc in MDA-MB-231 control cells and in MDA-MB-231 sh cells. After precipitation, we detected genomic DNA (gDNA) associated to CTCF or c-myc by PCR. C-myc protein binds unmethylated ICR on polymorphic site (Barsyte-Lovejoy, 2006). To identify the allele within c-myc interacts, we have used primer specific of polymorphism named myc A and myc B. In control cells (mock), c-myc is associated with allele B (Figure 4B). In *91H*-knockdown cells (sh), we observed that c-myc binds ICR on two alleles (A and B) (Figure 4B, 4C). In human, CTCF binds ICR through seven binding sites. We observed an increase of CTCF association to ICR in *91H*-invalidated cells (Figure 4D) probably due to a biallelic binding of CTCF, as c-myc (Figure 4E). Moreover, CTCF binds preferentially sites 4, 5, 6 and 7 (Figure 4D). So, *91H* RNA regulates CTCF and c-myc ICR-binding (Figure 4B, 4D).

As CTCF and c-myc interact with unmethylated ICR, we have determined DNA methylation on ICR and *H19* promoter in cells with or without *91H* knock-down. DNA methylation occurs on CpG-rich region. We used two DNA extraction kit allowing immunoprecipitation of gDNA fragments with 6 CpG unmethylated or 5 CpG methylated. According to number of CpG present in gDNA fragments, a same fragment could be immunoprecipitated by the two methods. Genomic DNA immunoprecipitated was quantified by qPCR. In *91H*-invalidated cells (sh), a significant high increase (3 folds) of unmethylated CpG was observed on ICR and *H19* promoter compared to control cells (Figure 5A, left and right panel). Surprisingly, we observed also a slight increase of ICR methylation (Figure 5A, left panel). At the *H19* promoter, we observed also a significant increase of CpG methylation (Figure 5A, right panel). These results argue in favor of a repartition of methylation in both alleles at ICR and *H19* promoter loci (Figure 5B) and are consistent with biallelic binding of CTCF and c-myc.

We studied epigenetic modifications affecting histone (acetylation and methylation) on ICR and *H19* promoter by ChIP. In *91H*-invalidated cells (sh), no significant epigenetic modifications

(histone acetylation and methylation) was observed on the ICR compared to control cells (mock) (Figure 6A, 6B, left panel). However, a slight increase of histone H3 methylation on ICR was observed in *91H*-knockdown cells. On the *H19* promoter, surprisingly, we observed a significant increase of histone H3 acetylation and methylation in *91H*-invalidated cells (sh) compared to control cells (mock) (Figure 6A, 6B, right panel).

Taken together, our results suggest that in *91H*-knockdown cells, there is perturbations of epigenetic modifications on two alleles at ICR and *H19* promoter sufficient to repress gene expression (Figure 7).

DISCUSSION

Numerous genomic imprinting loci are regulated by antisense lncRNA (Sanli & Feil, 2015). In the imprinted locus *H19/IGF2*, we have identified a new lncRNA antisense to *H19* gene and conserved between mammals, named *91H* (Berteaux et al., 2008; Tran et al., 2012). In mice, it was established that *91H* RNA regulates *IGF2* gene by promoting its expression from a novel promoter named Pm (Tran et al., 2012). However, the link between *H19* and *91H* remains unclear. Here, we have identified a potential mechanism of *91H* on the *H19/IGF2* locus. The genomic imprinting of *H19/IGF2* locus is regulated by the ICR, located 2kb upstream of *H19* gene. On the maternal allele, the CTCF (CCCTC-binding factor) protein binds the unmethylated ICR. This protein prevents enhancer interaction with the *IGF2* promoter; so only, the *H19* could be expressed from this allele. On the paternal allele, the methylated ICR repress *H19* expression and promotes *IGF2* expression. So, gene expression at the *H19/IGF2* locus is principally regulated by ICR methylation status. In this study, we have shown that *91H* regulates methylation of ICR and *H19* promoter on two alleles (Figure 7). Our results indicate that *91H* is involved in the maintenance of genomic imprinting. Indeed, in *91H*-knockdown cells, a repartition of methylation was observed on the two alleles on ICR and on *H19* promoter. We observed also a repartition of c-myc and CTCF binding (Figure 7). In our previous study, we show that *91H* downregulation by siRNA regulates negatively of *IGF2* and have few effects on *H19* expression (Berteaux et al., 2008). Here, by stably *91H* downregulation, we observed a significant downregulation of *IGF2* and *H19* gene expression. The discrepancies between our two studies are probably due to the technology used to perform loss-of-function experiments, transient or stably knock-down. We propose that *91H* lncRNA prevents methylation on maternal allele by masking methylation site.

Angrand and colleagues have described that lncRNA regulate gene expression through several mechanism (Angrand et al., 2015). One of these mechanism described is epigenetic regulation by lncRNA. LncRNA could regulate gene expression in *cis*. For example, the *HOTTIP* RNA, expressed from the HOX A locus, is able to recruit the histone modifiers complex WDR5-MLL which induces H3K4me3 methylation close to its transcription site. So,

HOTTIP RNA promotes gene expression. The lncRNA *HOTAIR*, expressed from the HOX C locus, regulates negatively expression of HOX D locus by recruiting two different chromatin modifiers complex, PRC2 and LSD1-CoREST. These complexes induce the repressive mark H3K27me3 and remove the active mark H3K4me2/3 respectively. So, *HOTAIR* RNA represses expression of HOX D gene. LncRNA could also be involved in genomic imprinting. For example, lncRNA *Kcnq1ot1*, *Air*, regulate genomic imprinting of their locus, respectively *Kcnq1/Kcnq1ot1*, *Igf2r/Air*, by epigenetic modifications (Sanli & Feil, 2015). Indeed, lncRNA *Kcnq1ot1*, transcript from the paternal allele, recruits histone methyltransferase (such as EZH2) to induce H3K27me3 or H3K9me2 on the same allele and so repress gene expression. The lncRNA *Air*, transcript from the paternal allele, interacts with DNA methyltransferase (G9a) to repress gene expression from the same allele. So, to identify the *91H* mechanism several studies are needed.

Moreover, recently it was observed that *H19* RNA regulates genomic imprinting of several loci by recruiting a methyl binding protein, MBD1 (Monnier et al., 2013). It should be interesting to determine if *91H* RNA could have similar functions.

Recently, two studies have revealed the *91H* role in colorectal and esophageal cancer (Deng et al., 2014; Gao et al., 2015). However, they observed contradictory functions of *91H*. Indeed, in colorectal cancer, it was reported that *91H* promotes tumorigenic cells properties whereas, in esophageal cancer, a tumor suppressor function of *91H* have been revealed. Here, we have determined the function of *91H* in breast cancer cells. In *91H*-invalidated cells, we observed a decrease of cell proliferation, clonogenicity, migration, invasion, and tumorigenesis. Similarly to Deng et al., we have shown that *91H* increases tumorigenic properties of breast cancer cells.

Previously, we have shown that *H19* promotes tumorigenic properties of breast cancer cells (Berteaux et al., 2005). Moreover, *IGF2* overexpression promotes also tumorigenesis (Bates et al., 1995; Moorehead et al., 2003). As, we have shown that *91H* regulates positively *H19* and *IGF2* expression, it is possible that phenotype observed in *91H*-invalidated cells were due to *H19* and *IGF2* repression.

Recently, *HOTS*, a new antisense transcript to *H19*, has been identified (Onyango & Feinberg, 2012). They have identified *HOTS* in several tissues except in mammary gland and it is possible that *HOTS* transcript and protein are not expressed in this tissue. Even if this gene is expressed in mammary gland, our *91H* shRNA does not overlap *HOTS*, so our observations are *HOTS*-independent.

In summary, in breast cancer cells, we identified a new mechanism of *91H* RNA to regulate expression of the *H19/IGF2* imprinted locus. We also identified that *91H* increases cell

tumorigenic capacities *in vitro* and *in vivo*. We reveal that 91H promotes expression of two oncogenes (*H19* and *IGF2*), so it is very important to identify all mechanisms associated to 91H to reveal a new potential therapeutic target.

BIBLIOGRAPHIE

Adriaenssens, E. et al. (1998) H19 overexpression in breast adenocarcinoma stromal cells is associated with tumor values and steroid receptor status but independent of p53 and Ki-67 expression. *Am. J. Pathol.* **153**, 1597–1607.

Adriaenssens, E. et al. (1999). Steroid hormones modulate H19 gene expression in both mammary gland and uterus. *Oncogene*, **18**, 4460–4473.

Angrand, P.-O. et al. (2015). The role of long non-coding RNAs in genome formatting and expression. *Front. Genet.* **165**.

Ariel, I. et al. (1995). The imprinted H19 gene as a tumor marker in bladder carcinoma. *Urology*, **45**, 335–338.

Ariel, I. et al. (1997). Genomic imprinting and the endometrial cycle. The expression of the imprinted gene H19 in the human female reproductive organs. *Diagn. Mol. Pathol.*, **6**, 17–25.

Barsyte-Lovejoy, D. et al. (2006). The c-Myc oncogene directly induces the H19 noncoding RNA by allele-specific binding to potentiate tumorigenesis. *Cancer Res.*, **66**, 5330–5337.

Bates, P. et al. (1995). Mammary cancer in transgenic mice expressing insulin-like growth factor II (IGF-II). *Br. J. Cancer*, **72**, 1189–1193.

Bergman, D. et al. (2013). Insulin-like growth factor 2 in development and disease: a mini-review. *Gerontology*, **59**, 240–249.

Berteaux, N. et al. (2005). H19 mRNA-like noncoding RNA promotes breast cancer cell proliferation through positive control by E2F1. *J. Biol. Chem.*, **280**, 29625–29636.

Berteaux, N. et al. (2008). A novel H19 antisense RNA overexpressed in breast cancer contributes to paternal IGF2 expression. *Mol. Cell. Biol.*, **28**, 6731–6745.

Cai, X. et al. (2007). The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA*, **13**, 313–316.

Deng, Q. et al. (2014). Up-regulation of 91H promotes tumor metastasis and predicts poor prognosis for patients with colorectal cancer. *PLoS ONE*, **9**, e103022.

Gao, T. et al. (2015). Long non-coding RNA 91H contributes to the occurrence and progression of esophageal squamous cell carcinoma by inhibiting IGF2 expression. *Mol. Carcinog.*, **54**, 359–367.

Hark, A.T. et al. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature*, **405**, 486–489.

Hibi, K. et al. (1996). Loss of H19 imprinting in esophageal cancer. *Cancer Res.*, **56**, 480–482.

Kondo, M. et al. (1995). Frequent loss of imprinting of the H19 gene is often associated with its overexpression in human lung cancers. *Oncogene*, **10**, 1193–1198.

Lagadec, C. et al. (2009). TrkA overexpression enhances growth and metastasis of breast cancer cells. *Oncogene*, **28**, 1960–1970.

Li, H. et al. (2014). Overexpression of lncRNA H19 enhances carcinogenesis and metastasis of gastric cancer. *Oncotarget*, **5**, 2318–2329.

Résultats

- Liu, J. et al. (1995). H19 and insulin-like growth factor-II gene expression in adrenal tumors and cultured adrenal cells. *J. Clin. Endocrinol. Metab.*, **80**, 492–496.
- Lottin, S. et al. (2002). Overexpression of an ectopic H19 gene enhances the tumorigenic properties of breast cancer cells. *Carcinogenesis*, **23**, 1885–1895.
- Monnier, P. et al. (2013). H19 lncRNA controls gene expression of the Imprinted Gene Network by recruiting MBD1. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 20693–20698.
- Moorehead, R.A. et al. (2003). Transgenic overexpression of IGF-II induces spontaneous lung tumors: a model for human lung adenocarcinoma. *Oncogene*, **22**, 853–857.
- Onyango, P. et al. (2011). A nucleolar protein, H19 opposite tumor suppressor (HOTS), is a tumor growth inhibitor encoded by a human imprinted H19 antisense transcript. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 16759–16764.
- Sanli, I. et al. (2015). Chromatin mechanisms in the developmental control of imprinted gene expression. *Int. J. Biochem. Cell Biol.*
- Takai, D. et al. (2001). Large scale mapping of methylcytosines in CTCF-binding sites in the human H19 promoter and aberrant hypomethylation in human bladder cancer. *Hum. Mol. Genet.*, **10**, 2619–2626.
- Tran, V.G. et al. (2012). H19 antisense RNA can up-regulate Igf2 transcription by activation of a novel promoter in mouse myoblasts. *PLoS ONE*, **7**, e37923.
- Tsang, W.P. et al. (2010). Oncofetal H19-derived miR-675 regulates tumor suppressor RB in human colorectal cancer. *Carcinogenesis*, **31**, 350–358.
- Zhang, E. et al. (2015). c-Myc-regulated long non-coding RNA H19 indicates a poor prognosis and affects cell proliferation in non-small-cell lung cancer. *Tumour Biol.*

FIGURES LEGENDS

Figure 1. *H19/IGF2* locus expression in human breast tumors. Relative expression of *H19*, *IGF2* and *91H* gene in 41 human breast tumors compared to normal breast cells (hMEC) determined by qRT-PCR. *** $p < 0.001$

Figure 2. *91H* promotes cells oncogenic properties. (A) Relative *91H* expression determined by qRT-PCR in MDA-MB-231 control cells (mock) or *91H*-invalidated cells (sh). (B) Proliferation of MDA-MB-231 control cells (mock) and *91H*-knockdown cells (sh). (C) Clonogenicity assay of breast cancer cells MDA-MB-231 control (mock) and *91H*-knockdown cells (sh). Two weeks after plating, colonies are fixed and stained with violet crystal 0.5%. After dissolution of coloration in acetic acid, absorbance was determined at 595nm. (D) Clonogenicity on agar of breast cancer cells control (mock) and *91H*-invalidated cells (sh) determined 14 days after plating. (E) Migratory capacities of breast cancer cells and MDA-MB-231 control (mock) and *91H*-invalidated cells (sh) determined by wound healing assay. (F) Invasive capacities of MDA-MB-231 control (mock) and *91H*-knockdown cells (sh) determined by transwell assay. * $p < 0.05$ *** $p < 0.001$

Figure 3. *91H* increases breast cancer cells tumorigenicity. (A) Breast cancer cells (MDA-MB-231) control (mock) and *91H*-invalidated cells (sh) were subcutaneously injected into SCID mice (11 mice per group). Tumor growth curve represents the mean of tumor volumes in each group and error bar, sem. (B) Detection of cells apoptosis (TUNEL) of paraffin-embedded sections of tumors. Scale bar is 100 μ m. ** $p < 0.01$ *** $p < 0.001$

Figure 4. *91H* upregulates CTCF, c-myc interaction on Imprinting Control Region. (A) Relative expression of *91H*, *H19* and *IGF2*, determined by qRT-PCR in MDA-MB-231 control cells or transfected with *91H* shRNA (sh). (B) ChIP c-myc in control cells (mock) and in *91H*-invalidated cells (sh). C-myc binding on polymorphic ICR site was detected by PCR was detected by PCR. (C) Schematic representation of c-myc binding in *91H*-invalidated cells (sh) versus normal cells (mock). (D) ChIP CTCF in control cells (mock) and in *91H*-invalidated cells (sh). CTCF fixation on the seven binding site was detected by PCR. (E) Schematic representation of CTCF binding in *91H*-depleted cells (sh) versus normal cells (mock).

Figure 5. *91H* regulates methylation on the Imprinting Control Region and the *H19* promoter. (A) Detection of ICR and *H19* promoter methylation in breast cancer cells (MDA-MB-231) control (mock) or stably transfected with *91H* shRNA (sh) by immunoprecipitation of methylated and un-methylated DNA. (B) Mechanism of *91H* RNA on the ICR and the *H19* promoter.

Figure 6. *91H* regulates epigenetic modifications on the Imprinting Control Region and the *H19* promoter. (A, B) Detection of histone H3 acetylation (A) and methylation (B) on the ICR and the *H19* promoter by ChIP followed by qPCR in MDA-MB-231 control cells (mock) and *91H*-invalidated cells (sh).

Figure 7. Mechanism of *91H* at the *H19/IGF2* locus. In normal condition, on the maternal allele, CTCF binds the unmethylated ICR. This protein act as an insulator, so only *H19* is expressed from this allele. The c-myc protein promotes *H19* expression from the maternal allele by binding on the unmethylated ICR. On the paternal allele, methylated ICR repress *H19* expression and prevents CTCF and c-myc interaction. So, only *IGF2* gene is express from this allele. In *91H*-knockdown condition (sh), we observed a repartition between the two alleles of CTCF binding. We observed also a repartition of methylation on ICR and *H19* promoter. These mechanism allow a decrease of expression of *H19* and *IGF2* gene. Gene expression is represented by arrow and transcriptional noise by dotted arrow.

SUPPLEMENTAL LEGENDS

Supplemental Figure 1. *91H* promotes T47D cells oncogenic properties. (A) Relative *91H* expression determined by qRT-PCR in T47D control cells (mock) or *91H*-invalidated cells (sh). (B) Proliferation of T47D control cells (mock) and *91H*-knockdown cells (sh). (C) Clonogenicity assay of breast cancer cells T47D control (mock) and *91H*-knockdown cells (sh). Two weeks after plating, colonies are fixed and stained with violet crystal 0.5%. After dissolution of coloration in acetic acid, absorbance was determined at 595nm. (D) Clonogenicity on agar of breast cancer cells control (mock) and *91H*-invalidated cells (sh) determined 14 days after plating. (E) Migratory capacities of breast cancer cells and MDA-MB-231 control (mock) and *91H*-invalidated cells (sh) determined by wound healing assay. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Supplemental Figure 2. Relative expression of *91H*, *H19* and *IGF2*, determined by qRT-PCR in T47D control cells or transfected with *91H* shRNA (sh). ** $p < 0.01$ *** $p < 0.001$

Supplemental Table 1. Primers used in this study.

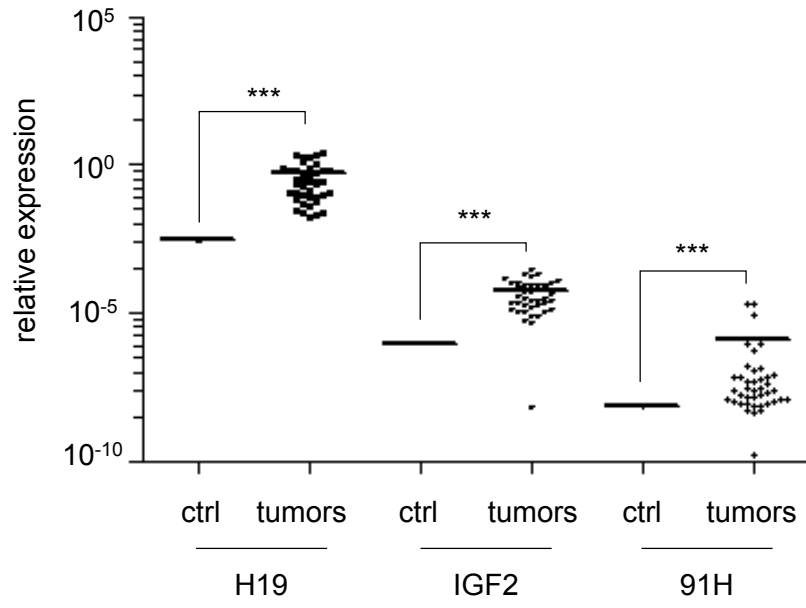


Figure 1.

Résultats

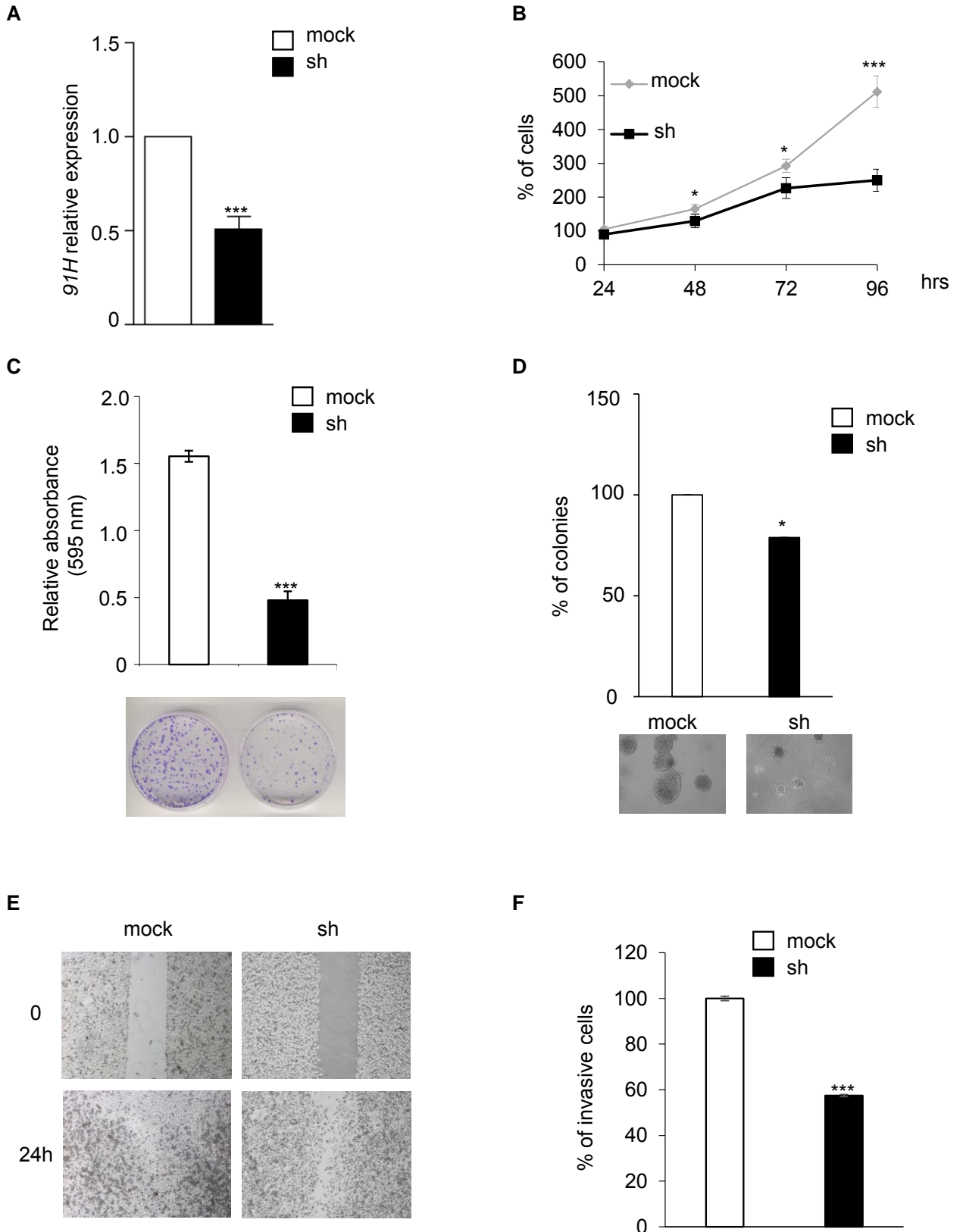


Figure 2.

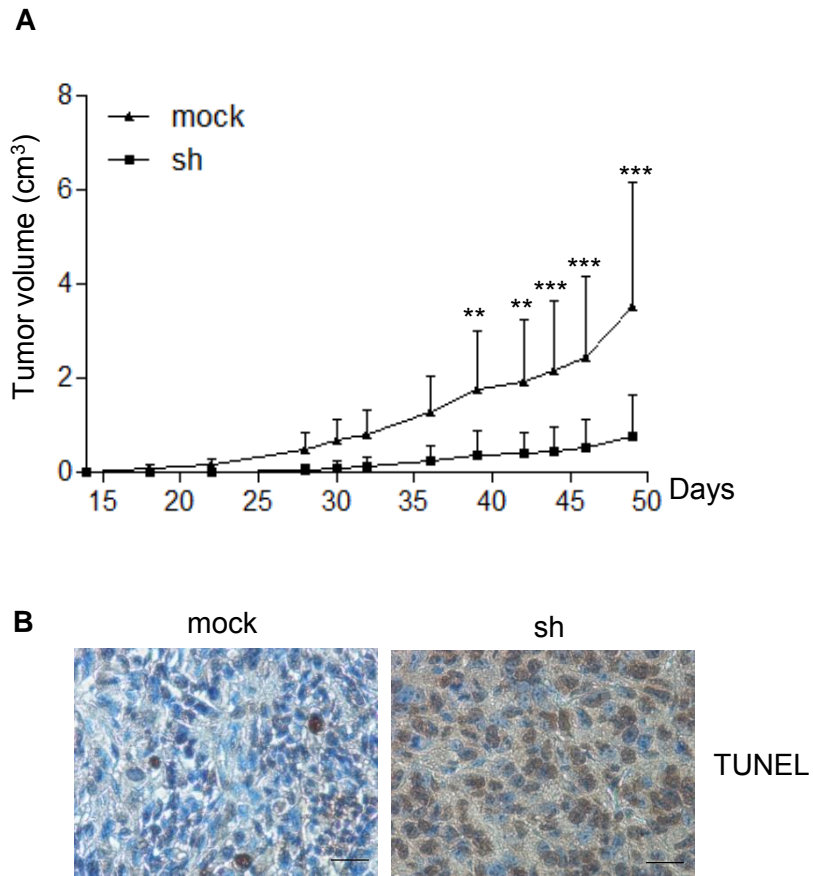


Figure 3.

Résultats

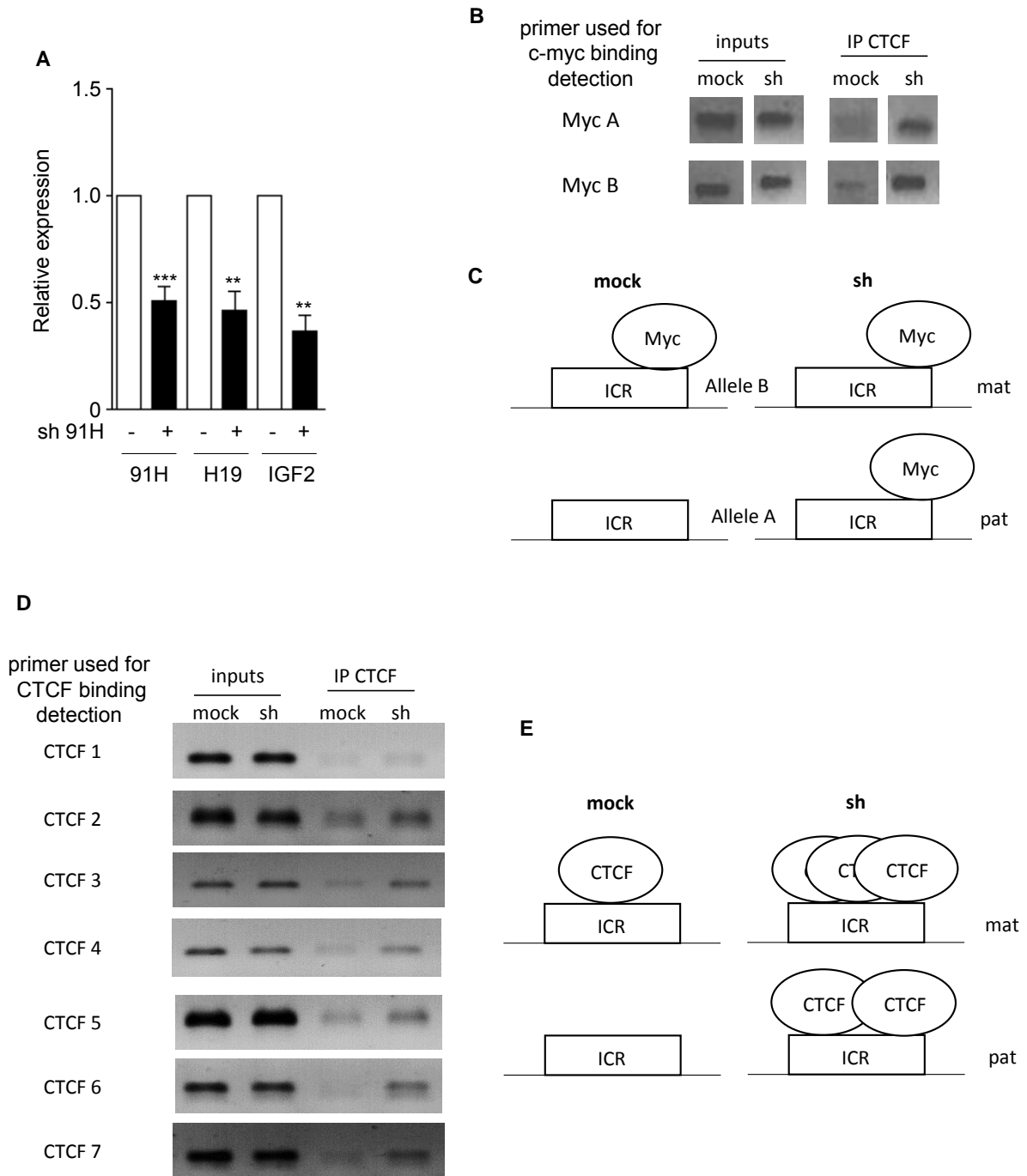


Figure 4.

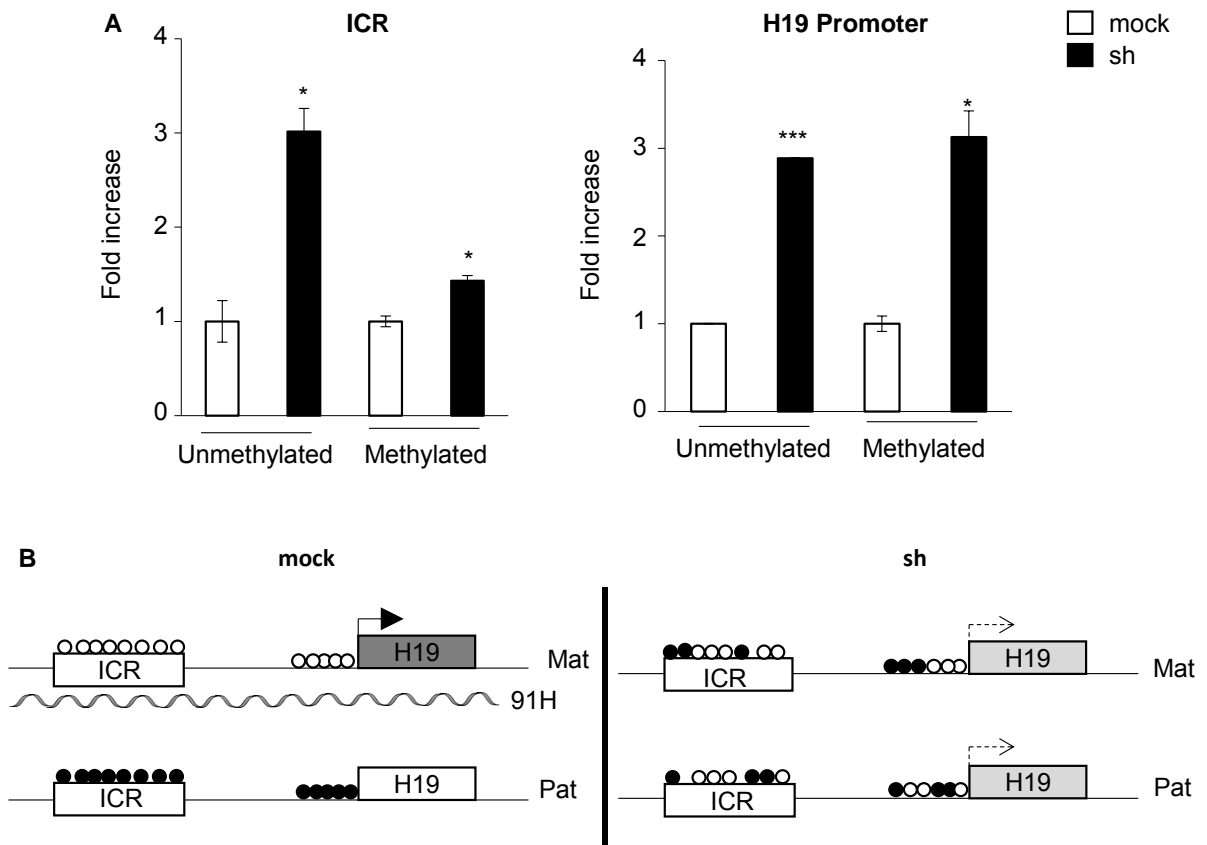


Figure 5.

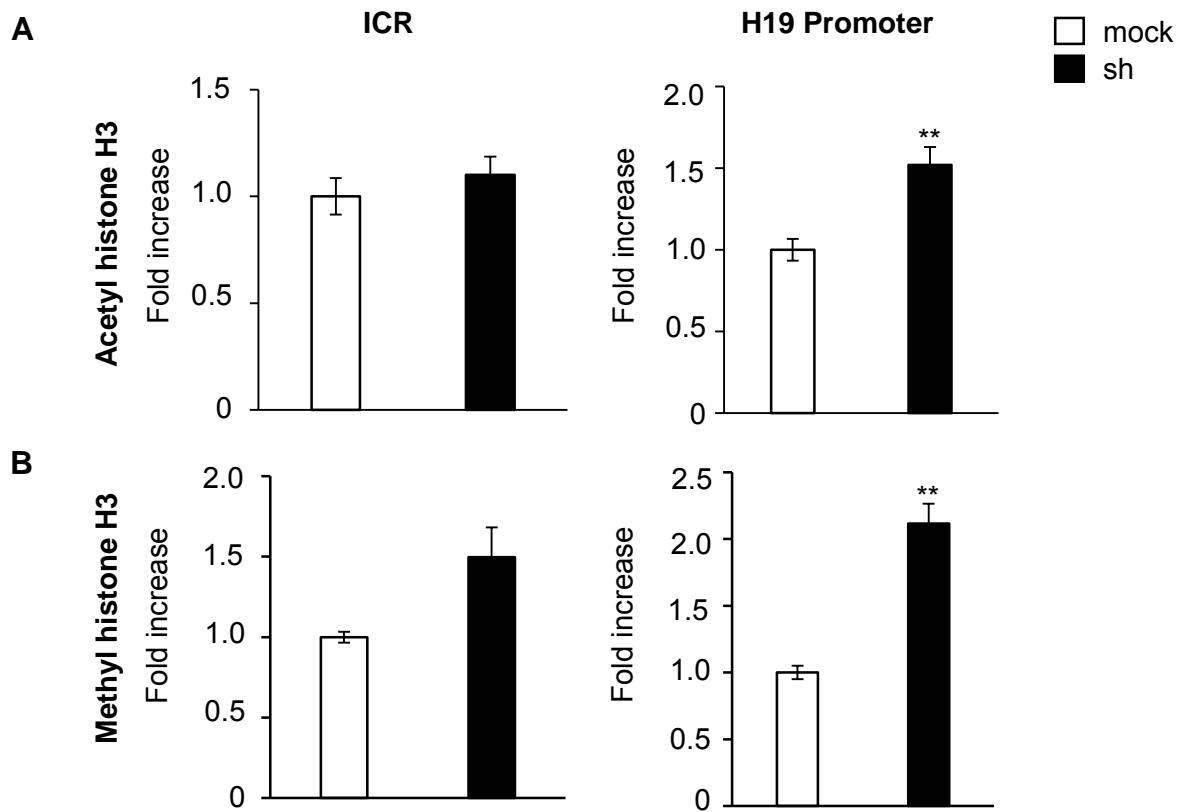


Figure 6.

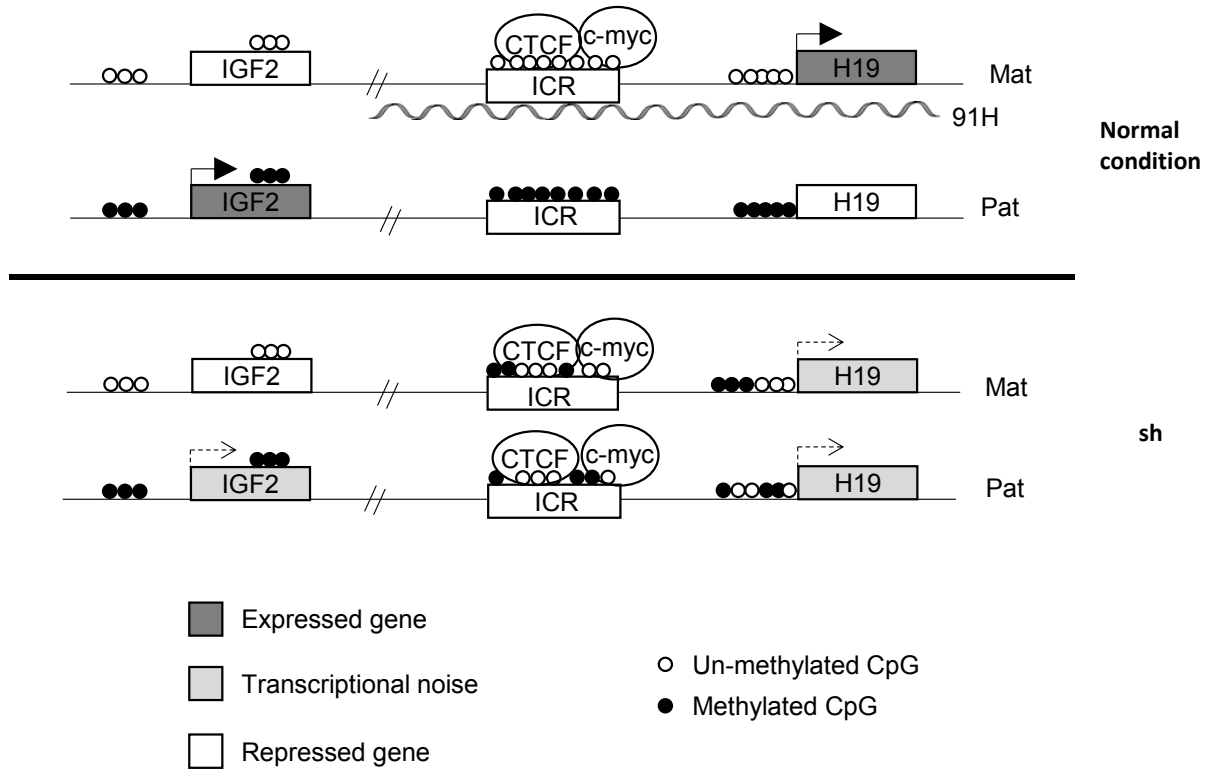
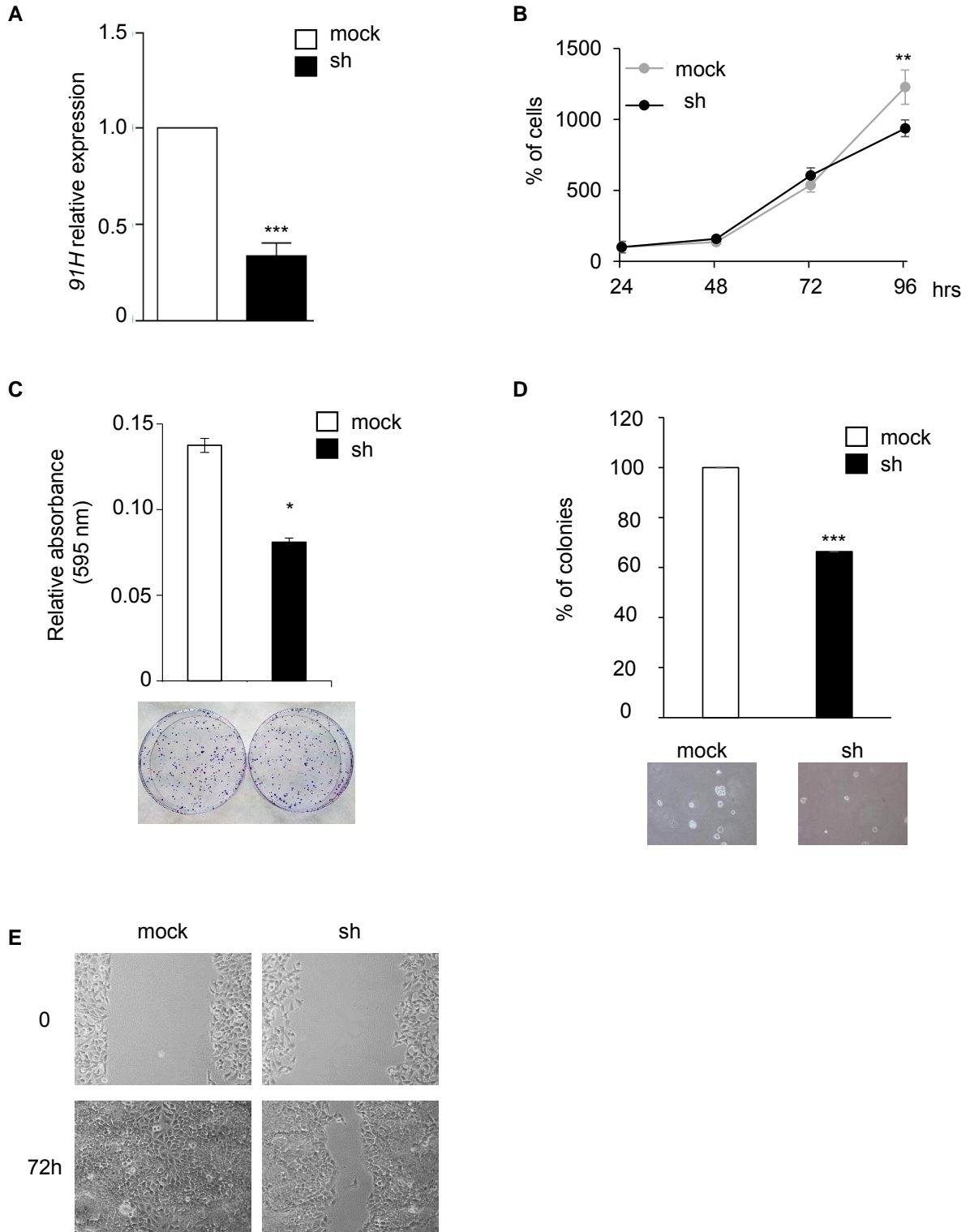
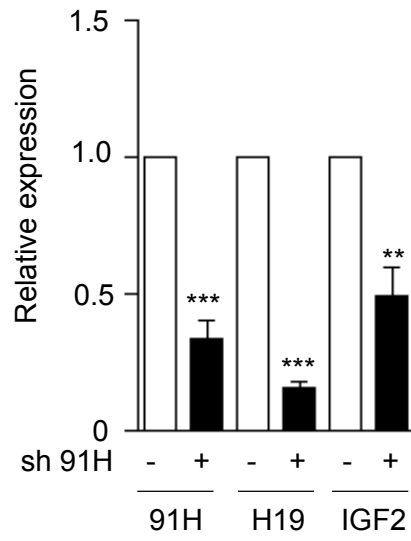


Figure 7.

Résultats



Supplemental Figure 1.



Supplemental Figure 2.

Name	5' to 3'
RPLP0 sense	GTGATGTGCAGCTGATCAAGACT
RPLP0 antisense	GATGACCAGCCCAAAGGAGA
H19 sense	TACAACCACTGCACTACCTG
H19 antisense	TGGCCATGAAGATGGAGTCG
IGF2 sense	GGCTTGTACTIONCTGGCAGGCC
IGF2 antisense	GAACCTGCCACGGGGTATC
91H sense	GCTTGTAGTAGAGTGCGCC
91H antisense	CATCCAGTTGACCGAGCTTG
promoteur H19 sense	GCTTGTAGTAGAGTGCGCC
promoteur H19 antisense	CATCCAGTTGACCGAGCTTG
ICR sense	GGGACTCAAGTCACGCCTAC
ICR antisense	CACCTCAAACCTGCATTGAA
CTCF1 sense	GCTTCCCATTGAGTCTTGGA
CTCF1 antisense	GAGACCTGGGACGTTTCTGT
CTCF2 sense	CACATTGGGCTACGTGTCTG
CTCF2 antisense	TGGTGAGACAGAAGGGGAAG
CTCF3 sense	TTCCCCTTCTGTCTCACCAC
CTCF3 antisense	GATATGGCCCGATACGAAGA
CTCF4 sense	GTGTCCCATTCTTTGGATGC
CTCF4 antisense	CATCTCCCAACCCTCAATA
CTCF5 sense	GTGTCCCATTCTTTGGATG
CTCF5 antisense	TTGGATGACCTGGGATGTT
CTCF6 sense	CTCTTGATAGCACATGGGT
CTCF6 antisense	GTGTCTATTCCAGATGAC
CTCF7 sense	GAGTCTCCCCTTGGGTCTTG
CTCF7 antisense	TGAGTGGACCCCAAGTTAG
c-myc A sense	CGCCTACTTATGTGATGATCAG
c-myc B sense	CGCCTACTTATGTGATGATCAC
c-myc antisense	CTGCACCCACGATAATGGATT

Supplemental Table S1.

Discussion

Dans cette étude, nous avons mis en lumière la fonction de l'ARN non codant 91H dans les cellules cancéreuses mammaires ainsi que sa fonction au locus *H19/IGF2*. En effet, nous montrons qu'une invalidation de l'expression de 91H par l'utilisation d'ARN interférant diminue les capacités tumorigènes des cellules cancéreuses mammaires (croissance, migration/invasion, clonogénicité, formation de tumeur en souris immunodéficientes). Ainsi, dans les cellules cancéreuses mammaires, 91H possède des propriétés oncogéniques.

Par ailleurs, dans une cohorte de tumeurs mammaires humaines, une forte expression des gènes *H19*, *IGF2* et *91H* a été constatée. De plus, nous observons une diminution de l'expression des gènes *H19* et *IGF2* en absence de *91H* dans les cellules cancéreuses mammaires. Le locus *H19/IGF2* étant soumis à l'empreinte génomique parentale, nous avons alors déterminé les fonctions de *91H* sur ce locus. Une répartition de la méthylation au niveau de l'ICR et du promoteur d'*H19* est observée en absence de *91H*. Nous remarquons également une augmentation de la fixation des protéines CTCF et c-myc sur l'ICR en absence d'expression de *91H*. Nous mettons ainsi en évidence pour la première fois que l'ARN *91H* est impliqué dans le maintien de l'empreinte génomique au locus *H19/IGF2*.

Dans cette étude, nous avons identifié la fonction de 91H au locus H19/IGF2 ainsi que sa fonction dans les cellules cancéreuses mammaires. Nous avons mis en évidence un nouveau mécanisme impliqué dans la tumorigénèse de la glande mammaire chez la Femme.

DISCUSSION-PERSPECTIVES

Les longs ARN non codants sont impliqués dans divers processus biologiques et pathologiques au sein d'une cellule. Pour exercer leurs fonctions, ils possèdent plusieurs mécanismes d'action qui seront différents selon leur localisation cellulaire. En effet, dans le noyau, les lncRNA peuvent réguler l'expression des gènes par modifications épigénétiques ou ils peuvent empêcher l'export des ARNm dans le cytosol en les maintenant dans des sous-domaines nucléaires : les paraspeckles. Dans le cytosol, les lncRNA peuvent être des précurseurs ou des éponges à microARN, ils peuvent également interagir avec des protéines pour en moduler leur fonction ou leur localisation cellulaire. Le locus *H19/IGF2* comporte plusieurs lncRNA : *H19*, *91H*, *PIHit*, *IGF2-as* dont les fonctions et modes d'action sont encore méconnus. Au laboratoire, nous nous intéressons aux ARN non codants transcrits à partir de l'allèle maternel : *H19* et *91H*. Nous avons déjà mis en évidence que *H19* est capable d'accroître les propriétés tumorigènes des cellules cancéreuses mammaires lorsqu'il est surexprimé (Lottin et al., 2002). Mais son mode d'action dans ce cancer n'était pas encore défini. Par ailleurs, lors de mon arrivée, l'équipe venait de mettre en évidence l'existence de l'ARN non codant *91H* dont la fonction et le rôle au cours de la tumorigenèse mammaire n'étaient pas identifiés (Berteaux et al., 2008).

➤ ***H19*, via son microARN, favorise la tumorigenèse**

Dans ce contexte, j'ai, dans un premier temps, identifié le(s) mode(s) d'action de l'ARN non codant *H19* au cours de la tumorigenèse de la glande mammaire. *H19* étant précurseur des microARN, miR-675-5p et miR-675-3p, dans les cellules cancéreuses mammaires, j'ai identifié plusieurs cibles du microARN miR-675-5p. En effet, j'ai mis en évidence que le **miR-675 régule négativement** l'expression des protéines de la famille **Cbl** (c-Cbl et Cbl-b) en se fixant sur les séquences codantes des ARNm. Cette nouvelle régulation permet une meilleure activation des récepteurs aux facteurs de croissance (EGF, HGF, pro-NGF, NGF). De plus, l'activation de ces récepteurs et de leurs voies de signalisation sous-jacente se maintient dans le temps lorsque les cellules surexpriment *H19* ou son microARN.

Les protéines c-Cbl et Cbl-b sont des E3-ubiquitine ligases appartenant à une même famille. Ces protéines sont composées de plusieurs domaines : un domaine de fixation aux tyrosines kinases, un domaine en doigt de zinc, une région riche en prolines et un domaine d'association à l'ubiquitine. Ces protéines reconnaissent des récepteurs activés ; elles permettent leur internalisation ainsi que leur dégradation par les lysosomes. Elles se fixent sur ces récepteurs par l'intermédiaire des protéines Grb2/SOS : deux protéines impliquées dans la transduction du signal. Il est établi que les protéines Cbl peuvent faire parties d'un complexe tripartite contenant les protéines Cdc42, p85Cool-1/ β Pix et c-Cbl (Hirsch et al., 2006). Par exemple, dans les cancers du sein, la phosphorylation de l'EGFR permet le recrutement de la protéine c-Cbl associée à la protéine Cdc42. Ce complexe protéique permet alors l'internalisation et la

dégradation de l'EGFR. Dans de nombreux cancers, une dérégulation de l'expression de c-Cbl ou une mutation sur cette protéine permet la prolifération, la migration ou encore la survie des cellules cancéreuses (Hirsch et al., 2006). Ici, dans cette étude, nous avons mis en évidence un nouveau mode de régulation de l'expression des protéines Cbl dans les cellules cancéreuses mammaires. Cette régulation permet une activation plus forte et prolongée des récepteurs aux facteurs de croissance. Ces effets moléculaires se traduisent par une prolifération, une migration mais aussi des capacités tumorigènes cellulaires augmentées. Précédemment, nous avons montré que les facteurs de croissance, notamment l'EGF et l'HGF, favorisent l'expression du gène *H19* (Adriaenssens et al., 2002). Nous montrons donc qu'il existe une boucle positive de rétrocontrôle entre les facteurs de croissance et le gène *H19*. Par ailleurs, de nombreuses thérapies ciblent les récepteurs aux facteurs de croissance et induisent la dégradation de ces derniers par l'intermédiaire de Cbl. Par exemple, l'icotinib, dans les cellules cancéreuses du poumon, favorise la dégradation d'EGFR par Cbl et permet un arrêt des cellules dans le cycle cellulaire (Mu et al., 2013). Dans notre étude, nous avons identifié un mécanisme de régulation des protéines Cbl pouvant être à l'origine des résistances des tumeurs aux thérapies usuelles. ***L'ensemble de nos données suggère que H19 est impliqué dans la résistance des tumeurs aux thérapies en exacerbant les voies d'autocrinie et de paracrinie.***

Dans le cancer du sein, le statut en récepteur est un outil de classification des tumeurs et peut être une cible thérapeutique. En effet, notre équipe a mis en évidence que les cellules tumorales sont capables de sécréter des facteurs de croissance, notamment du NGF (Descamps et al., 2001 ; Dollé et al., 2003 ; Vanhecke et al., 2011 ; Demont et al., 2012) et de modifier leur microenvironnement, notamment en activant les fibroblastes et les macrophages (fig 6), pour augmenter leurs capacités de prolifération et de migration/invasion. Par ailleurs, notre équipe a mis en évidence que les facteurs de croissance (NGF par exemple) sont capables d'induire une signalisation particulière, indépendante de leur récepteur afin de maintenir l'agressivité tumorale même en présence d'inhibiteurs (Lestaurtinib) (Aubert et al., 2015). ***Ainsi, l'activation des récepteurs aux facteurs de croissance est un mécanisme nécessaire à la croissance tumorale ; H19, par l'intermédiaire de son microARN, exacerbe ce mécanisme.***

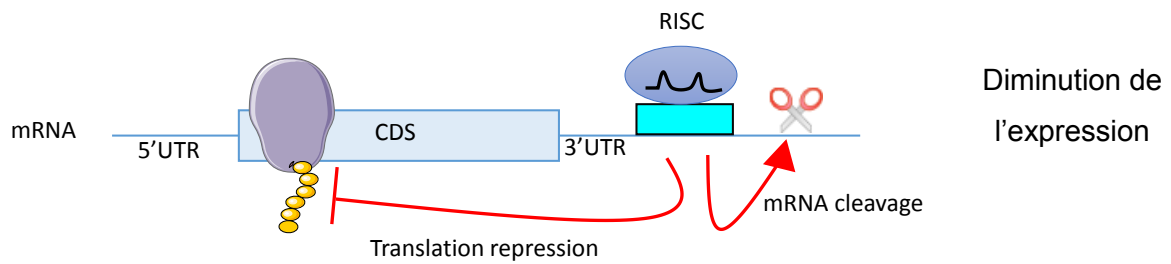
Par ailleurs, j'ai également mis en évidence que le miR-675 régule l'expression de l'ARNm FADD en se fixant sur l'UTR3' de celui-ci. Tout d'abord, mes données montrent une meilleure stabilité de l'ARNm FADD dans les cellules surexprimant *H19*. Ensuite, j'ai montré que le ***miR-675, en se fixant sur l'UTR3' FADD, favorise l'expression de ce gène*** dans les cellules cancéreuses mammaires. FADD étant une protéine principalement impliquée dans l'apoptose, j'ai étudié les effets de sa surexpression dans les cellules cancéreuses mammaires. De

manière surprenante, j'ai observé que les cellules surexprimant *H19* sont résistantes à l'apoptose induite par plusieurs apoptogènes (le trail, par exemple). Cette résistance peut être expliquée par la hausse de p-FADD dans ces cellules. Cette modification post-traductionnelle régule les fonctions de FADD. En effet, p-FADD favorise le cycle cellulaire, notamment la transition G2/M (Zhang et al., 2004b). Nous avons déjà montré que *H19* favorise la transition G1/S du cycle cellulaire, ici nous montrons qu'il permet également la transition G2/M (Berteaux et al., 2005).

J'ai également mis en évidence une fonction inattendue du miR-675 sur l'ARNm FADD. Une régulation positive du miR-675 sur cet ARNm a été observée. Quelques études ont déjà montré une régulation positive d'un microARN sur un ARNm cible : Vasudevan et al., 2007 ; Ørom et al., 2008 ; Henke et al., 2008. Cependant, ces études étaient réalisées dans des conditions particulières. En effet, Vasudevan et collaborateurs ont observé un effet positif de leur microARN sur l'ARNm du TNF α lorsque les cellules sont arrêtées dans le cycle. Dès l'ajout de sérum dans le milieu, comme attendu, leur microARN régule négativement cet ARNm. Ørom et collaborateurs ont observé un effet positif du miR-10a uniquement lorsqu'il se fixe sur l'UTR5' des ARNm codants pour les protéines ribosomales. De plus, cette régulation positive se produit uniquement en présence de motif dit « 5'TOP » sur l'UTR5' des ARNm. Enfin, la troisième étude est réalisée dans un contexte particulier : l'infection par le virus de l'Hépatite C. Nous avons donc recherché le mécanisme responsable de cette activation d'expression. Vasudevan et collaborateurs avaient mis en évidence l'existence de deux complexes RISC différents en fonction du cycle cellulaire. Lorsque les cellules sont privées en sérum, ils montrent que la protéine FXR1 est associée à la protéine AGO2 du RISC. Ils supposent que cette association est responsable de l'effet positif qu'ils observent. Dans mon analyse, lorsque j'ai identifié les trois cibles du miR-675, les cellules étaient en phase de prolifération. Ces connaissances ne sont donc pas applicables à mon étude.

FXR1 est une *RNA-binding protein*, membre de la famille FXR (*Fragile-X related protein*) qui est constituée de 3 protéines : FMRP, FXR1 et FXR2. Elles sont principalement impliquées dans le développement neuronal, musculaire ou encore cardiaque. Un défaut de leur expression est à l'origine du syndrome de l'X fragile. Ces protéines possèdent des domaines de fixation aux ARNm, nommés RGG box. Lorsqu'elles sont associées aux ARNm, elles peuvent moduler leur stabilité, leur traductibilité ou encore leur localisation cellulaire (Bardoni et al., 2001 ; Li et al., 2014). La protéine FXR1 se fixe sur les UTR3' des ARNm dans des régions riches en AU (ARE), notamment l'ARNm du TNF α lors de la réponse immune et favorise sa stabilité ainsi que sa traductibilité (Khera et al., 2010). De plus, l'association de la protéine FXR1 aux polyribosomes a été montrée.

A



B

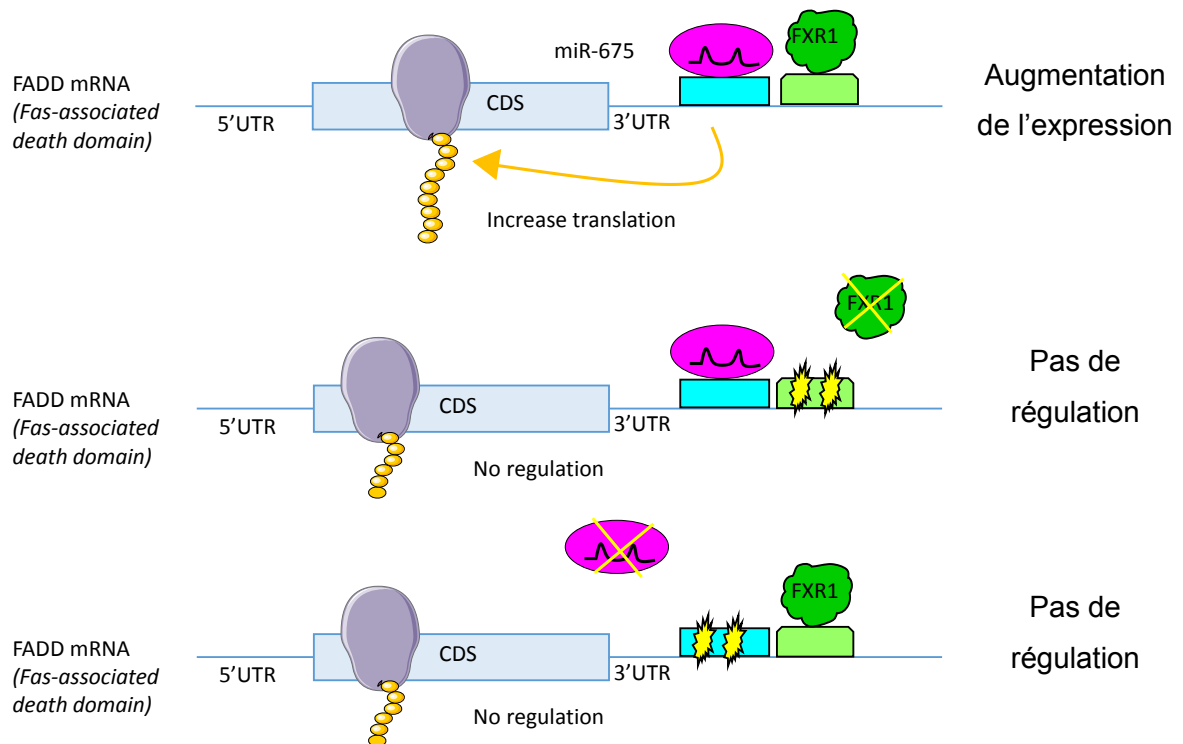


Figure 35. Nouveau mécanisme d'action des microARN. (A) Mode d'action connu des microARN. Le microARN associé au complexe RISC réprime son ARNm cible en empêchant sa traduction ou en le clivant. (B) Nouveau mode d'action des microARN. La protéine FXR1 associée à l'ARNm FADD permet le recrutement du miR-675 associé au complexe RISC. La protéine FXR1 et le miR-675 permettent une meilleure stabilité et traductibilité de l'ARNm FADD. En présence d'une mutation sur le site de fixation de FXR1 (schéma du milieu) ou d'une mutation sur le site de fixation du miR-675 (schéma du bas), un niveau basal de l'expression de FADD a été observé.

Dans notre étude, nous avons mis en évidence une nouvelle fonction pour cette protéine. En effet, nous montrons que l'effet positif du miR-675 sur l'ARNm FADD est FXR1-dépendant. Nous montrons également que la fixation de FXR1 sur l'ARNm n'empêche pas la fixation du miR-675 puisqu'une mutation de la *seed sequence* réprime l'action de ce microARN (fig 27). Nous montrons également que la présence de FXR1 sur l'ARNm FADD est nécessaire au recrutement du complexe miR-675/RISC.

Il est déjà établi que certaines protéines associées aux ARNm régulent la fonction des microARN (Loffreda et al., 2015). En effet, ces protéines peuvent favoriser le recrutement du complexe RISC sur un ARNm soit en se fixant au complexe RISC soit en permettant l'ouverture des structures secondaires de l'ARNm. Ces deux mécanismes permettent une meilleure reconnaissance et interaction entre un microARN et son ARNm cible. Nous pouvons citer par exemple les protéines FMRP et TDP-43 qui permettent le recrutement des miR-125a, miR-125b et miR-NID1 sur leur ARNm cible. En effet, Edbauer et collaborateurs ont montré que la protéine FMRP est essentielle au recrutement et à la fonction des miR-125b et miR-132 sur l'ARNm *NR2A* dans les cellules neuronales de souris (Edbauer et al., 2010). Les protéines associées aux ARNm peuvent également agir en compétition avec le complexe RISC et masquer le site de fixation du microARN. C'est le cas par exemple des protéines HuR, HuD et hnRNP L qui empêchent la fixation des microARN miR-494, miR-129 et miR-297, miR-299 sur leurs cibles.

Dans notre étude, nous montrons que la protéine FXR1 permet le recrutement du complexe RISC sur l'ARNm FADD. Nous montrons également que l'association miR-675/RISC-FXR1 sur l'ARNm FADD permet une meilleure traductibilité de l'ARNm (fig 35). Nous pouvons donc supposer l'existence de deux complexes RISC différents dont l'un favoriserait la répression des ARNm en inhibant leur traduction ou en les dégradant et l'autre favoriserait l'expression de ces derniers. Ces complexes seraient recrutés différemment en fonction des protéines associées à l'ARNm et à proximité du site de fixation des microARN.

Contrairement aux travaux de Vasudevan qui montraient l'existence de deux complexes RISC en fonction de l'état cellulaire (prolifération *versus* arrêt de cycle), nous montrons ici qu'un microARN, dans des cellules dans le même état (phase exponentielle de croissance), peut réguler positivement et négativement deux ARNm différents. Au regard des fonctions biologiques et pathologiques des microARN dans les cellules, il serait intéressant de généraliser ce nouveau mécanisme à l'ensemble des microARN et dans toutes les cellules.

Nous avons donc identifié un nouveau mode de régulation de l'action des microARN. Nous montrons que le recrutement des microARN sur leur cible peut être dirigé par les protéines associées aux ARNm (FXR1 par exemple). Nous montrons également que la

protéine FXR1 associée au miR-675 permet une régulation positive du miR-675 sur sa cible (fig 35).

Au cours de ma thèse, j'ai donc identifié trois cibles du microARN généré par H19 dans les cellules cancéreuses mammaires. La régulation d'expression de ces gènes par le miR-675 favorise la prolifération, la migration et la survie cellulaire. J'ai donc mis en évidence l'un des mécanismes permettant à H19 de promouvoir la progression tumorale.

Par ailleurs, nous montrons qu'une surexpression du miR-675, H19-indépendant, favorise la tumorigenèse et les capacités métastatiques des cellules cancéreuses mammaires lorsqu'elles sont injectées en souris immunodéficientes SCID. De plus, dans les tumeurs mammaires, ce microARN est fréquemment surexprimé et toutes les cibles déjà identifiées, dans tous les cancers, permettent la progression tumorale (Tableau 6 ; Zhai et al., 2015). Ce microARN pourrait donc faire l'objet de cibles thérapeutiques. Plusieurs études ont déjà testé les effets de microARN synthétiques ou d'inhibiteurs spécifiques sur la tumorigenèse (Naidu et al., 2015). La première étude permettait la ré-expression du miR-34a dans des modèles de souris porteuses de cancer du poumon ; une diminution du volume tumoral était observée (Xue et al., 2014). Vinall et collaborateurs ont testé les effets du miR-34a mimic administré en combinaison avec le cisplatine sur des cellules cancéreuses de vessie (Vinall et al., 2012). Ils ont mis en évidence une levée de la résistance à la drogue chimiothérapeutique lorsque les cellules sont traitées avec le microARN synthétique. Ensuite, d'autres études permettant l'expression d'autres microARN (let-7, miR-200c, miR520e, miR-375, miR-217) dans des souris xénotransplantées montrent également une diminution significative du volume tumoral (Kasinski, et al., 2014 ; Cortez et al., 2014 ; Zhang et al., 2012b ; He et al., 2012 ; Zhao et al., 2010b). Par ailleurs, quelques travaux ont étudié l'effet d'inhibiteur de microARN dans certains modèles tumoraux. Par exemple, en 2013, Sicard et collaborateurs ont testé un inhibiteur du miR-21 en combinaison avec la gemcitabine sur les adénocarcinomes ductaux du pancréas ; un bénéfice thérapeutique a été observé (Sicard et al., 2013). De plus, une répression de l'expression du miR-20b dans les cellules cancéreuses mammaires réprime la croissance tumorale *in vivo* (Zhou et al., 2014a).

Ces différentes études nous montrent l'importance d'identifier l'ensemble des microARN impliqués dans la tumorigenèse ainsi que leur fonction afin de mettre en évidence de nouvelles cibles thérapeutiques. Avant d'utiliser ces molécules en clinique, de nombreuses mises au point sont encore nécessaires. En effet, ces microARN synthétiques ou inhibiteurs doivent être stables dans l'organisme, biodisponibles et facilement délivrables. A l'heure actuelle, un essai clinique de phase I, se terminant en Décembre 2015, teste un liposome contenant le miR-34a synthétique. Au regard des données acquises concernant le miR-675, il serait donc

envisageable de tester les effets d'inhibiteurs spécifiques du miR-675 en association avec les thérapies conventionnelles sur des tumeurs mammaires xénotransplantées.

➤ ***H19* favorise l'instabilité génétique et la progression tumorale**

Toujours dans l'optique d'identifier le mécanisme d'action d'*H19* lors de la tumorigénèse de la glande mammaire, j'ai mis en évidence que *H19* est capable de se fixer sur une protéine impliquée dans la réparation des dommages à l'ADN. En effet, dans les cellules cancéreuses mammaires, une association entre l'ARN *H19* et la protéine P53 a été identifiée. De plus, cette association empêche la fonction de P53 lors de dommages à l'ADN. P53 est un facteur de transcription continuellement exprimé et dégradé par le protéasome dans la cellule. Lors de dommages à l'ADN, les cassures simples et doubles brins sont reconnues par plusieurs facteurs qui vont activer la voie ATM/ATR. Ces protéines permettent alors la phosphorylation de P53 sur la sérine 15 ; cette modification post-traductionnelle permet la dissociation du complexe P53/MDM2 et le transport de P53 dans le noyau où il va activer l'expression de plusieurs gènes cibles dont *P21* et *Bax* (Freeman & Espinosa, 2013 ; Shiloh et al., 2013). L'activation de ces gènes cibles est fonction du taux de cassures simples et doubles brins observé dans la cellule. En effet, lors de dommages facilement réparables, P53 va activer l'expression de *P21* qui permettra alors un arrêt du cycle cellulaire. Lors de dommages irréversibles, P53 activera l'expression du gène *Bax* et permettra alors la mort de la cellule par apoptose. J'ai mis en évidence, qu'en présence d'*H19* et lors de dommages à l'ADN, l'expression de ces deux gènes est diminuée. De plus, les cellules possédant des cassures simples ou doubles brins et surexprimant *H19* progressent plus facilement dans le cycle cellulaire que les cellules contrôles. *H19* favorise donc la prolifération des cellules tumorales porteuses de dommages à l'ADN. Par immunofluorescence, nous montrons que ***H19* empêche la fonction de P53 en piégeant cette protéine dans le cytosol.**

De plus, ***H19* favorise l'association P53/MDM2** dans les cellules cancéreuses mammaires. La protéine MDM2 est une ubiquitine ligase permettant la dégradation de P53 par le protéasome de façon continue. En effet, en condition normale, dès sa synthèse, P53 est directement pris en charge par la protéine MDM2 qui permettra alors sa dégradation. Ce mécanisme de régulation de P53 peut également faire intervenir une autre protéine, de structure homologue à MDM2, MDMX ou MDM4 (Toledo et al., 2007). Cette dernière ne possède pas de fonctions ubiquitines ligases, ainsi à elle seule, elle ne permet pas la dégradation de P53. Cependant, un KO MDM2 et MDM4 est léthal chez la souris (Toledo et al., 2007). Ces protéines sont donc essentielles et complémentaires à la régulation de P53 en condition normale. En effet, en 2006, Toledo et collaborateurs ont mis en évidence que la protéine MDM2 est essentielle à la dégradation de P53 tandis que la protéine MDM4 permet

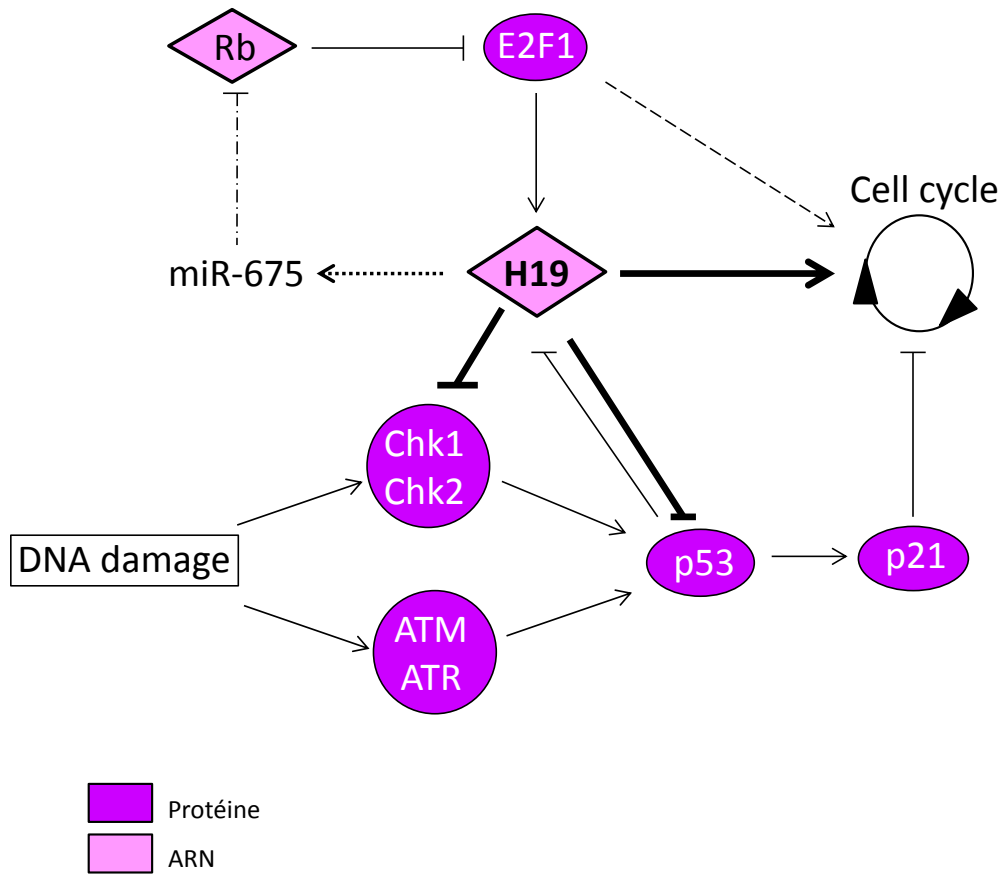


Figure 36. L'ARN *H19* et le cycle cellulaire. L'ARN *H19* empêche l'expression de la protéine Rb (un régulateur négatif du cycle cellulaire) par l'intermédiaire de son microARN (Tsang et al., 2010). Lors de dommages à l'ADN (cassures simples et/ou doubles brins), l'ARN *H19* empêche l'activation des protéines Chk1, Chk2 et P53. Ce mécanisme permet la prolifération des cellules porteuses de nouvelles mutations.

de réguler l'activité de P53 (Toledo et al., 2006). Les protéines MDM2 et MDM4 ont des structures homologues, il serait donc intéressant de déterminer si *H19* peut se fixer sur les protéines MDM2 et MDM4 pour réguler la dégradation et la fonction de P53. Dans ce cas, l'ARN *H19* exercerait une fonction « d'ARN *scaffold* ».

L'ARN *H19* est également capable de réguler l'action de la protéine P53 par l'intermédiaire de son microARN (Liu et al., 2015). En effet, dans les cancers de la vessie, l'expression ectopique du miR-675 est associée à une diminution d'expression de la protéine P53. Dans ces cellules, une augmentation de la prolifération ainsi qu'une diminution d'apoptose a été constatée. Cependant, aucune interaction directe entre le miR-675 et l'ARNm P53 n'a été identifiée ; il existe donc une protéine intermédiaire impliquée dans la régulation de P53 et régulée par le miR-675. ***Ainsi, H19 possède plusieurs mécanismes permettant la répression d'expression et de fonction de P53 dans les cellules cancéreuses.***

Dans les cellules cancéreuses mammaires, nous avons également mis en évidence que l'ARN *H19* est associé à la protéine BRCA1 (données non montrées). Cette protéine est également impliquée dans la réparation des dommages à l'ADN. En effet, de la même manière que P53, elle est activée par la voie ATM/ATR lors de cassures simples et doubles brins. Lorsqu'elle est activée, la protéine BRCA1, associée à la protéine RAD51, se fixe sur l'ADN et permet la réparation des cassures par recombinaison non homologue. Cette protéine possède donc une fonction nucléaire, nous pouvons donc supposer que *H19*, en s'y fixant, empêche sa translocation dans le noyau et réprime sa fonction.

H19 possède donc un rôle dans l'instabilité génétique tumorale puisqu'il est capable de s'associer avec plusieurs facteurs impliqués dans les mécanismes de réparation de l'ADN. Ces associations peuvent être responsables de la transformation cellulaire nécessaire à l'apparition des cellules cancéreuses.

Par ailleurs, plusieurs études ont déjà mis en évidence la fonction de l'ARN *H19* dans la régulation du cycle cellulaire. En effet, nous avons montré que *H19*, régulé positivement par le facteur E2F1, favorise la transition G1/S du cycle cellulaire (Berteaux et al., 2005) (fig 36). Par l'intermédiaire de son microARN, *H19* régule négativement l'expression de la protéine Rb (Tsang et al., 2010) (fig 36). Ici, nous montrons que l'ARN *H19*, lors de dommages à l'ADN, empêche l'activation des protéines Chk1, Chk2 et P53 (fig 36). Ce mécanisme permet alors la progression du cycle cellulaire bien que la cellule soit porteuse de mutations. Néanmoins, la protéine P53 réprime l'expression d'*H19* dans les cellules cancéreuses mammaires (Dugimont et al., 1998). Il existe donc une boucle négative de rétrocontrôle entre l'ARN *H19* et la protéine P53. ***L'ARN H19, grâce à l'ensemble de ses mécanismes, possède donc un rôle clé dans la régulation du cycle cellulaire*** (fig 36).

Selon Hanahan et Weinberg, une cellule devient tumorale lorsqu'elle possède au minimum 8 altérations génétiques. Néanmoins, au sein d'une tumeur, toutes les cellules ne possèdent pas les mêmes altérations. L'ARN *H19*, par l'intermédiaire de ses nouveaux mécanismes, peut donc participer à cette hétérogénéité tumorale en favorisant la prolifération des cellules ayant de nouvelles mutations. Dans de nombreux cancers dont le cancer du sein, une mutation de P53 ou BRCA1 est associée à un événement oncogénique. Ici, nous mettons en évidence un nouveau mécanisme de régulation de P53 qui peut être associé à un événement oncogénique. Ce mode d'action peut également être responsable de la transformation de cellules cancéreuses en cellules cancéreuses plus agressives. Ce nouveau mode d'action n'est pas spécifique des cellules cancéreuses mammaires puisque dans les cancers gastriques, l'ARN *H19* se fixe également sur la protéine P53 (Yang et al., 2012). De plus, dans ces tumeurs, une forte expression d'*H19* est associée à une très faible expression de la protéine P53.

De plus, dans les cancers, *H19* est fréquemment surexprimé. Au regard du mécanisme mis ici en évidence, il serait intéressant de déterminer si ce mécanisme permet l'initiation du cancer (la transformation cellulaire) ou s'il permet d'augmenter l'agressivité tumorale. Nous pourrions ainsi identifier la fonction d'*H19* au cours de la progression tumorale.

Au cours de ma thèse, j'ai donc mis en évidence plusieurs modes d'action de l'ARN H19 dans les cellules cancéreuses mammaires (fig 37). J'ai identifié plusieurs cibles de son microARN et j'ai également mis en évidence que H19 se fixe sur la protéine P53 et réprime sa fonction. Il est à noter que l'ensemble des mécanismes décrits participe au phénotype oncogénique associé à l'ARN H19. En effet, tous ces modes d'action permettent la prolifération, la survie, la migration/invasion, la tumorigenèse et/ou l'instabilité génétique des cellules tumorales mammaires.

➤ ***H19 dans les cellules souches***

En collaboration avec l'équipe du Dr Bourette, située à l'Institut de Biologie de Lille, nous avons mis en évidence l'implication de l'ARN *H19* dans le maintien des cellules souches normales. Cette équipe a utilisé un marqueur unique (s-SHIP) pour isoler une population de cellules prostatiques normales. Elle a ensuite caractérisé cette population et a mis en évidence que l'expression des marqueurs de surface : CD44, CD166, TROP2, et des facteurs de transcription : Sox2, Oct4, shh est augmentée dans cette population cellulaire isolée. De plus, ces cellules forment plus de sphères que les cellules contrôles et sont plus résistantes aux drogues, notamment à l'arsenic. Ainsi, sur la base de l'expression du gène *s-SHIP*, cette équipe a isolé une population de cellules souches prostatiques normales. En collaboration avec eux, nous avons mis en évidence que l'expression du gène *H19* est augmentée dans cette population cellulaire. Par invalidation de *H19* par ARN interférant ou surexpression de ce

gène, nous montrons que ***H19 est impliqué dans la formation et le maintien des cellules souches prostatiques humaines normales.*** Chez le rat castré, nous avons déjà mis en évidence que *H19* favorise le développement de la prostate normale (Berteaux et al., 2004).

Zeira et collaborateurs ont également mis en évidence la fonction de *H19* dans les cellules souches embryonnaires humaines (Zeira et al., 2015). Ils remarquent également une diminution de l'expression des facteurs Oct4 et nanog lorsque les cellules sont traitées avec un shRNA-*H19*. De la même manière, une forte expression d'*H19* est associée à une augmentation de l'invasion cellulaire, de l'angiogenèse, du nombre de cellules souches et de la tumorigénicité des cellules de glioblastomes (Jiang et al., 2015). ***Ainsi, H19 est également impliqué dans la formation et le maintien des cellules souches dans d'autres tissus. Cependant, son mode d'action dans la régulation de cet état physiologique n'est pas encore décrit.***

Par ailleurs, nous avons mis en lumière une régulation étroite entre le gène *H19* et le facteur de transcription Sox2. En effet, dans des cellules humaines prostatiques ou mammaires, normales ou cancéreuses, nous montrons qu'une forte expression d'*H19* est associée à une forte expression du marqueur Sox2 (fig 7 de l'article 2, P209). De plus, il est déjà mis en évidence que le facteur de transcription Sox2 régule positivement l'expression du gène *H19* en se fixant sur l'ICR de l'allèle maternel (Zimmerman et al., 2013). Ici, nous montrons qu'une variation de l'expression d'*H19* induit une modification de l'expression de Sox2. Cependant, le mode d'action permettant à *H19* de réguler l'expression de Sox2 n'est pas encore défini. De même, le mode d'action permettant à *H19* d'induire la formation et le maintien des cellules souches n'est pas encore décrit. Plusieurs mécanismes peuvent être associés à *H19* dans la formation des cellules souches.

Récemment, il a été mis en évidence que les longs ARN non codants sont impliqués dans la quiescence, la survie mais aussi dans l'auto-renouvellement des cellules souches. En effet, le long ARN non codant intergénique RoR (*Regulator of Reprogramming*) a été identifié dans les cellules souches reprogrammées et possède plusieurs modes d'action lui permettant de maintenir cette population cellulaire au sein d'un tissu (Loewer et al., 2010). Ses modes d'action sont similaires à ceux déjà décrits pour l'ARN *H19*. En effet, le *linc-RoR* peut agir comme une éponge à microARN, notamment au miR-145, dans les cellules souches embryonnaires (Wang et al., 2013). Ce microARN est connu comme étant un régulateur négatif de l'expression des facteurs Sox2, Oct4 et Klf4 (Xu et al., 2009). Ce microARN est donc un répresseur de l'auto-renouvellement des cellules souches. Ainsi, cet ARN non codant favorise l'expression de ces marqueurs dans les cellules souches. Ce mode d'action a également été retrouvé dans les cellules souches cancéreuses de l'endomètre chez l'Homme

(Zhou et al., 2014b). Le *linc-RoR* est également une éponge au miR-205 dans les cellules cancéreuses mammaires (Hou et al., 2014). Cette fonction lui permet de promouvoir la transition épithélio-mésenchymateuse des cellules ; il permet donc la dissémination des cellules cancéreuses en augmentant leurs capacités métastatiques.

Par ailleurs, comme *H19*, *linc-RoR* est capable de se fixer sur la protéine P53 lors de dommages à l'ADN et de réprimer sa fonction (Zhang et al., 2013). De plus, il est démontré que l'inhibition de la protéine P53 permet l'auto-renouvellement des cellules souches mammaires humaines et leur division symétriques (Cicalese et al., 2009 ; Bonizzi et al., 2012). P53 empêche également la reprogrammation des cellules souches en interagissant avec la protéine HDAC1 (Ma et al., 2015). Cette association permet de réprimer l'expression des gènes *Oct4*, *Nanog* et *H19*. Ce mécanisme agit donc également en faveur de la formation et du maintien des cellules souches. L'ARN non codant *RoR* possède donc beaucoup de similarités de fonction avec l'ARN non codant *H19* (fig 24).

Les mécanismes identifiés pour *linc-RoR* dans les cellules souches peuvent être transposés à *H19*. En effet, il est démontré que *H19* est une éponge au microARN let-7 (Kallen et al., 2013). Ce microARN est décrit comme étant un régulateur négatif de l'expression des facteurs Sox2 et Oct4. Dans des cellules souches hématopoïétiques, une régulation négative de let-7 par le microARN lin-28 permet l'expression de HMG2A. Cette protéine permet à son tour l'expression des facteurs Sox2, Oct4 et par conséquent, l'auto-renouvellement des cellules souches (Copley et al, 2013). Cette régulation a également été retrouvée dans les cellules cancéreuses des voies aéro-digestives supérieures où let-7 régule négativement l'expression des gènes *ARID3B* et *HMG2A* (Chien et al., 2015). De plus, il est démontré que l'expression de let-7 dans des cellules cancéreuses mammaires diminue la capacité des cellules à former des sphères et la tumorigenèse des cellules (Yu et al., 2007). L'ensemble des données montre que let-7 est un répresseur de la formation et du maintien des cellules souches dans un tissu. L'ARN *H19* peut donc promouvoir la formation des cellules souches grâce à sa fonction d'éponge au microARN let-7. Ce mécanisme peut également être responsable de la forte expression de Sox2 lorsque *H19* est surexprimé. De plus, dans notre étude, nous montrons que *H19* se fixe sur la protéine P53 et réprime sa fonction, de la même manière que *linc-RoR* (Wang et al., 2013). Ce mécanisme peut également être responsable de la formation des cellules souches dans nos cellules.

Par ailleurs, une relation entre les expressions d'*H19* et de Sox2 a été mise en évidence. *H19* est capable de se fixer sur la protéine EZH2 et par modifications épigénétiques, de réguler l'expression de gènes cibles. En effet, une répression de let-7 favorise l'expression de la protéine EZH2. Cette dernière permet alors l'expression de gènes spécifiques des cellules souches prostatiques (Kong et al., 2012). *H19*, grâce à ses interactions avec let-7 et EZH2,

peut donc favoriser l'expression de gènes signatures des cellules souches, dont Sox2 par exemple. De plus, il est possible que certaines cibles du miR-675, pas encore identifiées, soient impliquées dans la formation ou le maintien des cellules souches dans un tissu.

Dans notre étude, la relation d'expression entre Sox2 et *H19* a été identifiée dans les cellules normales et cancéreuses de la prostate mais aussi de la glande mammaire. Il serait donc intéressant de déterminer les effets de *H19* sur les cellules souches cancéreuses de ces tissus. Les résultats préliminaires montrent une augmentation de la capacité à former des sphères lorsque *H19* est surexprimé dans les cellules (données non présentées). De plus, il augmente l'expression de gènes impliqués dans les cellules souches (Sox2, Oct4, nanog). Néanmoins, des analyses complémentaires sont nécessaires pour confirmer ces données et identifier le(s) mode(s) d'action de l'ARN *H19*.

Les mécanismes déjà mis en évidence pour *H19* dans les cellules cancéreuses, peuvent donc être transposés aux cellules souches normales et/ou cancéreuses.

➤ **91H : un lncRNA oncogène impliqué dans l'empreinte génomique**

Pour finir, au cours de ma thèse, j'ai également étudié les fonctions du long ARN non codant *91H* dans les cellules cancéreuses mammaires. Cet ARN non codant, transcrit au locus *H19/IGF2*, a été découvert, chez l'Homme en 2008 (Berteaux et al., 2008). Il est conservé chez la souris où il permet l'expression d'*IGF2* à partir d'un nouveau promoteur (Tran et al., 2012). Les longs ARN non codants antisens sont généralement transcrits à partir de locus soumis à l'empreinte génomique parentale et sont impliqués dans la régulation de cette empreinte (Sanli & Feil, 2015). Au locus *H19/IGF2*, la fonction de *91H* dans les cellules cancéreuses mammaires n'était pas définie. En étudiant les modifications épigénétiques au niveau du centre de l'empreinte (ICR) et du promoteur d'*H19* dans des cellules invalidées stablement pour *91H*, j'ai mis en évidence que **91H permet le maintien de l'empreinte génomique parentale au locus *H19/IGF2***. En effet, en absence de *91H*, la méthylation de l'ICR et du promoteur d'*H19* se répartit sur les deux allèles. Nous observons également une répartition de la fixation des protéines CTCF et c-myc sur les deux allèles en absence de *91H*. Ces deux mécanismes sont responsables de la perte d'expression des gènes *H19* et *IGF2* observée en absence de *91H*. Au regard de nos données et des connaissances déjà acquises sur le locus *H19/IGF2*, nous supposons que l'expression des gènes *H19* et *IGF2* observée en absence de *91H* correspond à un bruit de fond transcriptionnel (Fig 7 de l'article 3, voir P 243). Nous pensons que sur l'allèle maternel, *91H* empêche la méthylation de l'ICR et du promoteur d'*H19* en masquant les sites de méthylation.

De nombreux loci soumis à l'empreinte génomique sont régulés par des longs ARN non codants antisens (Sanli et Feil, 2015). Cependant, en général, ces lncRNA permettent le

recrutement de protéines impliquées dans la méthylation de l'ADN ou la chromatine, EZH2 ou encore G9a. En effet, les lncRNA *Kcnq1ot1* et *Air*, transcrits respectivement à partir des allèles paternels des loci *Kcnq1/Kcnq1ot1*, *Igf2r/Air* recrutent respectivement les protéines EZH2 et G9a. Ils permettent ensuite la méthylation de l'allèle paternel de leur loci d'origine et empêchent donc l'expression des gènes présents sur cet allèle. ***Ici, nous mettons en évidence un nouveau mécanisme d'action des lncRNA antisens pour réguler l'empreinte génomique. Nos données suggèrent que le lncRNA 91H empêche la méthylation sur l'allèle maternel en masquant les îlots CpG.***

Par ailleurs, au cours de cette étude, l'implication de *91H* au cours de la tumorigenèse de la glande mammaire a été définie. En effet, les capacités de prolifération, de migration/invasion et de croissance à basse densité et sans ancrage des cellules sont diminuées en absence d'expression de *91H*. De plus, les capacités tumorigènes des cellules cancéreuses mammaires invalidées pour *91H* diminuent lorsqu'elles sont xénotreffées en souris immunodéficientes. Ainsi, l'ensemble de ces résultats montre que ***91H* potentialise les capacités tumorigènes des cellules cancéreuses mammaires** ; il agit donc comme un oncogène dans ce type cellulaire. Cependant, nous n'avons pas identifié le(s) mode(s) d'action de *91H* pour favoriser la tumorigenèse des cellules. Mais nous montrons que *91H* est un régulateur positif de l'expression d'*H19* et d'*IGF2*. Or, nous avons déjà établi que *H19* augmente les capacités tumorigènes des cellules cancéreuses mammaires (Lottin et al., 2002, Berteaux et al., 2005, Vennin et al., 2015). De plus, des animaux transgéniques surexprimant *IGF2* ont des risques augmentés de développer des tumeurs mammaires et pulmonaires (Bates et al., 1995 ; Moorehead et al., 2003). Dans les cancers de l'endomètre, une forte expression d'*IGF2* est même associée à un mauvais pronostic (Dong et al., 2015). Ainsi, dans notre étude, nous ne pouvons discerner les phénotypes associés à l'inactivation de *91H* et les phénotypes associés à la répression d'*H19* ou d'*IGF2*. Cependant, nous pouvons imaginer que les ARN *H19* et *91H* agissent de manière coopérative ou synergique pour augmenter les capacités tumorigènes des cellules cancéreuses mammaires.

➤ Perspectives

Le gène *H19* est transcrit en un long ARN non codant exprimé pendant le développement embryonnaire et réprimé à la naissance excepté dans certains tissus, notamment la glande mammaire et l'utérus. Dans ces tissus, son expression est souvent associée aux variations hormonales de l'individu. Plusieurs études ont mis en évidence qu'une expression aberrante du gène *H19* favorise la progression tumorale dans de nombreux cancers : glande mammaire, utérus, vessie, foie, colon par exemple. Nous avons mis en évidence plusieurs modes d'action associés à l'ARN *H19* ou son microARN dans le cancer du sein. Cependant, ces mécanismes ont été identifiés dans les cellules tumorales. Afin d'identifier de nouvelles thérapeutiques, il serait intéressant de déterminer la fonction de l'ARN *H19* au cours de la transformation cellulaire. *Cet ARN est-il impliqué dès l'initiation du cancer ou ses fonctions oncogéniques apparaissent lors de la progression tumorale ?*

A l'heure actuelle, il existe un modèle d'évolution tumorale selon lequel les cellules souches, après acquisition de mutations, permettraient l'apparition de la masse tumorale et seraient également responsables des récurrences observées après traitements.

Dans ce manuscrit, nous mettons en évidence que l'ARN *H19* est impliqué dans la formation et le maintien des cellules souches prostatiques normales humaines. Un des objectifs du groupe à l'heure actuelle est de déterminer la fonction de l'ARN *H19* dans la biologie des cellules souches normales et cancéreuses de la glande mammaire. Les premiers tests fonctionnels réalisés sont plutôt encourageants et concluants. En effet, il semblerait que la surexpression du gène *H19* dans les cellules cancéreuses mammaires favorisent les capacités des cellules à former des sphères (une caractéristique de cellules souches) et augmentent l'expression de facteurs de transcription associés aux cellules souches (Sox2, Oct4, nanog) (données non montrées). Pour renforcer ces résultats, nous injecterons en souris immunodéficientes SCID des cellules surexprimant *H19* en dilution limite. Nous nous attendons à observer l'apparition de tumeur à partir de ces cellules même lorsqu'elles sont injectées en petites quantités (500 cellules par exemple). Après avoir confirmé l'action de *H19* dans la biologie des cellules souches cancéreuses (CSC) de la glande mammaire, l'équipe identifiera le(s) mode(s) d'action permettant à l'ARN *H19* de promouvoir la formation et le maintien de ces cellules.

Nous avons déjà mis en évidence que *H19* empêche la fonction de la protéine P53 dans les cellules cancéreuses mammaires. La signalisation de la protéine P53 est impliquée dans la régulation des cellules souches. En effet, plusieurs études ont mis en évidence que l'expression de cette protéine empêche la formation et le maintien de cette population cellulaire. Nous pouvons donc supposer que la régulation de P53 par l'ARN *H19* est l'un des mécanismes permettant à cet ARN de promouvoir les cellules souches cancéreuses.

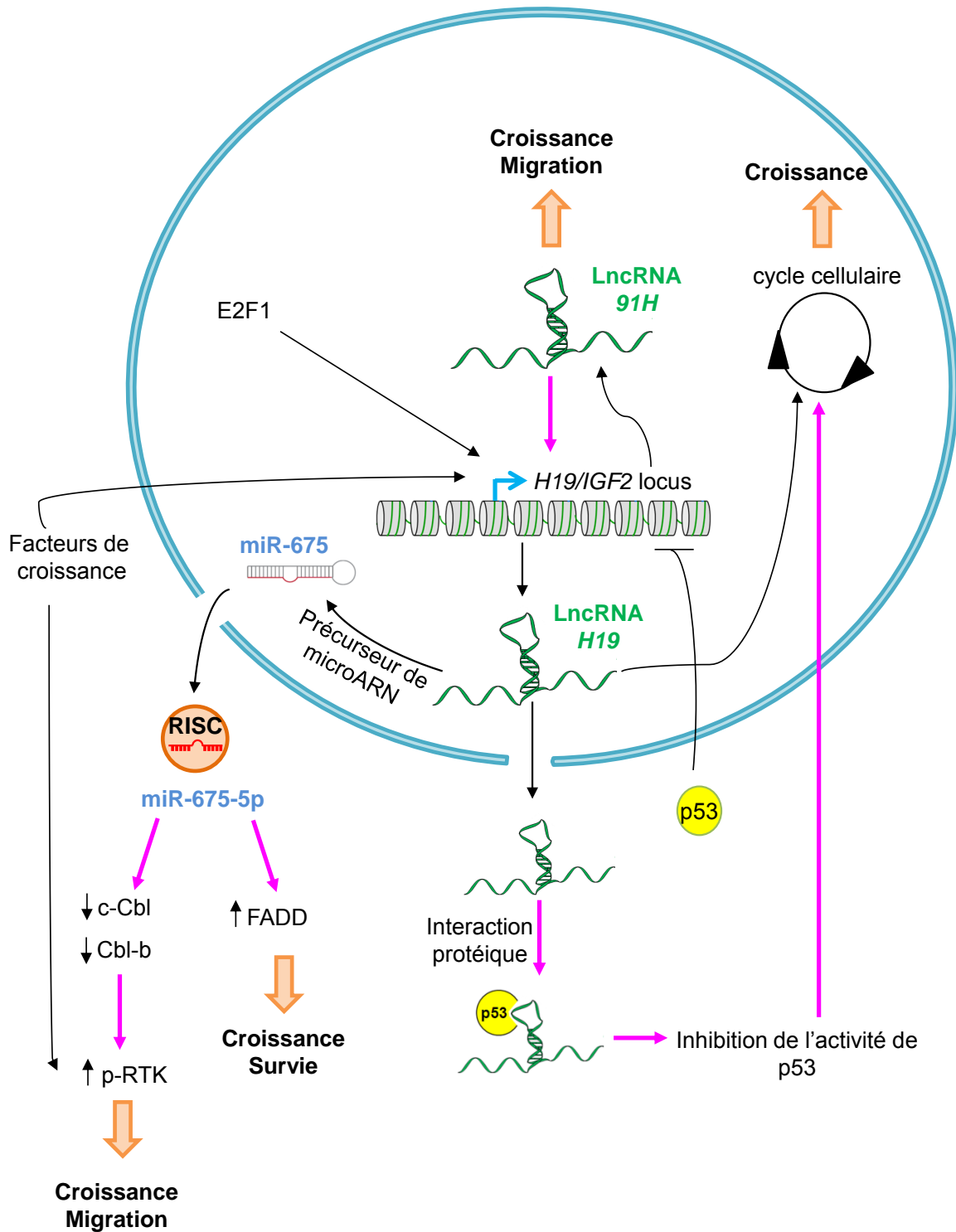


Figure 37. Rôle des ARN non codants H19 et 91H dans une cellule cancéreuse mammaire. Trois cibles du microARN généré par l'ARN H19 ont été identifiées. Cette régulation mise en évidence permet les capacités tumorigènes des cellules cancéreuses mammaires. L'ARN H19 favorise l'instabilité génétique en empêchant les fonctions de la protéine P53. L'ARN non codant 91H permet le maintien de l'empreinte génomique au locus H19/IGF2 et favorise la tumorigénèse des cellules cancéreuses mammaires. En rose : les nouveaux modes d'action. En noir : les fonctions déjà identifiées. (adapté de Angrand et al., 2015).

Pour vérifier cette hypothèse, l'ensemble des analyses sera réalisée dans des cellules cancéreuses mammaires avec différents statut P53. D'autres mécanismes déjà décrits pour l'ARN *H19* peuvent également réguler les CSC : précurseur de microARN, éponge à microARN, interaction avec la protéine EZH2. L'implication de ces mécanismes dans la biologie des CSC de la glande mammaire sera également déterminée.

A l'heure actuelle, l'équipe a mis en place un modèle de souris transgénique dite C3(1)Tag. Ce modèle a été obtenu après microinjection d'un transgène permettant l'expression de l'oncogène large T de SV40 sous le contrôle du promoteur C3(1) dans des ovocytes fertilisés de souris FVB/N. Ainsi, le transgène est exprimé uniquement dans la prostate et dans la glande mammaire. Chez les souris femelles, une hyperplasie spontanée de la glande mammaire apparaît 8 semaines après la naissance (Maroulakou et al., 1994). Toutes les souris femelles développent des adénocarcinomes mammaires *in situ* à l'âge de 12 semaines, et invasif à l'âge de 16 semaines. Pour finir, 20 semaines après la naissance, soit environ 5 mois, les souris femelles présentes des métastases. Chez les souris mâles, un adénocarcinome prostatique est observé après l'âge de 8 mois. Ces tumeurs apparaissent donc plus tardivement comparées aux tumeurs mammaires. Les gènes transcrits au locus *H19/IGF2*, notamment *H19*, *IGF2* et *91H*, étant surexprimés dans les adénocarcinomes mammaires ; nous utiliserons ce modèle de souris afin d'identifier leur fonction au cours de la tumorigenèse. Dans un premier temps, nous quantifierons leur niveau d'expression à différents stades de la tumorigenèse : hyperplasique, adénocarcinome et dans les métastases par exemple. Nous pourrions ainsi déterminer si la surexpression de ces gènes intervient de manière séquentielle et dès les premiers stades de la tumorigenèse. Selon les résultats obtenus, les mécanismes régulant l'expression du locus *H19/IGF2* lors de la progression tumorale seront alors étudiés. Cette analyse nous permettra une meilleure compréhension des mécanismes impliqués dans l'initiation et la progression des adénocarcinomes mammaires.

Enfin, dans ce manuscrit, nous avons mis en lumière un nouveau mécanisme de recrutement et de régulation de l'action des microARN dans les cellules cancéreuses mammaires. Nous montrons que la protéine FXR1, associée à l'ARNm FADD, favorise le recrutement du microARN généré par *H19* et permet une régulation positive de ce miR sur cette cible. Afin d'identifier de nouvelles cibles thérapeutiques, nous pourrions essayer de généraliser ce surprenant mécanisme dans plusieurs types cellulaires et avec d'autres cibles.

Par ailleurs, nous avons mis en évidence dans ce manuscrit des mécanismes associés à l'ARN *H19* et d'autres associés au miR-675. Afin de discriminer les effets propres à l'ARN *H19* de ceux associés au miR-675, il serait intéressant de générer un ARN *H19* muté dans le miR-675 en utilisant la technologie des TALENs par exemple. Ainsi, nous aurions des cellules

surexprimant *H19* et son microARN, des cellules surexprimant uniquement le microARN et des cellules surexprimant *H19* mais pas le microARN. L'ensemble des études seraient alors réalisées dans ces trois lignées. Nous pourrions alors déterminer l'impact de l'ARN *H19* et/ou de son microARN sur la biologie des cellules cancéreuses non souches et souches de la glande mammaire. Nous nous attendons à observer un effet coopératif entre l'ARN et son miR dans les cellules cancéreuses afin de promouvoir la tumorigenèse.

L'étude des ARN non codants est un domaine de recherche complexe et très varié. En effet, ces molécules, contrairement aux protéines, sont dotées de fonctions très diverses ; elles possèdent plusieurs modes d'action différents mais permettant d'aboutir à un même phénotype cellulaire. Afin de pouvoir cibler ces molécules lors de traitements thérapeutiques, il est donc nécessaire d'identifier l'ensemble de leurs mécanismes pour mieux comprendre leur rôle dans les cellules normales et cancéreuses.

Pour conclure, j'ai mis en lumière plusieurs mécanismes associés à l'ARN H19 dans les cellules tumorales (Fig 37). Ces différents mécanismes (précurseur de microARN, interagant protéique) permettent la prolifération, la migration/invasion, l'instabilité génétique et la progression tumorale des cellules cancéreuses mammaires. Ils sont donc responsables du phénotype oncogène de l'ARN H19 dans ce tissu. De plus, j'ai mis en évidence l'implication de l'ARN H19 dans les cellules souches prostatiques humaines normales. Les cellules souches pouvant être impliquées dans l'initiation du cancer ou sa récurrence, H19 peut donc également avoir un rôle dans ces étapes de la tumorigenèse. Pour finir, j'ai identifié le rôle du long ARN non codant 91H dans la régulation de l'expression des gènes du H19/IGF2. J'ai également mis en évidence que cet ARN diminue les capacités tumorigènes des cellules cancéreuses mammaires lorsqu'il est invalidé. Il possède donc des fonctions d'oncogène.

Les mécanismes associés à H19 mis en évidence dans ce manuscrit permettent de mieux comprendre la fonction de cet ARN lors de la tumorigenèse. Ces mécanismes peuvent faire l'objet de nouvelles cibles thérapeutiques dans le traitement du cancer.

BIBLIOGRAPHIE

Abi Habib, W., Azzi, S., Brioude, F., Steunou, V., Thibaud, N., Das Neves, C., Le Jule, M., Chantot-Bastaraud, S., Keren, B., Lyonnet, S., et al. (2014). Extensive investigation of the IGF2/H19 imprinting control region reveals novel OCT4/SOX2 binding site defects associated with specific methylation patterns in Beckwith-Wiedemann syndrome. *Hum. Mol. Genet.* 23, 5763–5773.

Adriaenssens, E., Dumont, L., Lottin, S., Bolle, D., Leprêtre, A., Delobelle, A., Bouali, F., Dugimont, T., Coll, J., and Cury, J.J. (1998). H19 overexpression in breast adenocarcinoma stromal cells is associated with tumor values and steroid receptor status but independent of p53 and Ki-67 expression. *Am. J. Pathol.* 153, 1597–1607.

Adriaenssens, E., Lottin, S., Dugimont, T., Fauquette, W., Coll, J., Dupouy, J.P., Boilly, B., and Cury, J.J. (1999). Steroid hormones modulate H19 gene expression in both mammary gland and uterus. *Oncogene* 18, 4460–4473.

Adriaenssens, E., Lottin, S., Berteaux, N., Hornez, L., Fauquette, W., Fafeur, V., Peyrat, J.-P., Le Bourhis, X., Hondermarck, H., Coll, J., et al. (2002). Cross-talk between mesenchyme and epithelium increases H19 gene expression during scattering and morphogenesis of epithelial cells. *Exp. Cell Res.* 275, 215–229.

Adriaenssens, E., Vanhecke, E., Saule, P., Mougel, A., Page, A., Romon, R., Nurcombe, V., Le Bourhis, X., and Hondermarck, H. (2008). Nerve growth factor is a potential therapeutic target in breast cancer. *Cancer Res.* 68, 346–351.

Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., and Clarke, M.F. (2003). Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. U.S.A.* 100, 3983–3988.

Angrand, P.-O., Vennin, C., Le Bourhis, X., and Adriaenssens, E. (2015). The role of long non-coding RNAs in genome formatting and expression. *Front. Genet.* 165.

Apostolou, P., and Fostira, F. (2013). Hereditary breast cancer: the era of new susceptibility genes. *Biomed Res Int* 2013, 747318.

Ariel, I., Lustig, O., Schneider, T., Pizov, G., Sappir, M., De-Groot, N., and Hochberg, A. (1995). The imprinted H19 gene as a tumor marker in bladder carcinoma. *Urology* 45, 335–338.

Ariel, I., Ayesch, S., Perlman, E.J., Pizov, G., Tanos, V., Schneider, T., Erdmann, V.A., Podeh, D., Komitowski, D., Quasem, A.S., et al. (1997a). The product of the imprinted H19 gene is an oncofetal RNA. *MP, Mol. Pathol.* 50, 34–44.

Ariel, I., Weinstein, D., Voutilainen, R., Schneider, T., Lustig-Yariv, O., de Groot, N., and Hochberg, A. (1997b). Genomic imprinting and the endometrial cycle. The expression of the imprinted gene H19 in the human female reproductive organs. *Diagn. Mol. Pathol.* *6*, 17–25.

Ariel, I., Miao, H.Q., Ji, X.R., Schneider, T., Roll, D., de Groot, N., Hochberg, A., and Ayeshe, S. (1998). Imprinted H19 oncofetal RNA is a candidate tumour marker for hepatocellular carcinoma. *MP, Mol. Pathol.* *51*, 21–25.

Ariel, I., de Groot, N., and Hochberg, A. (2000a). Imprinted H19 gene expression in embryogenesis and human cancer: the oncofetal connection. *Am. J. Med. Genet.* *91*, 46–50.

Ariel, I., Sughayer, M., Fellig, Y., Pizov, G., Ayeshe, S., Podeh, D., Libdeh, B.A., Levy, C., Birman, T., Tykocinski, M.L., et al. (2000b). The imprinted H19 gene is a marker of early recurrence in human bladder carcinoma. *MP, Mol. Pathol.* *53*, 320–323.

Arnold, A., and Papanikolaou, A. (2005). Cyclin D1 in breast cancer pathogenesis. *J. Clin. Oncol.* *23*, 4215–4224.

Artandi, S.E., and DePinho, R.A. (2010). Telomeres and telomerase in cancer. *Carcinogenesis* *31*, 9–18.

Askarian-Amiri, M.E., Crawford, J., French, J.D., Smart, C.E., Smith, M.A., Clark, M.B., Ru, K., Mercer, T.R., Thompson, E.R., Lakhani, S.R., et al. (2011). SNORD-host RNA Zfas1 is a regulator of mammary development and a potential marker for breast cancer. *RNA* *17*, 878–891.

Aubert, L., Guilbert, M., Corbet, C., Génot, E., Adriaenssens, E., Chassat, T., Bertucci, F., Daubon, T., Magné, N., Le Bourhis, X., et al. (2015). NGF-induced TrkA/CD44 association is involved in tumor aggressiveness and resistance to lestaurtinib. *Oncotarget* *6*, 9807–9819.

Auyeung, V.C., Ulitsky, I., McGeary, S.E., and Bartel, D.P. (2013). Beyond secondary structure: primary-sequence determinants license pri-miRNA hairpins for processing. *Cell* *152*, 844–858.

Ayala, Y.M., Pagani, F., and Baralle, F.E. (2006). TDP43 depletion rescues aberrant CFTR exon 9 skipping. *FEBS Lett.* *580*, 1339–1344.

Azuma-Mukai, A., Oguri, H., Mituyama, T., Qian, Z.R., Asai, K., Siomi, H., and Siomi, M.C. (2008). Characterization of endogenous human Argonautes and their miRNA partners in RNA silencing. *Proc. Natl. Acad. Sci. U.S.A.* *105*, 7964–7969.

- Babiarz, J.E.**, Ruby, J.G., Wang, Y., Bartel, D.P., and Blelloch, R. (2008). Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev.* *22*, 2773–2785.
- Bai, L.**, and Rohrschneider, L.R. (2010). s-SHIP promoter expression marks activated stem cells in developing mouse mammary tissue. *Genes Dev.* *24*, 1882–1892.
- Bardoni, B.**, Schenck, A., and Mandel, J.-L. (2001). The Fragile X mental retardation protein. *Brain Research Bulletin* *56*, 375–382.
- Barlow, D.P.**, and Bartolomei, M.S. (2014). Genomic imprinting in mammals. *Cold Spring Harb Perspect Biol* *6*.
- Barlow, D.P.**, Stöger, 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*, 84–87.
- 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*, 5330–5337.
- Bartke, A.**, Chandrashekar, V., Bailey, B., Zaczek, D., and Turyn, D. (2002). Consequences of growth hormone (GH) overexpression and GH resistance. *Neuropeptides* *36*, 201–208.
- Bartolomei, M.S.**, Zemel, S., and Tilghman, S.M. (1991). Parental imprinting of the mouse H19 gene. *Nature* *351*, 153–155.
- Baselga, J.** (2010). Treatment of HER2-overexpressing breast cancer. *Ann. Oncol.* *21 Suppl 7*, vii36–vii40.
- Baselga, J.**, Gelmon, K.A., Verma, S., Wardley, A., Conte, P., Miles, D., Bianchi, G., Cortes, J., McNally, V.A., Ross, G.A., et al. (2010). Phase II trial of pertuzumab and trastuzumab in patients with human epidermal growth factor receptor 2-positive metastatic breast cancer that progressed during prior trastuzumab therapy. *J. Clin. Oncol.* *28*, 1138–1144.
- Bates, P.**, Fisher, R., Ward, A., Richardson, L., Hill, D.J., and Graham, C.F. (1995). Mammary cancer in transgenic mice expressing insulin-like growth factor II (IGF-II). *Br. J. Cancer* *72*, 1189–1193.
- Bauderlique-Le Roy, H.**, Vennin, C., Brocqueville, G., Spruyt, N., Adriaenssens, E., and Bourette, R.P. (2015). Enrichment of Human Stem-Like Prostate Cells with s-SHIP Promoter Activity Uncovers a Role in Stemness for the Long Noncoding RNA H19. *Stem Cells Dev.* *24*, 1252–1262.

- Behm-Ansmant, I.**, Rehwinkel, J., and Izaurralde, E. (2006). MicroRNAs silence gene expression by repressing protein expression and/or by promoting mRNA decay. *Cold Spring Harb. Symp. Quant. Biol.* 71, 523–530.
- Berezikov, E.**, Chung, W.-J., Willis, J., Cuppen, E., and Lai, E.C. (2007). Mammalian mirtron genes. *Mol. Cell* 28, 328–336.
- Bergmann, J.H.**, and Spector, D.L. (2014). Long non-coding RNAs: Modulators of nuclear structure and function. *Curr Opin Cell Biol* 0, 10–18.
- Bergman, D.**, Halje, M., Nordin, M., and Engström, W. (2013). Insulin-like growth factor 2 in development and disease: a mini-review. *Gerontology* 59, 240–249.
- Bernard, D.**, Prasanth, K.V., Tripathi, V., Colasse, S., Nakamura, T., Xuan, Z., Zhang, M.Q., Sedel, F., Jourden, L., Couplier, F., et al. (2010). A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. *EMBO J.* 29, 3082–3093.
- 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, 363–366.
- Bernstein, E.**, Kim, S.Y., Carmell, M.A., Murchison, E.P., Alcorn, H., Li, M.Z., Mills, A.A., Elledge, S.J., Anderson, K.V., and Hannon, G.J. (2003). Dicer is essential for mouse development. *Nat. Genet.* 35, 215–217.
- Berteaux, N.**, Lottin, S., Adriaenssens, E., Van Coppenolle, F., Van Coppenolle, F., Leroy, X., Coll, J., Dugimont, T., and Cury, J.-J. (2004). Hormonal regulation of H19 gene expression in prostate epithelial cells. *J. Endocrinol.* 183, 69–78.
- Berteaux, N.**, Lottin, S., Monté, D., Pinte, S., Quatannens, B., Coll, J., Hondermarck, H., Cury, 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, 29625–29636.
- Berteaux, N.**, Aptel, N., Cathala, G., Genton, C., Coll, J., Daccache, A., Spruyt, N., Hondermarck, H., Dugimont, T., Cury, J.-J., et al. (2008). A novel H19 antisense RNA overexpressed in breast cancer contributes to paternal IGF2 expression. *Mol. Cell. Biol.* 28, 6731–6745.
- Berteaux, N.**, Spruyt, N., and Adriaenssens, E. (2010). Noncoding RNAs at H19/IGF2 Locus: Role in Imprinting, Gene Expression, and Associated Pathologies. *Rna Technologies & Their Applications* 419.
- Bestor, T.H.** (1988). Cloning of a mammalian DNA methyltransferase. *Gene* 74, 9–12.

- Bhattacharyya, S.N.**, Habermacher, R., Martine, U., Closs, E.I., and Filipowicz, W. (2006). Relief of microRNA-mediated translational repression in human cells subjected to stress. *Cell* *125*, 1111–1124.
- Biran, H.**, Ariel, I., de Groot, N., Shani, A., and Hochberg, A. (1994). Human imprinted genes as oncodevelopmental markers. *Tumour Biol.* *15*, 123–134.
- Blaszczyk, J.**, Tropea, J.E., Bubunenko, M., Routzahn, K.M., Waugh, D.S., Court, D.L., and Ji, X. (2001). Crystallographic and modeling studies of RNase III suggest a mechanism for double-stranded RNA cleavage. *Structure* *9*, 1225–1236.
- Bohnsack, M.T.**, Czaplinski, K., and Gorlich, D. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* *10*, 185–191.
- Boldin, M.P.**, Taganov, K.D., Rao, D.S., Yang, L., Zhao, J.L., Kalwani, M., Garcia-Flores, Y., Luong, M., Devrekanli, A., Xu, J., et al. (2011). miR-146a is a significant brake on autoimmunity, myeloproliferation, and cancer in mice. *J. Exp. Med.* *208*, 1189–1201.
- Bonizzi, G.**, Cicalese, A., Insinga, A., and Pelicci, P.G. (2012). The emerging role of p53 in stem cells. *Trends Mol Med* *18*, 6–12.
- Bonnet, D.**, and Dick, J.E. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat. Med.* *3*, 730–737.
- Bosco, E.E.**, and Knudsen, E.S. (2007). RB in breast cancer: at the crossroads of tumorigenesis and treatment. *Cell Cycle* *6*, 667–671.
- Bourhis, X.L.**, Toillon, R.A., Boilly, B., and Hondermarck, H. (2000). Autocrine and paracrine growth inhibitors of breast cancer cells. *Breast Cancer Res. Treat.* *60*, 251–258.
- Bowman, B.M.**, Sebolt, K.A., Hoff, B.A., Boes, J.L., Daniels, D.L., Heist, K.A., Galbán, C.J., Patel, R.M., Zhang, J., Beer, D.G., et al. (2015). Phosphorylation of FADD by the kinase CK1 α promotes KRASG12D-induced lung cancer. *Sci Signal* *8*, ra9.
- 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.
- Brockdorff, N.** (2013). Noncoding RNA and Polycomb recruitment. *RNA* *19*, 429–442.
- Brockdorff, N.**, Ashworth, A., Kay, G.F., McCabe, V.M., Norris, D.P., Cooper, P.J., Swift, S., and Rastan, S. (1992). The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell* *71*, 515–526.

- Brown, C.J.**, Hendrich, B.D., Rupert, J.L., Lafrenière, R.G., Xing, Y., Lawrence, J., and Willard, H.F. (1992). The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* 71, 527–542.
- Cai, X.**, and Cullen, B.R. (2007). The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA* 13, 313–316.
- Calin, G.A.**, Dumitru, C.D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., et al. (2002). Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15524–15529.
- Caretti, G.**, Schiltz, R.L., Dilworth, F.J., Di Padova, M., Zhao, P., Ogryzko, V., Fuller-Pace, F.V., Hoffman, E.P., Tapscott, S.J., and Sartorelli, V. (2006). The RNA helicases p68/p72 and the noncoding RNA SRA are coregulators of MyoD and skeletal muscle differentiation. *Dev. Cell* 11, 547–560.
- Carmeliet, P.** (2003). Angiogenesis in health and disease. *Nat. Med.* 9, 653–660.
- Casola, S.**, Pedone, P.V., Cavazzana, A.O., Basso, G., Luksch, R., d' Amore, E.S., Carli, M., Bruni, C.B., and Riccio, A. (1997). Expression and parental imprinting of the H19 gene in human rhabdomyosarcoma. *Oncogene* 14, 1503–1510.
- Cesana, M.**, Cacchiarelli, D., Legnini, I., Santini, T., Sthandier, O., Chinappi, M., Tramontano, A., and Bozzoni, I. (2011). A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* 147, 358–369.
- Chao, C.C.-K.** (2015). Mechanisms of p53 degradation. *Clinica Chimica Acta* 438, 139–147.
- Cheloufi, S.**, Dos Santos, C.O., Chong, M.M.W., and Hannon, G.J. (2010). A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* 465, 584–589.
- Chen, L.-L.**, and Carmichael, G.G. (2009). Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol. Cell* 35, 467–478.
- Chen, Y.**, and Olopade, O.I. (2008). MYC in breast tumor progression. *Expert Rev Anticancer Ther* 8, 1689–1698.
- Chendrimada, T.P.**, Gregory, R.I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., and Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* 436, 740–744.

- Chien, C.-S.**, Wang, M.-L., Chu, P.-Y., Chang, Y.-L., Liu, W.-H., Yu, C.-C., Lan, Y.-T., Huang, P.-I., Lee, Y.-Y., Chen, Y.-W., et al. (2015). Lin28B/Let-7 Regulates Expression of Oct4 and Sox2 and Reprograms Oral Squamous Cell Carcinoma Cells to a Stem-like State. *Cancer Res.* 75, 2553–2565.
- Chong, M.M.W.**, Zhang, G., Cheloufi, S., Neubert, T.A., Hannon, G.J., and Littman, D.R. (2010). Canonical and alternate functions of the microRNA biogenesis machinery. *Genes Dev.* 24, 1951–1960.
- Choueiri, T.K.**, Mayer, E.L., Je, Y., Rosenberg, J.E., Nguyen, P.L., Azzi, G.R., Bellmunt, J., Burstein, H.J., and Schutz, F.A.B. (2011). Congestive heart failure risk in patients with breast cancer treated with bevacizumab. *J. Clin. Oncol.* 29, 632–638.
- Chu, C.**, and Rana, T.M. (2006). Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. *PLoS Biol.* 4, e210.
- Cicalese, A.**, Bonizzi, G., Pasi, C.E., Faretta, M., Ronzoni, S., Giulini, B., Brisken, C., Minucci, S., Di Fiore, P.P., and Pelicci, P.G. (2009). The tumor suppressor p53 regulates polarity of self-renewing divisions in mammary stem cells. *Cell* 138, 1083–1095.
- Cifuentes, D.**, Xue, H., Taylor, D.W., Patnode, H., Mishima, Y., Cheloufi, S., Ma, E., Mane, S., Hannon, G.J., Lawson, N.D., et al. (2010). A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* 328, 1694–1698.
- Clark, B.S.**, and Blackshaw, S. (2014). Long non-coding RNA-dependent transcriptional regulation in neuronal development and disease. *Front Genet* 5, 164.
- Clevers, H.** (2011). The cancer stem cell: premises, promises and challenges. *Nat. Med.* 17, 313–319.
- Constância, 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, 203–206.
- Cooper, D.W.**, VandeBerg, J.L., Sharman, G.B., and Poole, W.E. (1971). Phosphoglycerate kinase polymorphism in kangaroos provides further evidence for paternal X inactivation. *Nature New Biol.* 230, 155–157.
- Copley, M.R.**, Babovic, S., Benz, C., Knapp, D.J.H.F., Beer, P.A., Kent, D.G., Wohrer, S., Treloar, D.Q., Day, C., Rowe, K., et al. (2013). The Lin28b-let-7-Hmga2 axis determines the higher self-renewal potential of fetal haematopoietic stem cells. *Nat. Cell Biol.* 15, 916–925.

Cortez, M.A., Valdecanas, D., Zhang, X., Zhan, Y., Bhardwaj, V., Calin, G.A., Komaki, R., Giri, D.K., Quini, C.C., Wolfe, T., et al. (2014). Therapeutic delivery of miR-200c enhances radiosensitivity in lung cancer. *Mol. Ther.* 22, 1494–1503.

Court, F., Baniol, M., Hagege, H., Petit, J.S., Lelay-Taha, M.-N., Carbonell, F., Weber, M., Cathala, G., and Forne, T. (2011). Long-range chromatin interactions at the mouse *Igf2/H19* locus reveal a novel paternally expressed long non-coding RNA. *Nucleic Acids Res.* 39, 5893–5906.

Crouse, H.V., Brown, A., and Mumford, B.C. (1971). --chromosome inheritance and the problem of chromosome “imprinting” in *Sciara* (Sciaridae, Diptera). *Chromosoma* 34, 324–339 8.

Cruz-Correa, M., Cui, H., Giardiello, F.M., Powe, N.R., Hylind, L., Robinson, A., Hutcheon, D.F., Kafonek, D.R., Brandenburg, S., Wu, Y., et al. (2004). Loss of imprinting of insulin growth factor II gene: a potential heritable biomarker for colon neoplasia predisposition. *Gastroenterology* 126, 964–970.

Cui, H. (2007). Loss of imprinting of IGF2 as an epigenetic marker for the risk of human cancer. *Dis. Markers* 23, 105–112.

Cui, H., Onyango, P., Brandenburg, S., Wu, Y., Hsieh, C.-L., and Feinberg, A.P. (2002). Loss of imprinting in colorectal cancer linked to hypomethylation of H19 and IGF2. *Cancer Res.* 62, 6442–6446.

Cui, H., Cruz-Correa, M., Giardiello, F.M., Hutcheon, D.F., Kafonek, D.R., Brandenburg, S., Wu, Y., He, X., Powe, N.R., and Feinberg, A.P. (2003). Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. *Science* 299, 1753–1755.

Curtin, J.F., and Cotter, T.G. (2004). JNK regulates HIPK3 expression and promotes resistance to Fas-mediated apoptosis in DU 145 prostate carcinoma cells. *J. Biol. Chem.* 279, 17090–17100.

Cusanelli, E., Romero, C.A.P., and Chartrand, P. (2013). Telomeric noncoding RNA TERRA is induced by telomere shortening to nucleate telomerase molecules at short telomeres. *Mol. Cell* 51, 780–791.

Czech, B., Zhou, R., Erlich, Y., Brennecke, J., Binari, R., Villalta, C., Gordon, A., Perrimon, N., and Hannon, G.J. (2009). Hierarchical rules for Argonaute loading in *Drosophila*. *Mol. Cell* 36, 445–456.

- Dahiya, N.**, Sherman-Baust, C.A., Wang, T.-L., Davidson, B., Shih, I.-M., Zhang, Y., Wood, W., Becker, K.G., and Morin, P.J. (2008). MicroRNA expression and identification of putative miRNA targets in ovarian cancer. *PLoS ONE* 3, e2436.
- Dean-Colomb, W.**, and Esteva, F.J. (2008). Her2-positive breast cancer: herceptin and beyond. *Eur. J. Cancer* 44, 2806–2812.
- 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, 849–859.
- Dekker, J.**, Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. *Science* 295, 1306–1311.
- DeLeo, A.B.**, Jay, G., Appella, E., Dubois, G.C., Law, L.W., and Old, L.J. (1979). Detection of a transformation-related antigen in chemically induced sarcomas and other transformed cells of the mouse. *Proc. Natl. Acad. Sci. U.S.A.* 76, 2420–2424.
- Demont, Y.**, Corbet, C., Page, A., Ataman-Önal, Y., Choquet-Kastylevsky, G., Fliniaux, I., Le Bourhis, X., Toillon, R.-A., Bradshaw, R.A., and Hondermarck, H. (2012). Pro-nerve growth factor induces autocrine stimulation of breast cancer cell invasion through tropomyosin-related kinase A (TrkA) and sortilin protein. *J. Biol. Chem.* 287, 1923–1931.
- Deng, Q.**, He, B., Gao, T., Pan, Y., Sun, H., Xu, Y., Li, R., Ying, H., Wang, F., Liu, X., et al. (2014). Up-regulation of 91H promotes tumor metastasis and predicts poor prognosis for patients with colorectal cancer. *PLoS ONE* 9, e103022.
- Denli, A.M.**, Tops, B.B.J., Plasterk, R.H.A., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. *Nature* 432, 231–235.
- Deplus, R.**, Brenner, C., Burgers, W.A., Putmans, P., Kouzarides, T., de Launoit, Y., and Fuks, F. (2002). Dnmt3L is a transcriptional repressor that recruits histone deacetylase. *Nucleic Acids Res.* 30, 3831–3838.
- Derrien, T.**, Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G., et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 22, 1775–1789.

- Descamps, S.**, Toillon, R.A., Adriaenssens, E., Pawlowski, V., Cool, S.M., Nurcombe, V., Le Bourhis, X., Boilly, B., Peyrat, J.P., and Hondermarck, H. (2001). Nerve growth factor stimulates proliferation and survival of human breast cancer cells through two distinct signaling pathways. *J. Biol. Chem.* 276, 17864–17870.
- Dey, B.K.**, Pfeifer, K., and Dutta, A. (2014). The H19 long noncoding RNA gives rise to microRNAs miR-675-3p and miR-675-5p to promote skeletal muscle differentiation and regeneration. *Genes Dev.* 28, 491–501.
- Diederichs, S.**, and Haber, D.A. (2007). Dual role for argonautes in microRNA processing and posttranscriptional regulation of microRNA expression. *Cell* 131, 1097–1108.
- Dieterich, M.**, Stubert, J., Reimer, T., Erickson, N., and Berling, A. (2014). Influence of lifestyle factors on breast cancer risk. *Breast Care (Basel)* 9, 407–414.
- Di Fiore, R.**, D’Anneo, A., Tesoriere, G., and Vento, R. (2013). RB1 in cancer: different mechanisms of RB1 inactivation and alterations of pRb pathway in tumorigenesis. *J. Cell. Physiol.* 228, 1676–1687.
- Dippold, W.G.**, Jay, G., DeLeo, A.B., Khoury, G., and Old, L.J. (1981). p53 transformation-related protein: detection by monoclonal antibody in mouse and human cells. *Proc. Natl. Acad. Sci. U.S.A.* 78, 1695–1699.
- Dirat, B.**, Bochet, L., Dabek, M., Daviaud, D., Dauvillier, S., Majed, B., Wang, Y.Y., Meulle, A., Salles, B., Le Gonidec, S., et al. (2011). Cancer-associated adipocytes exhibit an activated phenotype and contribute to breast cancer invasion. *Cancer Res.* 71, 2455–2465.
- Dollé, L.**, El Yazidi-Belkoura, I., Adriaenssens, E., Nurcombe, V., and Hondermarck, H. (2003). Nerve growth factor overexpression and autocrine loop in breast cancer cells. *Oncogene* 22, 5592–5601.
- Dong, Y.**, Li, J., Han, F., Chen, H., Zhao, X., Qin, Q., Shi, R., and Liu, J. (2015). High IGF2 expression is associated with poor clinical outcome in human ovarian cancer. *Oncology Reports*.
- Douc-Rasy, S.**, Barrois, M., Fogel, S., Ahomadegbe, J.C., Stéhelin, D., Coll, J., and Riou, G. (1996). High incidence of loss of heterozygosity and abnormal imprinting of H19 and IGF2 genes in invasive cervical carcinomas. Uncoupling of H19 and IGF2 expression and biallelic hypomethylation of H19. *Oncogene* 12, 423–430.

- 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, 3419–3428.
- 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, 1205–1213.
- 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, 1139–1144.
- Duart-Garcia, C.**, and Braunschweig, M.H. (2013). The Igf2as transcript is exported into cytoplasm and associated with polysomes. *Biochem. Genet.* 51, 119–130.
- Dudek, K.A.**, Lafont, J.E., Martinez-Sanchez, A., and Murphy, C.L. (2010). Type II collagen expression is regulated by tissue-specific miR-675 in human articular chondrocytes. *J. Biol. Chem.* 285, 24381–24387.
- Dueck, A.**, Ziegler, C., Eichner, A., Berezikov, E., and Meister, G. (2012). microRNAs associated with the different human Argonaute proteins. *Nucleic Acids Res.* 40, 9850–9862.
- Dugimont, T.**, Cury, J.J., Wernert, N., Delobelle, A., Raes, M.B., Joubel, A., Stehelin, D., and Coll, J. (1995). The H19 gene is expressed within both epithelial and stromal components of human invasive adenocarcinomas. *Biol. Cell* 85, 117–124.
- Dugimont, T.**, Montpellier, C., Adriaenssens, E., Lottin, S., Dumont, L., Iotsova, V., Lagrou, C., Stéhelin, D., Coll, J., and Cury, J.J. (1998). The H19 TATA-less promoter is efficiently repressed by wild-type tumor suppressor gene product p53. *Oncogene* 16, 2395–2401.
- Edbauer, D.**, Neilson, J.R., Foster, K.A., Wang, C.-F., Seeburg, D.P., Battersby, M.N., Tada, T., Dolan, B.M., Sharp, P.A., and Sheng, M. (2010). Regulation of synaptic structure and function by FMRP-associated microRNAs miR-125b and miR-132. *Neuron* 65, 373–384.
- Eißmann, M.**, Gutschner, T., Hämmerle, M., Günther, S., Caudron-Herger, M., Groß, M., Schirmacher, P., Rippe, K., Braun, T., Zörnig, M., et al. (2012). Loss of the abundant nuclear non-coding RNA MALAT1 is compatible with life and development. *RNA Biol* 9, 1076–1087.
- Elenbaas, B.**, and Weinberg, R.A. (2001). Heterotypic signaling between epithelial tumor cells and fibroblasts in carcinoma formation. *Exp. Cell Res.* 264, 169–184.
- ENCODE Project Consortium** (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74.

- Engel, N.**, Raval, A.K., Thorvaldsen, J.L., and Bartolomei, S.M. (2008). Three-dimensional conformation at the H19/Igf2 locus supports a model of enhancer tracking. *Hum. Mol. Genet.* *17*, 3021–3029.
- Eulalio, A.**, Behm-Ansmant, I., and Izaurralde, E. (2007). P bodies: at the crossroads of post-transcriptional pathways. *Nat. Rev. Mol. Cell Biol.* *8*, 9–22.
- Fabbri, M.**, Garzon, R., Cimmino, A., Liu, Z., Zanasi, N., Callegari, E., Liu, S., Alder, H., Costinean, S., Fernandez-Cymering, C., et al. (2007). MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. *Proc. Natl. Acad. Sci. U.S.A.* *104*, 15805–15810.
- Faghihi, M.A.**, Modarresi, F., Khalil, A.M., Wood, D.E., Sahagan, B.G., Morgan, T.E., Finch, C.E., St Laurent, G., Kenny, P.J., and Wahlestedt, C. (2008). Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat. Med.* *14*, 723–730.
- Faghihi, M.A.**, Zhang, M., Huang, J., Modarresi, F., Van der Brug, M.P., Nalls, M.A., Cookson, M.R., St-Laurent, G., and Wahlestedt, C. (2010). Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biol.* *11*, R56.
- Feil, R.**, Walter, J., Allen, N.D., and Reik, W. (1994). Developmental control of allelic methylation in the imprinted mouse Igf2 and H19 genes. *Development* *120*, 2933–2943.
- Ferguson-Smith, A.C.**, Cattanach, B.M., Barton, S.C., Beechey, C.V., and Surani, M.A. (1991). Embryological and molecular investigations of parental imprinting on mouse chromosome 7. *Nature* *351*, 667–670.
- Feuerhahn, S.**, Iglesias, N., Panza, A., Porro, A., and Lingner, J. (2010). TERRA biogenesis, turnover and implications for function. *FEBS Lett.* *584*, 3812–3818.
- Fitzpatrick, G.V.**, Pugacheva, E.M., Shin, J.-Y., Abdullaev, Z., Yang, Y., Khatod, K., Lobanenkova, V.V., and Higgins, M.J. (2007). Allele-specific binding of CTCF to the multipartite imprinting control region KvDMR1. *Mol. Cell. Biol.* *27*, 2636–2647.
- Flynt, A.S.**, Greimann, J.C., Chung, W.-J., Lima, C.D., and Lai, E.C. (2010). MicroRNA biogenesis via splicing and exosome-mediated trimming in *Drosophila*. *Mol. Cell* *38*, 900–907.
- Förstemann, K.**, Horwich, M.D., Wee, L., Tomari, Y., and Zamore, P.D. (2007). *Drosophila* microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1. *Cell* *130*, 287–297.

Foulkes, W.D., and Shuen, A.Y. (2013). In brief: BRCA1 and BRCA2. *J. Pathol.* 230, 347–349.

Freeman, J.A., and Espinosa, J.M. (2013). The impact of post-transcriptional regulation in the p53 network. *Brief Funct Genomics* 12, 46–57.

Fujita, N., Watanabe, S., Ichimura, T., Tsuruzoe, S., Shinkai, Y., Tachibana, M., Chiba, T., and Nakao, M. (2003). Methyl-CpG binding domain 1 (MBD1) interacts with the Suv39h1-HP1 heterochromatic complex for DNA methylation-based transcriptional repression. *J. Biol. Chem.* 278, 24132–24138.

Fuks, F. (2003). [DNA methyltransferases: from chromatin remodeling to cancer]. *Med Sci (Paris)* 19, 477–480.

Fuks, F., Burgers, W.A., Brehm, A., Hughes-Davies, L., and Kouzarides, T. (2000). DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nat. Genet.* 24, 88–91.

Fuks, F., Burgers, W.A., Godin, N., Kasai, M., and Kouzarides, T. (2001). Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. *EMBO J.* 20, 2536–2544.

Fuks, F., Hurd, P.J., Wolf, D., Nan, X., Bird, A.P., and Kouzarides, T. (2003). The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *J. Biol. Chem.* 278, 4035–4040.

Fukuzawa, R., Umezawa, A., Ochi, K., Urano, F., Ikeda, H., and Hata, J. (1999). High frequency of inactivation of the imprinted H19 gene in “sporadic” hepatoblastoma. *Int. J. Cancer* 82, 490–497.

Gao, T., He, B., Pan, Y., Xu, Y., Li, R., Deng, Q., Sun, H., and Wang, S. (2015). Long non-coding RNA 91H contributes to the occurrence and progression of esophageal squamous cell carcinoma by inhibiting IGF2 expression. *Mol. Carcinog.* 54, 359–367.

Gao, W.-L., Liu, M., Yang, Y., Yang, H., Liao, Q., Bai, Y., Li, Y.-X., Li, D., Peng, C., and Wang, Y.-L. (2012). The imprinted H19 gene regulates human placental trophoblast cell proliferation via encoding miR-675 that targets Nodal Modulator 1 (NOMO1). *RNA Biol* 9, 1002–1010.

Garcia, A.I., Buisson, M., Bertrand, P., Rimokh, R., Rouleau, E., Lopez, B.S., Lidereau, R., Mikaelian, I., and Mazoyer, S. (2011). Down-regulation of BRCA1 expression by miR-146a and miR-146b-5p in triple negative sporadic breast cancers. *EMBO Mol Med* 3, 279–290.

Garzon, R., Fabbri, M., Cimmino, A., Calin, G.A., and Croce, C.M. (2006). MicroRNA expression and function in cancer. *Trends Mol Med* 12, 580–587.

Garzon, R., Heaphy, C.E.A., Havelange, V., Fabbri, M., Volinia, S., Tsao, T., Zanesi, N., Kornblau, S.M., Marcucci, G., Calin, G.A., et al. (2009). MicroRNA 29b functions in acute myeloid leukemia. *Blood* *114*, 5331–5341.

Geisler, S., and Coller, J. (2013). RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. *Nat Rev Mol Cell Biol* *14*, 699–712.

Ghildiyal, M., Xu, J., Seitz, H., Weng, Z., and Zamore, P.D. (2010). Sorting of Drosophila small silencing RNAs partitions microRNA* strands into the RNA interference pathway. *RNA* *16*, 43–56.

Ginger, M.R., Gonzalez-Rimbau, M.F., Gay, J.P., and Rosen, J.M. (2001). Persistent changes in gene expression induced by estrogen and progesterone in the rat mammary gland. *Mol. Endocrinol.* *15*, 1993–2009.

Ginger, M.R., Shore, A.N., Contreras, A., Rijnkels, M., Miller, J., Gonzalez-Rimbau, M.F., and Rosen, J.M. (2006). A noncoding RNA is a potential marker of cell fate during mammary gland development. *Proc. Natl. Acad. Sci. U.S.A.* *103*, 5781–5786.

Giraldez, A.J., Mishima, Y., Rihel, J., Grocock, R.J., Van Dongen, S., Inoue, K., Enright, A.J., and Schier, A.F. (2006). Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* *312*, 75–79.

Gordon, F.E., Nutt, C.L., Cheunschon, P., Nakayama, Y., Provencher, K.A., Rice, K.A., Zhou, Y., Zhang, X., and Klibanski, A. (2010). Increased expression of angiogenic genes in the brains of mouse meg3-null embryos. *Endocrinology* *151*, 2443–2452.

Goswami, S., Sahai, E., Wyckoff, J.B., Cammer, M., Cox, D., Pixley, F.J., Stanley, E.R., Segall, J.E., and Condeelis, J.S. (2005). Macrophages promote the invasion of breast carcinoma cells via a colony-stimulating factor-1/epidermal growth factor paracrine loop. *Cancer Res.* *65*, 5278–5283.

Grandjean, V., O'Neill, L., Sado, T., Turner, B., and Ferguson-Smith, A. (2001). Relationship between DNA methylation, histone H4 acetylation and gene expression in the mouse imprinted Igf2-H19 domain. *FEBS Lett.* *488*, 165–169.

Gregory, R.I., Yan, K.-P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. *Nature* *432*, 235–240.

Gregory, R.I., Chendrimada, T.P., Cooch, N., and Shiekhattar, R. (2005). Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* *123*, 631–640.

Grishok, A., Pasquinelli, A.E., Conte, D., Li, N., Parrish, S., Ha, I., Baillie, D.L., Fire, A., Ruvkun, G., and Mello, C.C. (2001). Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* 106, 23–34.

Grotwinkel, J.T., Wild, K., Segnitz, B., and Sinning, I. (2014). SRP RNA remodeling by SRP68 explains its role in protein translocation. *Science* 344, 101–104.

Gumireddy, K., Li, A., Yan, J., Setoyama, T., Johannes, G.J., Orom, U.A., Tchou, J., Liu, Q., Zhang, L., Speicher, D.W., et al. (2013). Identification of a long non-coding RNA-associated RNP complex regulating metastasis at the translational step. *EMBO J.* 32, 2672–2684.

Ha, M., and Kim, V.N. (2014). Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* 15, 509–524.

Haase, A.D., Jaskiewicz, L., Zhang, H., Lainé, S., Sack, R., Gagnol, A., and Filipowicz, W. (2005). TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. *EMBO Rep.* 6, 961–967.

Hagège, H., Klous, P., Braem, C., Splinter, E., Dekker, J., Cathala, G., de Laat, W., and Forné, T. (2007). Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nat Protoc* 2, 1722–1733.

Hammond, S.M., Boettcher, S., Caudy, A.A., Kobayashi, R., and Hannon, G.J. (2001). Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* 293, 1146–1150.

Han, J., Lee, Y., Yeom, K.-H., Kim, Y.-K., Jin, H., and Kim, V.N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev.* 18, 3016–3027.

Han, J., Lee, Y., Yeom, K.-H., Nam, J.-W., Heo, I., Rhee, J.-K., Sohn, S.Y., Cho, Y., Zhang, B.-T., and Kim, V.N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* 125, 887–901.

Han, K., Gennarino, V.A., Lee, Y., Pang, K., Hashimoto-Torii, K., Choufani, S., Raju, C.S., Oldham, M.C., Weksberg, R., Rakic, P., et al. (2013). Human-specific regulation of MeCP2 levels in fetal brains by microRNA miR-483-5p. *Genes Dev.* 27, 485–490.

Hanahan, D., and Coussens, L.M. (2012). Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* 21, 309–322.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. *Cell* 100, 57–70.

- Hanahan, D.**, and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674.
- Hansen, T.B.**, Jensen, T.I., Clausen, B.H., Bramsen, J.B., Finsen, B., Damgaard, C.K., and Kjems, J. (2013). Natural RNA circles function as efficient microRNA sponges. *Nature* 495, 384–388.
- Hansji, H.**, Leung, E.Y., Baguley, B.C., Finlay, G.J., and Askarian-Amiri, M.E. (2014). Keeping abreast with long non-coding RNAs in mammary gland development and breast cancer. *Front Genet* 5, 379.
- Hao, Y.**, Crenshaw, T., Moulton, T., Newcomb, E., and Tycko, B. (1993). Tumour-suppressor activity of H19 RNA. *Nature* 365, 764–767.
- Hark, A.T.**, and Tilghman, S.M. (1998). Chromatin conformation of the H19 epigenetic mark. *Hum. Mol. Genet.* 7, 1979–1985.
- Hark, A.T.**, Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 405, 486–489.
- Hartmann, L.C.**, Sellers, T.A., Frost, M.H., Lingle, W.L., Degnim, A.C., Ghosh, K., Vierkant, R.A., Maloney, S.D., Pankratz, V.S., Hillman, D.W., et al. (2005). Benign breast disease and the risk of breast cancer. *N. Engl. J. Med.* 353, 229–237.
- Hashimoto, K.**, Azuma, C., Tokugawa, Y., Nobunaga, T., Aki, T.A., Matsui, Y., Yanagida, T., Izumi, H., Saji, F., and Murata, Y. (1997). Loss of H19 imprinting and up-regulation of H19 and SNRPN in a case with malignant mixed Müllerian tumor of the uterus. *Hum. Pathol.* 28, 862–865.
- He, H.**, Jazdzewski, K., Li, W., Liyanarachchi, S., Nagy, R., Volinia, S., Calin, G.A., Liu, C.-G., Franssila, K., Suster, S., et al. (2005). The role of microRNA genes in papillary thyroid carcinoma. *Proc. Natl. Acad. Sci. U.S.A.* 102, 19075–19080.
- He, X.-X.**, Chang, Y., Meng, F.-Y., Wang, M.-Y., Xie, Q.-H., Tang, F., Li, P.-Y., Song, Y.-H., and Lin, J.-S. (2012). MicroRNA-375 targets AEG-1 in hepatocellular carcinoma and suppresses liver cancer cell growth in vitro and in vivo. *Oncogene* 31, 3357–3369.
- Hemberger, M.**, Redies, C., Krause, R., Oswald, J., Walter, J., and Fundele, R.H. (1998). H19 and Igf2 are expressed and differentially imprinted in neuroectoderm-derived cells in the mouse brain. *Dev. Genes Evol.* 208, 393–402.

- Henke, J.I.**, Goergen, D., Zheng, J., Song, Y., Schüttler, C.G., Fehr, C., Jünemann, C., and Niepmann, M. (2008). microRNA-122 stimulates translation of hepatitis C virus RNA. *EMBO J.* 27, 3300–3310.
- Heo, J.B.**, and Sung, S. (2011). Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science* 331, 76–79.
- Hernandez, J.M.**, Elahi, A., Clark, C.W., Wang, J., Humphries, L.A., Centeno, B., Bloom, G., Fuchs, B.C., Yeatman, T., and Shibata, D. (2013). miR-675 Mediates Downregulation of Twist1 and Rb in AFP-Secreting Hepatocellular Carcinoma. *Ann. Surg. Oncol.*
- Hibi, K.**, Nakamura, H., Hirai, A., Fujikake, Y., Kasai, Y., Akiyama, S., Ito, K., and Takagi, H. (1996). Loss of H19 imprinting in esophageal cancer. *Cancer Res.* 56, 480–482.
- Higgins, M.J.**, and Baselga, J. (2011). Targeted therapies for breast cancer. *J. Clin. Invest.* 121, 3797–3803.
- Hirsch, D.S.**, Shen, Y., and Wu, W.J. (2006). Growth and motility inhibition of breast cancer cells by epidermal growth factor receptor degradation is correlated with inactivation of Cdc42. *Cancer Res.* 66, 3523–3530.
- Hori, N.**, Yamane, M., Kouno, K., and Sato, K. (2012). Induction of DNA demethylation depending on two sets of Sox2 and adjacent Oct3/4 binding sites (Sox-Oct motifs) within the mouse H19/insulin-like growth factor 2 (Igf2) imprinted control region. *J. Biol. Chem.* 287, 44006–44016.
- Hou, P.**, Zhao, Y., Li, Z., Yao, R., Ma, M., Gao, Y., Zhao, L., Zhang, Y., Huang, B., and Lu, J. (2014). LincRNA-ROR induces epithelial-to-mesenchymal transition and contributes to breast cancer tumorigenesis and metastasis. *Cell Death Dis* 5, e1287.
- Howell, A.**, Cuzick, J., Baum, M., Buzdar, A., Dowsett, M., Forbes, J.F., Hochtin-Boes, G., Houghton, J., Locker, G.Y., Tobias, J.S., et al. (2005). Results of the ATAC (Arimidex, Tamoxifen, Alone or in Combination) trial after completion of 5 years' adjuvant treatment for breast cancer. *Lancet* 365, 60–62.
- Hu, M.**, and Polyak, K. (2008). Molecular characterisation of the tumour microenvironment in breast cancer. *Eur. J. Cancer* 44, 2760–2765.
- Hu, H.Y.**, Yan, Z., Xu, Y., Hu, H., Menzel, C., Zhou, Y.H., Chen, W., and Khaitovich, P. (2009). Sequence features associated with microRNA strand selection in humans and flies. *BMC Genomics* 10, 413.

- Huang, C.**, Cao, L., Qiu, L., Dai, X., Ma, L., Zhou, Y., Li, H., Gao, M., Li, W., Zhang, Q., et al. (2015). Upregulation of H19 promotes invasion and induces epithelial-to-mesenchymal transition in esophageal cancer. *Oncol Lett* 10, 291–296.
- Huang, J.**, Zhou, N., Watabe, K., Lu, Z., Wu, F., Xu, M., and Mo, Y.-Y. (2014). Long non-coding RNA UCA1 promotes breast tumor growth by suppression of p27 (Kip1). *Cell Death Dis* 5, e1008.
- Humphreys, D.T.**, Westman, B.J., Martin, D.I.K., and Preiss, T. (2005). MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. *Proc. Natl. Acad. Sci. U.S.A.* 102, 16961–16966.
- Hutvagner, G.**, and Zamore, P.D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. *Science* 297, 2056–2060.
- Hutvagner, G.**, McLachlan, J., Pasquinelli, A.E., Bálint, E., Tuschl, T., and Zamore, P.D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 293, 834–838.
- Ideraabdullah, F.Y.**, Vigneau, S., and Bartolomei, M.S. (2008). Genomic imprinting mechanisms in mammals. *Mutat. Res.* 647, 77–85.
- Iglehart, J.D.**, and Silver, D.P. (2009). Synthetic lethality--a new direction in cancer-drug development. *N. Engl. J. Med.* 361, 189–191.
- Imig, J.**, Brunschweiler, A., Brümmer, A., Guennewig, B., Mittal, N., Kishore, S., Tsikrika, P., Gerber, A.P., Zavolan, M., and Hall, J. (2015). miR-CLIP capture of a miRNA targetome uncovers a lincRNA H19-miR-106a interaction. *Nat. Chem. Biol.* 11, 107–114.
- Iorio, M.V.**, Ferracin, M., Liu, C.-G., Veronese, A., Spizzo, R., Sabbioni, S., Magri, E., Pedriali, M., Fabbri, M., Campiglio, M., et al. (2005). MicroRNA gene expression deregulation in human breast cancer. *Cancer Res.* 65, 7065–7070.
- Ishihara, K.**, and Sasaki, H. (2002). An evolutionarily conserved putative insulator element near the 3' boundary of the imprinted Igf2/H19 domain. *Hum. Mol. Genet.* 11, 1627–1636.
- Ito, Y.**, Nativio, R., and Murrell, A. (2013). Induced DNA demethylation can reshape chromatin topology at the IGF2-H19 locus. *Nucleic Acids Res.* 41, 5290–5302.
- Iyengar, P.**, Espina, V., Williams, T.W., Lin, Y., Berry, D., Jelicks, L.A., Lee, H., Temple, K., Graves, R., Pollard, J., et al. (2005). Adipocyte-derived collagen VI affects early mammary

tumor progression in vivo, demonstrating a critical interaction in the tumor/stroma microenvironment. *J. Clin. Invest.* 115, 1163–1176.

Jackson, R.J., and Standart, N. (2007). How do microRNAs regulate gene expression? *Sci. STKE* 2007, re1.

Jeon, Y., and Lee, J.T. (2011). YY1 tethers Xist RNA to the inactive X nucleation center. *Cell* 146, 119–133.

Ji, P., Diederichs, S., Wang, W., Böing, S., Metzger, R., Schneider, P.M., Tidow, N., Brandt, B., Buerger, H., Bulk, E., et al. (2003). MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* 22, 8031–8041.

Jiang, M., Huang, O., Xie, Z., Wu, S., Zhang, X., Shen, A., Liu, H., Chen, X., Wu, J., Lou, Y., et al. (2014). A novel long non-coding RNA-ARA: adriamycin resistance-associated. *Biochem. Pharmacol.* 87, 254–283.

Jiang, X., Yan, Y., Hu, M., Chen, X., Wang, Y., Dai, Y., Wu, D., Wang, Y., Zhuang, Z., and Xia, H. (2015). Increased level of H19 long noncoding RNA promotes invasion, angiogenesis, and stemness of glioblastoma cells. *J. Neurosurg.* 1–8.

Joubel, A., Curgy, J.J., Pelczar, H., Begue, A., Lagrou, C., Stehelin, D., and Coll, J. (1996). The 5' part of the human H19 RNA contains cis-acting elements hampering its translatability. *Cell. Mol. Biol. (Noisy-Le-Grand)* 42, 1159–1172.

Jouvenot, Y., Poirier, F., Jami, J., and Paldi, A. (1999). Biallelic transcription of Igf2 and H19 in individual cells suggests a post-transcriptional contribution to genomic imprinting. *Curr. Biol.* 9, 1199–1202.

Juan, V., Crain, C., and Wilson, C. (2000). Evidence for evolutionarily conserved secondary structure in the H19 tumor suppressor RNA. *Nucleic Acids Res.* 28, 1221–1227.

Kallen, A.N., Zhou, X.-B., Xu, J., Qiao, C., Ma, J., Yan, L., Lu, L., Liu, C., Yi, J.-S., Zhang, H., et al. (2013). The imprinted H19 lncRNA antagonizes let-7 microRNAs. *Mol. Cell* 52, 101–112.

Kalluri, R., and Zeisberg, M. (2006). Fibroblasts in cancer. *Nat. Rev. Cancer* 6, 392–401.

Kanduri, C., Thakur, N., and Pandey, R.R. (2006). The length of the transcript encoded from the Kcnq1ot1 antisense promoter determines the degree of silencing. *EMBO J.* 25, 2096–2106.

- Kaneko, S.**, Bonasio, R., Saldaña-Meyer, R., Yoshida, T., Son, J., Nishino, K., Umezawa, A., and Reinberg, D. (2014). Interactions between JARID2 and noncoding RNAs regulate PRC2 recruitment to chromatin. *Mol. Cell* 53, 290–300.
- Kanellopoulou, C.**, Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev.* 19, 489–501.
- Kanhere, A.**, and Jenner, R.G. (2012). Noncoding RNA localisation mechanisms in chromatin regulation. *Silence* 3, 2.
- Kartha, R.V.**, and Subramanian, S. (2014). Competing endogenous RNAs (ceRNAs): new entrants to the intricacies of gene regulation. *Front Genet* 5, 8.
- Kasinski, A.L.**, and Slack, F.J. (2011). Epigenetics and genetics. MicroRNAs en route to the clinic: progress in validating and targeting microRNAs for cancer therapy. *Nat. Rev. Cancer* 11, 849–864.
- Kasinski, A.L.**, Kelnar, K., Stahlhut, C., Orellana, E., Zhao, J., Shimer, E., Dysart, S., Chen, X., Bader, A.G., and Slack, F.J. (2014). A combinatorial microRNA therapeutics approach to suppressing non-small cell lung cancer. *Oncogene*.
- Kato, Y.**, and Sasaki, H. (2005). Imprinting and looping: epigenetic marks control interactions between regulatory elements. *Bioessays* 27, 1–4.
- Kawamata, T.**, and Tomari, Y. (2010). Making RISC. *Trends Biochem. Sci.* 35, 368–376.
- Kedde, M.**, Strasser, M.J., Boldajipour, B., Oude Vrielink, J.A.F., Slanchev, K., le Sage, C., Nagel, R., Voorhoeve, P.M., van Duijse, J., Ørom, U.A., et al. (2007). RNA-binding protein Dnd1 inhibits microRNA access to target mRNA. *Cell* 131, 1273–1286.
- Keniry, A.**, Oxley, D., Monnier, P., Kyba, M., Dandolo, L., Smits, G., and Reik, W. (2012). The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. *Nat. Cell Biol.* 14, 659–665.
- Ketting, R.F.**, Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.* 15, 2654–2659.
- Khera, T.K.**, Dick, A.D., and Nicholson, L.B. (2010). Fragile X-related protein FXR1 controls post-transcriptional suppression of lipopolysaccharide-induced tumour necrosis factor- α production by transforming growth factor- β 1. *FEBS J.* 277, 2754–2765.

- Khvorova, A.**, Reynolds, A., and Jayasena, S.D. (2003). Functional siRNAs and miRNAs exhibit strand bias. *Cell* 115, 209–216.
- Kim, K.S.**, and Lee, Y.I. (1997). Biallelic expression of the H19 and IGF2 genes in hepatocellular carcinoma. *Cancer Lett.* 119, 143–148.
- Kim, Y.-K.**, and Kim, V.N. (2007). Processing of intronic microRNAs. *EMBO J.* 26, 775–783.
- Kim, H.H.**, Kuwano, Y., Srikantan, S., Lee, E.K., Martindale, J.L., and Gorospe, M. (2009). HuR recruits let-7/RISC to repress c-Myc expression. *Genes Dev.* 23, 1743–1748.
- Kim, H.T.**, Choi, B.H., Niikawa, N., Lee, T.S., and Chang, S.I. (1998). Frequent loss of imprinting of the H19 and IGF-II genes in ovarian tumors. *Am. J. Med. Genet.* 80, 391–395.
- Kim, N.-H.**, Choi, S.-H., Kim, C.-H., Lee, C.H., Lee, T.R., and Lee, A.-Y. (2014a). Reduced MiR-675 in exosome in H19 RNA-related melanogenesis via MITF as a direct target. *J. Invest. Dermatol.* 134, 1075–1082.
- Kim, N.-H.**, Choi, S.-H., Lee, T.R., Lee, C.-H., and Lee, A.-Y. (2014b). Cadherin 11, a miR-675 target, induces N-cadherin expression and epithelial-mesenchymal transition in melasma. *J. Invest. Dermatol.* 134, 2967–2976.
- Kim, H.**, Lee, H.-J., Oh, Y., Choi, S.-G., Hong, S.-H., Kim, H.-J., Lee, S.-Y., Choi, J.-W., Su Hwang, D., Kim, K.-S., et al. (2014c). The DUSP26 phosphatase activator adenylate kinase 2 regulates FADD phosphorylation and cell growth. *Nat Commun* 5, 3351.
- Kino, T.**, Hurt, D.E., Ichijo, T., Nader, N., and Chrousos, G.P. (2010). Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci Signal* 3, ra8.
- Kleer, C.G.**, Bloushtain-Qimron, N., Chen, Y.-H., Carrasco, D., Hu, M., Yao, J., Kraeft, S.-K., Collins, L.C., Sabel, M.S., Argani, P., et al. (2008). Epithelial and stromal cathepsin K and CXCL14 expression in breast tumor progression. *Clin. Cancer Res.* 14, 5357–5367.
- Knight, S.W.**, and Bass, B.L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* 293, 2269–2271.
- Knudson, A.G.** (1971). Mutation and cancer: statistical study of retinoblastoma. *Proc. Natl. Acad. Sci. U.S.A.* 68, 820–823.
- Knudson, A.G.** (2001). Two genetic hits (more or less) to cancer. *Nat. Rev. Cancer* 1, 157–162.

Kondo, M., Suzuki, H., Ueda, R., Osada, H., Takagi, K., Takahashi, T., and Takahashi, T. (1995). Frequent loss of imprinting of the H19 gene is often associated with its overexpression in human lung cancers. *Oncogene* 10, 1193–1198.

Kong, D., Heath, E., Chen, W., Cher, M.L., Powell, I., Heilbrun, L., Li, Y., Ali, S., Sethi, S., Hassan, O., et al. (2012). Loss of let-7 up-regulates EZH2 in prostate cancer consistent with the acquisition of cancer stem cell signatures that are attenuated by BR-DIM. *PLoS ONE* 7, e33729.

Kopf, E., Bibi, O., Ayesb, S., Tykocinski, M., Vitner, K., Looijenga, L.H., de Groot, N., and Hochberg, A. (1998). The effect of retinoic acid on the activation of the human H19 promoter by a 3' downstream region. *FEBS Lett.* 432, 123–127.

Krol, J., Loedige, I., and Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.* 11, 597–610.

Kung, J.T.Y., Colognori, D., and Lee, J.T. (2013). Long noncoding RNAs: past, present, and future. *Genetics* 193, 651–669.

Kurukuti, S., Tiwari, V.K., Tavoosidana, G., Pugacheva, E., Murrell, A., Zhao, Z., Lobanenko, 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, 10684–10689.

Lacroix, M., Toillon, R.-A., and Leclercq, G. (2006). p53 and breast cancer, an update. *Endocr. Relat. Cancer* 13, 293–325.

Lagadec, C., Meignan, S., Adriaenssens, E., Foveau, B., Vanhecke, E., Romon, R., Toillon, R.-A., Oxombre, B., Hondermarck, H., and Le Bourhis, X. (2009). TrkA overexpression enhances growth and metastasis of breast cancer cells. *Oncogene* 28, 1960–1970.

Landers, M., Bancescu, D.L., Le Meur, E., Rougeulle, C., Glatt-Deeley, H., Brannan, C., Muscatelli, F., and Lalande, M. (2004). Regulation of the large (approximately 1000 kb) imprinted murine Ube3a antisense transcript by alternative exons upstream of Snurf/Snrpn. *Nucleic Acids Res.* 32, 3480–3492.

Lanz, R.B., McKenna, N.J., Onate, S.A., Albrecht, U., Wong, J., Tsai, S.Y., Tsai, M.J., and O'Malley, B.W. (1999). A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. *Cell* 97, 17–27.

- Lanz, R.B.**, Chua, S.S., Barron, N., Söder, B.M., DeMayo, F., and O'Malley, B.W. (2003). Steroid receptor RNA activator stimulates proliferation as well as apoptosis in vivo. *Mol. Cell Biol.* *23*, 7163–7176.
- Laoui, D.**, Movahedi, K., Van Overmeire, E., Van den Bossche, J., Schouppe, E., Mommer, C., Nikolaou, A., Morias, Y., De Baetselier, P., and Van Ginderachter, J.A. (2011). Tumor-associated macrophages in breast cancer: distinct subsets, distinct functions. *Int. J. Dev. Biol.* *55*, 861–867.
- Lau, N.C.**, Lim, L.P., Weinstein, E.G., and Bartel, D.P. (2001). An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* *294*, 858–862.
- Lee, J.T.** (2012). Epigenetic regulation by long noncoding RNAs. *Science* *338*, 1435–1439.
- Lee, J.T.**, Davidow, L.S., and Warshawsky, D. (1999). Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat. Genet.* *21*, 400–404.
- Lee, R.C.**, Feinbaum, R.L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* *75*, 843–854.
- Lee, Y.**, Hur, I., Park, S.-Y., Kim, Y.-K., Suh, M.R., and Kim, V.N. (2006). The role of PACT in the RNA silencing pathway. *EMBO J.* *25*, 522–532.
- Lee, Y.S.**, Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J., and Carthew, R.W. (2004). Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* *117*, 69–81.
- Leibovitch, M.P.**, Solhonne, B., Guillier, M., Verrelle, P., Leibovitch, S.A., and Verelle P [corrected to Verrelle, P. (1995). Direct relationship between the expression of tumor suppressor H19 mRNA and *c-mos* proto-oncogene during myogenesis. *Oncogene* *10*, 251–260.
- 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*, 34–39.
- Leonhardt, H.**, Page, A.W., Weier, H.U., and Bestor, T.H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* *71*, 865–873.
- Lerebours, F.**, and Lidereau, R. (2002). Molecular alterations in sporadic breast cancer. *Crit. Rev. Oncol. Hematol.* *44*, 121–141.
- Leuschner, P.J.F.**, Ameres, S.L., Kueng, S., and Martinez, J. (2006). Cleavage of the siRNA passenger strand during RISC assembly in human cells. *EMBO Rep.* *7*, 314–320.

Li, C., Lei, B., Huang, S., Zheng, M., Liu, Z., Li, Z., and Deng, Y. (2015). H19 derived microRNA-675 regulates cell proliferation and migration through CDK6 in glioma. *Am J Transl Res* 7, 1747–1764.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69, 915–926.

Li, H., Li, J., Jia, S., Wu, M., An, J., Zheng, Q., Zhang, W., and Lu, D. (2015). miR675 upregulates long noncoding RNA H19 through activating EGR1 in human liver cancer. *Oncotarget* 6, 31958–31984.

Li, H., Yu, B., Li, J., Su, L., Yan, M., Zhu, Z., and Liu, B. (2014). Overexpression of lncRNA H19 enhances carcinogenesis and metastasis of gastric cancer. *Oncotarget* 5, 2318–2329.

Li, S., Yu, Z., Chen, S.-S., Li, F., Lei, C.-Y., Chen, X.-X., Bao, J.-M., Luo, Y., Lin, G.-Z., Pang, S.-Y., et al. (2015). The YAP1 oncogene contributes to bladder cancer cell proliferation and migration by regulating the H19 long noncoding RNA. *Urol. Oncol.* 33, 427.e1–e10.

Li, T., Hu, J.-F., Qiu, X., Ling, J., Chen, H., Wang, S., Hou, A., Vu, T.H., and Hoffman, A.R. (2008). CTCF regulates allelic expression of *Igf2* by orchestrating a promoter-polycomb repressive complex 2 intrachromosomal loop. *Mol. Cell. Biol.* 28, 6473–6482.

Li, W., Notani, D., Ma, Q., Tanasa, B., Nunez, E., Chen, A.Y., Merkurjev, D., Zhang, J., Ohgi, K., Song, X., et al. (2013). Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* 498, 516–520.

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, 28247–28252.

Li, Y., and Zhao, X. (2014). Concise review: Fragile X proteins in stem cell maintenance and differentiation. *Stem Cells* 32, 1724–1733.

Liang, W.-C., Fu, W.-M., Wong, C.-W., Wang, Y., Wang, W.-M., Hu, G.-X., Zhang, L., Xiao, L.-J., Wan, D.C.-C., Zhang, J.-F., et al. (2015). The lncRNA H19 promotes epithelial to mesenchymal transition by functioning as miRNA sponges in colorectal cancer. *Oncotarget* 6, 22513–22525.

Liao, B., Hu, Y., Herrick, D.J., and Brewer, G. (2005). The RNA-binding protein IMP-3 is a translational activator of insulin-like growth factor II leader-3 mRNA during proliferation of human K562 leukemia cells. *J. Biol. Chem.* 280, 18517–18524.

- Lin, R.**, Roychowdhury-Saha, M., Black, C., Watt, A.T., Marcusson, E.G., Freier, S.M., and Edgington, T.S. (2011). Control of RNA processing by a large non-coding RNA over-expressed in carcinomas. *FEBS Lett.* *585*, 671–676.
- Liu, C.**, Chen, Z., Fang, J., Xu, A., Zhang, W., and Wang, Z. (2015). H19-derived miR-675 contributes to bladder cancer cell proliferation by regulating p53 activation. *Tumour Biol.*
- Liu, J.**, Kahri, A.I., Heikkilä, P., Ilvesmäki, V., and Voutilainen, R. (1995). H19 and insulin-like growth factor-II gene expression in adrenal tumors and cultured adrenal cells. *J. Clin. Endocrinol. Metab.* *80*, 492–496.
- Liu, J.**, Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.-J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. *Science* *305*, 1437–1441.
- Liu, T.**, Yacoub, R., Taliaferro-Smith, L.D., Sun, S.-Y., Graham, T.R., Dolan, R., Lobo, C., Tighiouart, M., Yang, L., Adams, A., et al. (2011). Combinatorial effects of lapatinib and rapamycin in triple-negative breast cancer cells. *Mol. Cancer Ther.* *10*, 1460–1469.
- Liu, Y.**, Oakeley, E.J., Sun, L., and Jost, J.P. (1998). Multiple domains are involved in the targeting of the mouse DNA methyltransferase to the DNA replication foci. *Nucleic Acids Res.* *26*, 1038–1045.
- Liu, Y.**, Ye, X., Jiang, F., Liang, C., Chen, D., Peng, J., Kinch, L.N., Grishin, N.V., and Liu, Q. (2009). C3PO, an endoribonuclease that promotes RNAi by facilitating RISC activation. *Science* *325*, 750–753.
- Lochter, A.**, Galosy, S., Muschler, J., Freedman, N., Werb, Z., and Bissell, M.J. (1997). Matrix metalloproteinase stromelysin-1 triggers a cascade of molecular alterations that leads to stable epithelial-to-mesenchymal conversion and a premalignant phenotype in mammary epithelial cells. *J. Cell Biol.* *139*, 1861–1872.
- Loewer, S.**, Cabili, M.N., Guttman, M., Loh, Y.-H., Thomas, K., Park, I.H., Garber, M., Curran, M., Onder, T., Agarwal, S., et al. (2010). Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nat. Genet.* *42*, 1113–1117.
- Loffreda, A.**, Rigamonti, A., Barabino, S.M.L., and Lenzken, S.C. (2015). RNA-Binding Proteins in the Regulation of miRNA Activity: A Focus on Neuronal Functions. *Biomolecules* *5*, 2363–2387.

Lottin, S., Adriaenssens, E., Dupressoir, T., Berteaux, N., Montpellier, C., Coll, J., Dugimont, T., and Cury, J.J. (2002). Overexpression of an ectopic H19 gene enhances the tumorigenic properties of breast cancer cells. *Carcinogenesis* 23, 1885–1895.

Lottin, S., Adriaenssens, E., Berteaux, N., Leprêtre, A., Vilain, M.-O., Denhez, E., Coll, J., Dugimont, T., and Cury, J.-J. (2005). The human H19 gene is frequently overexpressed in myometrium and stroma during pathological endometrial proliferative events. *Eur. J. Cancer* 41, 168–177.

Lund, E., Güttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear export of microRNA precursors. *Science* 303, 95–98.

Luo, M., Li, Z., Wang, W., Zeng, Y., Liu, Z., and Qiu, J. (2013). Long non-coding RNA H19 increases bladder cancer metastasis by associating with EZH2 and inhibiting E-cadherin expression. *Cancer Lett.* 333, 213–221.

Lustig, O., Ariel, I., Ilan, J., Lev-Lehman, E., De-Groot, N., and Hochberg, A. (1994). Expression of the imprinted gene H19 in the human fetus. *Mol. Reprod. Dev.* 38, 239–246.

Lyle, R., Watanabe, D., te Vruchte, D., Lerchner, W., Smrzka, O.W., Wutz, A., Schageman, J., Hahner, L., Davies, C., and Barlow, D.P. (2000). The imprinted antisense RNA at the *Igf2r* locus overlaps but does not imprint *Mas1*. *Nat. Genet.* 25, 19–21.

Lynch, C.A., Tycko, B., Bestor, T.H., and Walsh, C.P. (2002). Reactivation of a silenced H19 gene in human rhabdomyosarcoma by demethylation of DNA but not by histone hyperacetylation. *Mol. Cancer* 1, 2.

Ma, J.-B., Yuan, Y.-R., Meister, G., Pei, Y., Tuschl, T., and Patel, D.J. (2005). Structural basis for 5'-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. *Nature* 434, 666–670.

Ma, L., Reinhardt, F., Pan, E., Soutschek, J., Bhat, B., Marcusson, E.G., Teruya-Feldstein, J., Bell, G.W., and Weinberg, R.A. (2010). Therapeutic silencing of miR-10b inhibits metastasis in a mouse mammary tumor model. *Nat. Biotechnol.* 28, 341–347.

Ma, P.J., Zhang, H., Li, R., Wang, Y.S., Zhang, Y., and Hua, S. (2015). P53-Mediated Repression of the Reprogramming in Cloned Bovine Embryos Through Direct Interaction with HDAC1 and Indirect Interaction with DNMT3A. *Reprod. Domest. Anim.* 50, 400–409.

Macrae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.Z., Adams, P.D., and Doudna, J.A. (2006). Structural basis for double-stranded RNA processing by Dicer. *Science* 311, 195–198.

- MacRae, I.J.**, Ma, E., Zhou, M., Robinson, C.V., and Doudna, J.A. (2008). In vitro reconstitution of the human RISC-loading complex. *Proc. Natl. Acad. Sci. U.S.A.* *105*, 512–517.
- Mahdi, K.M.**, Nassiri, M.R., and Nasiri, K. (2013). Hereditary genes and SNPs associated with breast cancer. *Asian Pac. J. Cancer Prev.* *14*, 3403–3409.
- Maniataki, E.**, and Mourelatos, Z. (2005). A human, ATP-independent, RISC assembly machine fueled by pre-miRNA. *Genes Dev.* *19*, 2979–2990.
- Mao, Y.**, Keller, E.T., Garfield, D.H., Shen, K., and Wang, J. (2013). Stromal cells in tumor microenvironment and breast cancer. *Cancer Metastasis Rev.* *32*, 303–315.
- Mao, Y.S.**, Sunwoo, H., Zhang, B., and Spector, D.L. (2011). Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat. Cell Biol.* *13*, 95–101.
- Maroney, P.A.**, Yu, Y., Fisher, J., and Nilsen, T.W. (2006). Evidence that microRNAs are associated with translating messenger RNAs in human cells. *Nat. Struct. Mol. Biol.* *13*, 1102–1107.
- Maroulakou, I.G.**, Anver, M., Garrett, L., and Green, J.E. (1994). Prostate and mammary adenocarcinoma in transgenic mice carrying a rat C3(1) simian virus 40 large tumor antigen fusion gene. *Proc. Natl. Acad. Sci. U.S.A.* *91*, 11236–11240.
- Martianov, I.**, Ramadass, A., Serra Barros, A., Chow, N., and Akoulitchev, A. (2007). Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* *445*, 666–670.
- Martinez, J.**, and Tuschl, T. (2004). RISC is a 5' phosphomonoester-producing RNA endonuclease. *Genes Dev.* *18*, 975–980.
- Mathonnet, G.**, Fabian, M.R., Svitkin, Y.V., Parsyan, A., Huck, L., Murata, T., Biffo, S., Merrick, W.C., Darzynkiewicz, E., Pillai, R.S., et al. (2007). MicroRNA inhibition of translation initiation in vitro by targeting the cap-binding complex eIF4F. *Science* *317*, 1764–1767.
- Matouk, I.J.**, Mezan, S., Mizrahi, A., Ohana, P., Abu-Lail, R., Fellig, Y., Degroot, N., Galun, E., and Hochberg, A. (2010). The oncofetal H19 RNA connection: hypoxia, p53 and cancer. *Biochim. Biophys. Acta* *1803*, 443–451.
- Matouk, I.J.**, Raveh, E., Abu-lail Rasha, Mezan, S., Gilon, M., Gershtain, E., Birman, T., Gallula, J., Schneider, T., Barkali, M., et al. (2014). Oncofetal H19 RNA promotes tumor metastasis. *Biochim. Biophys. Acta* *1843*, 1414–1426.

- Matranga, C.**, Tomari, Y., Shin, C., Bartel, D.P., and Zamore, P.D. (2005). Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell* 123, 607–620.
- Mavaddat, N.**, Antoniou, A.C., Easton, D.F., and Garcia-Closas, M. (2010). Genetic susceptibility to breast cancer. *Mol Oncol* 4, 174–191.
- McGrath, J.**, and Solter, D. (1984a). Maternal Thp lethality in the mouse is a nuclear, not cytoplasmic, defect. *Nature* 308, 550–551.
- McGrath, J.**, and Solter, D. (1984b). Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* 37, 179–183.
- 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, 185–197.
- Melo, S.A.**, Moutinho, C., Ropero, S., Calin, G.A., Rossi, S., Spizzo, R., Fernandez, A.F., Davalos, V., Villanueva, A., Montoya, G., et al. (2010). A genetic defect in exportin-5 traps precursor microRNAs in the nucleus of cancer cells. *Cancer Cell* 18, 303–315.
- Memczak, S.**, Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S.D., Gregersen, L.H., Munschauer, M., et al. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 495, 333–338.
- Miller, E.**, Lee, H.J., Lulla, A., Hernandez, L., Gokare, P., and Lim, B. (2014). Current treatment of early breast cancer: adjuvant and neoadjuvant therapy. *F1000Res* 3, 198.
- Miller, K.**, Wang, M., Gralow, J., Dickler, M., Cobleigh, M., Perez, E.A., Shenkier, T., Cella, D., and Davidson, N.E. (2007). Paclitaxel plus bevacizumab versus paclitaxel alone for metastatic breast cancer. *N. Engl. J. Med.* 357, 2666–2676.
- Milligan, L.**, Antoine, E., Bisbal, C., Weber, M., Brunel, C., Forné, T., and Cathala, G. (2000). H19 gene expression is up-regulated exclusively by stabilization of the RNA during muscle cell differentiation. *Oncogene* 19, 5810–5816.
- Mishima, Y.**, Giraldez, A.J., Takeda, Y., Fujiwara, T., Sakamoto, H., Schier, A.F., and Inoue, K. (2006). Differential regulation of germline mRNAs in soma and germ cells by zebrafish miR-430. *Curr. Biol.* 16, 2135–2142.
- Miyoshi, K.**, Tsukumo, H., Nagami, T., Siomi, H., and Siomi, M.C. (2005). Slicer function of *Drosophila* Argonautes and its involvement in RISC formation. *Genes Dev.* 19, 2837–2848.

- Moasser, M.M.** (2007). The oncogene HER2: its signaling and transforming functions and its role in human cancer pathogenesis. *Oncogene* 26, 6469–6487.
- Monnier, P.**, Martinet, C., Pontis, J., Stancheva, I., Ait-Si-Ali, S., and Dandolo, L. (2013). H19 lncRNA controls gene expression of the Imprinted Gene Network by recruiting MBD1. *Proc. Natl. Acad. Sci. U.S.A.* 110, 20693–20698.
- Monteys, A.M.**, Spengler, R.M., Wan, J., Tecedor, L., Lennox, K.A., Xing, Y., and Davidson, B.L. (2010). Structure and activity of putative intronic miRNA promoters. *RNA* 16, 495–505.
- 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, 12509–12514.
- Moorehead, R.A.**, Sanchez, O.H., Baldwin, R.M., and Khokha, R. (2003). Transgenic overexpression of IGF-II induces spontaneous lung tumors: a model for human lung adenocarcinoma. *Oncogene* 22, 853–857.
- Moulton, T.**, Crenshaw, T., Hao, Y., Moosikasawan, J., Lin, N., Dembitzer, F., Hensle, T., Weiss, L., McMorro, L., and Loew, T. (1994). Epigenetic lesions at the H19 locus in Wilms' tumour patients. *Nat. Genet.* 7, 440–447.
- Mourelatos, Z.**, Dostie, J., Paushkin, S., Sharma, A., Charroux, B., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev.* 16, 720–728.
- Mourtada-Maarabouni, M.**, Pickard, M.R., Hedge, V.L., Farzaneh, F., and Williams, G.T. (2009). GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. *Oncogene* 28, 195–208.
- Mousavi, K.**, Zare, H., Dell'orso, S., Grontved, L., Gutierrez-Cruz, G., Derfoul, A., Hager, G.L., and Sartorelli, V. (2013). eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci. *Mol. Cell* 51, 606–617.
- Mu, X.**, Zhang, Y., Qu, X., Hou, K., Kang, J., Hu, X., and Liu, Y. (2013). Ubiquitin ligase Cbl-b is involved in icotinib (BPI-2009H)-induced apoptosis and G1 phase arrest of EGFR mutation-positive non-small-cell lung cancer. *Biomed Res Int* 2013, 726375.
- Mueller, M.M.**, and Fusenig, N.E. (2004). Friends or foes - bipolar effects of the tumour stroma in cancer. *Nat. Rev. Cancer* 4, 839–849.

- Murchison, E.P.**, Partridge, J.F., Tam, O.H., Cheloufi, S., and Hannon, G.J. (2005). Characterization of Dicer-deficient murine embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* *102*, 12135–12140.
- Murphy, S.K.**, Huang, Z., Wen, Y., Spillman, M.A., Whitaker, R.S., Simel, L.R., Nichols, T.D., Marks, J.R., and Berchuck, A. (2006). Frequent IGF2/H19 domain epigenetic alterations and elevated IGF2 expression in epithelial ovarian cancer. *Mol. Cancer Res.* *4*, 283–292.
- Murrell, A.**, Heeson, S., Bowden, L., Constância, M., Dean, W., Kelsey, G., and Reik, W. (2001). An intragenic methylated region in the imprinted *Igf2* gene augments transcription. *EMBO Rep.* *2*, 1101–1106.
- 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*, 889–893.
- Naidu, S.**, Magee, P., and Garofalo, M. (2015). MiRNA-based therapeutic intervention of cancer. *J Hematol Oncol* *8*, 68.
- Nakagawa, S.**, Ip, J.Y., Shioi, G., Tripathi, V., Zong, X., Hirose, T., and Prasanth, K.V. (2012). *Malat1* is not an essential component of nuclear speckles in mice. *RNA* *18*, 1487–1499.
- Nativio, R.**, Wendt, K.S., Ito, Y., Huddleston, J.E., Uribe-Lewis, S., Woodfine, K., Krueger, C., Reik, W., Peters, J.-M., and Murrell, A. (2009). Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. *PLoS Genet.* *5*, e1000739.
- Nielsen, F.C.**, Nielsen, J., and Christiansen, J. (2001). A family of IGF-II mRNA binding proteins (IMP) involved in RNA trafficking. *Scand. J. Clin. Lab. Invest. Suppl.* *234*, 93–99.
- Nkondjock, A.**, and Ghadirian, P. (2005). [Risk factors and risk reduction of breast cancer]. *Med Sci (Paris)* *21*, 175–180.
- Nottrott, S.**, Simard, M.J., and Richter, J.D. (2006). Human let-7a miRNA blocks protein production on actively translating polyribosomes. *Nat. Struct. Mol. Biol.* *13*, 1108–1114.
- Obeid, E.**, Nanda, R., Fu, Y.-X., and Olopade, O.I. (2013). The role of tumor-associated macrophages in breast cancer progression (review). *Int. J. Oncol.* *43*, 5–12.
- Ogawa, Y.**, Sun, B.K., and Lee, J.T. (2008). Intersection of the RNA interference and X-inactivation pathways. *Science* *320*, 1336–1341.

- Ohlsson, R.**, Hedborg, F., Holmgren, L., Walsh, C., and Ekström, T.J. (1994). Overlapping patterns of IGF2 and H19 expression during human development: biallelic IGF2 expression correlates with a lack of H19 expression. *Development* 120, 361–368.
- Okada, C.**, Yamashita, E., Lee, S.J., Shibata, S., Katahira, J., Nakagawa, A., Yoneda, Y., and Tsukihara, T. (2009). A high-resolution structure of the pre-microRNA nuclear export machinery. *Science* 326, 1275–1279.
- Okamura, K.**, Ishizuka, A., Siomi, H., and Siomi, M.C. (2004). Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes Dev.* 18, 1655–1666.
- Okamura, K.**, Hagen, J.W., Duan, H., Tyler, D.M., and Lai, E.C. (2007). The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. *Cell* 130, 89–100.
- Okamura, K.**, Liu, N., and Lai, E.C. (2009). Distinct mechanisms for microRNA strand selection by *Drosophila* Argonautes. *Mol. Cell* 36, 431–444.
- Okano, M.**, Xie, S., and Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat. Genet.* 19, 219–220.
- Okutsu, T.**, Kuroiwa, Y., Kagitani, F., Kai, M., Aisaka, K., Tsutsumi, O., Kaneko, Y., Yokomori, K., Surani, M.A., Kohda, T., et al. (2000). Expression and imprinting status of human PEG8/IGF2AS, a paternally expressed antisense transcript from the IGF2 locus, in Wilms' tumors. *J. Biochem.* 127, 475–483.
- Olsen, P.H.**, and Ambros, V. (1999). The lin-4 regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation. *Dev. Biol.* 216, 671–680.
- Onyango, P.**, and Feinberg, A.P. (2011). A nucleolar protein, H19 opposite tumor suppressor (HOTS), is a tumor growth inhibitor encoded by a human imprinted H19 antisense transcript. *Proc. Natl. Acad. Sci. U.S.A.* 108, 16759–16764.
- Orimo, A.**, Gupta, P.B., Sgroi, D.C., Arenzana-Seisdedos, F., Delaunay, T., Naeem, R., Carey, V.J., Richardson, A.L., and Weinberg, R.A. (2005). Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. *Cell* 121, 335–348.
- Ørom, U.A.**, Nielsen, F.C., and Lund, A.H. (2008). MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Mol. Cell* 30, 460–471.

Ørom, U.A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G., Lai, F., Zytznicki, M., Notredame, C., Huang, Q., et al. (2010). Long noncoding RNAs with enhancer-like function in human cells. *Cell* 143, 46–58.

Osborne, C., Wilson, P., and Tripathy, D. (2004). Oncogenes and tumor suppressor genes in breast cancer: potential diagnostic and therapeutic applications. *Oncologist* 9, 361–377.

O’Shaughnessy, J., Osborne, C., Pippen, J.E., Yoffe, M., Patt, D., Rocha, C., Koo, I.C., Sherman, B.M., and Bradley, C. (2011). Iniparib plus chemotherapy in metastatic triple-negative breast cancer. *N. Engl. J. Med.* 364, 205–214.

Ozsolak, F., Poling, L.L., Wang, Z., Liu, H., Liu, X.S., Roeder, R.G., Zhang, X., Song, J.S., and Fisher, D.E. (2008). Chromatin structure analyses identify miRNA promoters. *Genes Dev.* 22, 3172–3183.

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, 5523–5527.

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, 673–681.

Park, J.-E., Heo, I., Tian, Y., Simanshu, D.K., Chang, H., Jee, D., Patel, D.J., and Kim, V.N. (2011). Dicer recognizes the 5' end of RNA for efficient and accurate processing. *Nature* 475, 201–205.

Parker, J.S., Roe, S.M., and Barford, D. (2005). Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex. *Nature* 434, 663–666.

Perincheri, S., Dingle, R.W.C., Peterson, M.L., and Spear, B.T. (2005). Hereditary persistence of alpha-fetoprotein and H19 expression in liver of BALB/cJ mice is due to a retrovirus insertion in the Zfx2 gene. *Proc. Natl. Acad. Sci. U.S.A.* 102, 396–401.

Perou, C.M., Sørlie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross, D.T., Johnsen, H., Akslen, L.A., et al. (2000). Molecular portraits of human breast tumours. *Nature* 406, 747–752.

Pillai, R.S., Bhattacharyya, S.N., Artus, C.G., Zoller, T., Cougot, N., Basyuk, E., Bertrand, E., and Filipowicz, W. (2005). Inhibition of translational initiation by Let-7 MicroRNA in human cells. *Science* 309, 1573–1576.

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, 1105–1114.

- Poliseno, L.**, Salmena, L., Zhang, J., Carver, B., Haveman, W.J., and Pandolfi, P.P. (2010). A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* 465, 1033–1038.
- Polyak, K.**, and Kalluri, R. (2010). The role of the microenvironment in mammary gland development and cancer. *Cold Spring Harb Perspect Biol* 2, a003244.
- Pravtcheva, D.D.**, and Wise, T.L. (1998). Metastasizing mammary carcinomas in H19 enhancers-Igf2 transgenic mice. *J. Exp. Zool.* 281, 43–57.
- Rainier, S.**, Johnson, L.A., Dobry, C.J., Ping, A.J., Grundy, P.E., and Feinberg, A.P. (1993). Relaxation of imprinted genes in human cancer. *Nature* 362, 747–749.
- Rand, T.A.**, Petersen, S., Du, F., and Wang, X. (2005). Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. *Cell* 123, 621–629.
- Raveh, E.**, Matouk, I.J., Gilon, M., and Hochberg, A. (2015). The H19 Long non-coding RNA in cancer initiation, progression and metastasis - a proposed unifying theory. *Mol. Cancer* 14, 184.
- Reeve, A.E.** (1996). Role of genomic imprinting in Wilms' tumour and overgrowth disorders. *Med. Pediatr. Oncol.* 27, 470–475.
- Reik, W.** (1989). Genomic imprinting and genetic disorders in man. *Trends Genet.* 5, 331–336.
- Reik, W.**, Brown, K.W., Schneid, H., Le Bouc, Y., Bickmore, W., and Maher, E.R. (1995). Imprinting mutations in the Beckwith-Wiedemann syndrome suggested by altered imprinting pattern in the IGF2-H19 domain. *Hum. Mol. Genet.* 4, 2379–2385.
- Reya, T.**, Morrison, S.J., Clarke, M.F., and Weissman, I.L. (2001). Stem cells, cancer, and cancer stem cells. *Nature* 414, 105–111.
- Rinn, J.L.**, Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Bruggmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E., et al. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 129, 1311–1323.
- 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, 1596–1604.
- Rogler, C.E.**, Yang, D., Rossetti, L., Donohoe, J., Alt, E., Chang, C.J., Rosenfeld, R., Neely, K., and Hintz, R. (1994). Altered body composition and increased frequency of diverse malignancies in insulin-like growth factor-II transgenic mice. *J. Biol. Chem.* 269, 13779–13784.

- Rohrschneider, L.R.**, Custodio, J.M., Anderson, T.A., Miller, C.P., and Gu, H. (2005). The intron 5/6 promoter region of the *ship1* gene regulates expression in stem/progenitor cells of the mouse embryo. *Dev. Biol.* 283, 503–521.
- Rose, D.P.**, and Vona-Davis, L. (2014). Biochemical and molecular mechanisms for the association between obesity, chronic inflammation, and breast cancer. *Biofactors* 40, 1–12.
- Rotter, V.**, Abutbul, H., and Ben-Ze'ev, A. (1983). P53 transformation-related protein accumulates in the nucleus of transformed fibroblasts in association with the chromatin and is found in the cytoplasm of non-transformed fibroblasts. *EMBO J.* 2, 1041–1047.
- Rougeulle, C.**, Cardoso, C., Fontés, M., Colleaux, L., and Lalande, M. (1998). An imprinted antisense RNA overlaps UBE3A and a second maternally expressed transcript. *Nat. Genet.* 19, 15–16.
- Roy, P.G.**, and Thompson, A.M. (2006). Cyclin D1 and breast cancer. *Breast* 15, 718–727.
- Rubio, E.D.**, Reiss, D.J., Welch, P.L., Distech, C.M., Filippova, G.N., Baliga, N.S., Aebersold, R., Ranish, J.A., and Krumm, A. (2008). CTCF physically links cohesin to chromatin. *Proc. Natl. Acad. Sci. U.S.A.* 105, 8309–8314.
- Ruby, J.G.**, Jan, C.H., and Bartel, D.P. (2007). Intronic microRNA precursors that bypass Drosha processing. *Nature* 448, 83–86.
- 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, 29562–29569.
- Runte, M.**, Hüttenhofer, A., Gross, S., Kiefmann, M., Horsthemke, B., and Buiting, K. (2001). The IC-SNURF-SNRPN transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for UBE3A. *Hum. Mol. Genet.* 10, 2687–2700.
- Sandhu, S.K.**, Volinia, S., Costinean, S., Galasso, M., Neinast, R., Santhanam, R., Parthun, M.R., Perrotti, D., Marcucci, G., Garzon, R., et al. (2012). miR-155 targets histone deacetylase 4 (HDAC4) and impairs transcriptional activity of B-cell lymphoma 6 (BCL6) in the E μ -miR-155 transgenic mouse model. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20047–20052.
- Sanli, I.**, and Feil, R. (2015). Chromatin mechanisms in the developmental control of imprinted gene expression. *Int. J. Biochem. Cell Biol.*
- Santanam, U.**, Zanesi, N., Efanov, A., Costinean, S., Palamarchuk, A., Hagan, J.P., Volinia, S., Alder, H., Rassenti, L., Kipps, T., et al. (2010). Chronic lymphocytic leukemia modeled in mouse by targeted miR-29 expression. *Proc. Natl. Acad. Sci. U.S.A.* 107, 12210–12215.

- Scheuer, W.**, Friess, T., Burtcher, H., Bossenmaier, B., Endl, J., and Hasmann, M. (2009). Strongly enhanced antitumor activity of trastuzumab and pertuzumab combination treatment on HER2-positive human xenograft tumor models. *Cancer Res.* 69, 9330–9336.
- Schwarz, D.S.**, Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P.D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 115, 199–208.
- Searle, A.G.**, and Beechey, C.V. (1978). Complementation studies with mouse translocations. *Cytogenet. Cell Genet.* 20, 282–303.
- Serra, V.**, Markman, B., Scaltriti, M., Eichhorn, P.J.A., Valero, V., Guzman, M., Botero, M.L., Llouch, E., Atzori, F., Di Cosimo, S., et al. (2008). NVP-BEZ235, a dual PI3K/mTOR inhibitor, prevents PI3K signaling and inhibits the growth of cancer cells with activating PI3K mutations. *Cancer Res.* 68, 8022–8030.
- Seto, E.**, and Yoshida, M. (2014). Erasers of histone acetylation: the histone deacetylase enzymes. *Cold Spring Harb Perspect Biol* 6, a018713.
- Shi, X.**, Sun, M., Liu, H., Yao, Y., and Song, Y. (2013). Long non-coding RNAs: a new frontier in the study of human diseases. *Cancer Lett.* 339, 159–166.
- Shiloh, Y.**, and Ziv, Y. (2013). The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. *Nat. Rev. Mol. Cell Biol.* 14, 197–210.
- 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, 168–178.
- Shipitsin, M.**, and Polyak, K. (2008). The cancer stem cell hypothesis: in search of definitions, markers, and relevance. *Lab. Invest.* 88, 459–463.
- Shore, A.N.**, and Rosen, J.M. (2014). Regulation of mammary epithelial cell homeostasis by lncRNAs. *Int. J. Biochem. Cell Biol.* 54, 318–330.
- Shore, A.N.**, Herschkowitz, J.I., and Rosen, J.M. (2012). Noncoding RNAs involved in mammary gland development and tumorigenesis: there's a long way to go. *J Mammary Gland Biol Neoplasia* 17, 43–58.
- Sicard, F.**, Gayral, M., Lulka, H., Buscail, L., and Cordelier, P. (2013). Targeting miR-21 for the therapy of pancreatic cancer. *Mol. Ther.* 21, 986–994.
- Silva, J.M.**, Boczek, N.J., Berres, M.W., Ma, X., and Smith, D.I. (2011). LSINCT5 is over expressed in breast and ovarian cancer and affects cellular proliferation. *RNA Biol* 8, 496–505.

- Sleutels, F.**, Zwart, R., and Barlow, D.P. (2002). The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* 415, 810–813.
- Smits, G.**, Mungall, A.J., Griffiths-Jones, S., Smith, P., Beury, D., Matthews, L., Rogers, J., Pask, A.J., Shaw, G., VandeBerg, J.L., et al. (2008). Conservation of the H19 noncoding RNA and H19-IGF2 imprinting mechanism in therians. *Nat. Genet.* 40, 971–976.
- Sohn, S.Y.**, Bae, W.J., Kim, J.J., Yeom, K.-H., Kim, V.N., and Cho, Y. (2007). Crystal structure of human DGCR8 core. *Nat. Struct. Mol. Biol.* 14, 847–853.
- Song, J.-J.**, Smith, S.K., Hannon, G.J., and Joshua-Tor, L. (2004). Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* 305, 1434–1437.
- Song, Q.**, Xu, Y., Yang, C., Chen, Z., Jia, C., Chen, J., Zhang, Y., Lai, P., Fan, X., Zhou, X., et al. (2014). miR-483-5p promotes invasion and metastasis of lung adenocarcinoma by targeting RhoGDI1 and ALCAM. *Cancer Res.* 74, 3031–3042.
- Soria, G.**, and Ben-Baruch, A. (2008). The inflammatory chemokines CCL2 and CCL5 in breast cancer. *Cancer Lett.* 267, 271–285.
- 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, 1186–1195.
- Stefani, G.**, and Slack, F.J. (2008). Small non-coding RNAs in animal development. *Nat. Rev. Mol. Cell Biol.* 9, 219–230.
- Stendahl, M.**, Kronblad, A., Rydén, L., Emdin, S., Bengtsson, N.O., and Landberg, G. (2004). Cyclin D1 overexpression is a negative predictive factor for tamoxifen response in postmenopausal breast cancer patients. *Br. J. Cancer* 90, 1942–1948.
- Su, H.**, Trombly, M.I., Chen, J., and Wang, X. (2009). Essential and overlapping functions for mammalian Argonautes in microRNA silencing. *Genes Dev.* 23, 304–317.
- Sun, S.**, Del Rosario, B.C., Szanto, A., Ogawa, Y., Jeon, Y., and Lee, J.T. (2013). Jpx RNA activates Xist by evicting CTCF. *Cell* 153, 1537–1551.
- Sundaram, G.M.**, Common, J.E.A., Gopal, F.E., Srikanta, S., Lakshman, K., Lunny, D.P., Lim, T.C., Tanavde, V., Lane, E.B., and Sampath, P. (2013). “See-saw” expression of microRNA-198 and FSTL1 from a single transcript in wound healing. *Nature* 495, 103–106.
- Sunwoo, H.**, Dinger, M.E., Wilusz, J.E., Amaral, P.P., Mattick, J.S., and Spector, D.L. (2009). MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res.* 19, 347–359.

- Surani, M.A.**, Barton, S.C., and Norris, M.L. (1984). Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* 308, 548–550.
- Szabó, 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, 607–610.
- Tabara, H.**, Sarkissian, M., Kelly, W.G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C.C. (1999). The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. *Cell* 99, 123–132.
- Taganov, K.D.**, Boldin, M.P., Chang, K.-J., and Baltimore, D. (2006). NF-kappaB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12481–12486.
- Takai, D.**, Gonzales, F.A., Tsai, Y.C., Thayer, M.J., and Jones, P.A. (2001). Large scale mapping of methylcytosines in CTCF-binding sites in the human H19 promoter and aberrant hypomethylation in human bladder cancer. *Hum. Mol. Genet.* 10, 2619–2626.
- Tanos, V.**, Prus, D., Ayeshe, S., Weinstein, D., Tykocinski, M.L., De-Groot, N., Hochberg, A., and Ariel, I. (1999). Expression of the imprinted H19 oncofetal RNA in epithelial ovarian cancer. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 85, 7–11.
- Tanos, V.**, Ariel, I., Prus, D., De-Groot, N., and Hochberg, A. (2004). H19 and IGF2 gene expression in human normal, hyperplastic, and malignant endometrium. *Int. J. Gynecol. Cancer* 14, 521–525.
- Taunton, J.**, Hassig, C.A., and Schreiber, S.L. (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* 272, 408–411.
- Taylor, E.R.**, Seleiro, E.A., and Brickell, P.M. (1991). Identification of antisense transcripts of the chicken insulin-like growth factor-II gene. *J. Mol. Endocrinol.* 7, 145–154.
- Tessier, C.R.**, Doyle, G.A., Clark, B.A., Pitot, H.C., and Ross, J. (2004). Mammary tumor induction in transgenic mice expressing an RNA-binding protein. *Cancer Res.* 64, 209–214.
- Tétreault, N.**, and De Guire, V. (2013). miRNAs: their discovery, biogenesis and mechanism of action. *Clin. Biochem.* 46, 842–845.
- Thompson, D.**, and Easton, D. (2004). The genetic epidemiology of breast cancer genes. *J Mammary Gland Biol Neoplasia* 9, 221–236.
- Thompson, D.**, Easton, D., and Breast Cancer Linkage Consortium (2001). Variation in cancer risks, by mutation position, in BRCA2 mutation carriers. *Am. J. Hum. Genet.* 68, 410–419.

- Tian, Y.**, Simanshu, D.K., Ma, J.-B., Park, J.-E., Heo, I., Kim, V.N., and Patel, D.J. (2014). A phosphate-binding pocket within the platform-PAZ-connector helix cassette of human Dicer. *Mol. Cell* 53, 606–616.
- Toillon, R.-A.**, Lagadec, C., Page, A., Chopin, V., Sautière, P.-E., Ricort, J.-M., Lemoine, J., Zhang, M., Hondermarck, H., and Le Bourhis, X. (2007). Proteomics demonstration that normal breast epithelial cells can induce apoptosis of breast cancer cells through insulin-like growth factor-binding protein-3 and maspin. *Mol. Cell Proteomics* 6, 1239–1247.
- Toledo, F.**, and Wahl, G.M. (2007). MDM2 and MDM4: p53 regulators as targets in anticancer therapy. *Int. J. Biochem. Cell Biol.* 39, 1476–1482.
- Toledo, F.**, Krummel, K.A., Lee, C.J., Liu, C.-W., Rodewald, L.-W., Tang, M., and Wahl, G.M. (2006). A mouse p53 mutant lacking the proline-rich domain rescues Mdm4 deficiency and provides insight into the Mdm2-Mdm4-p53 regulatory network. *Cancer Cell* 9, 273–285.
- Tomao, F.**, Papa, A., Zaccarelli, E., Rossi, L., Caruso, D., Minozzi, M., Vici, P., Frati, L., and Tomao, S. (2015). Triple-negative breast cancer: new perspectives for targeted therapies. *Onco Targets Ther* 8, 177–193.
- Tomari, Y.**, Du, T., Haley, B., Schwarz, D.S., Bennett, R., Cook, H.A., Koppetsch, B.S., Theurkauf, W.E., and Zamore, P.D. (2004). RISC assembly defects in the *Drosophila* RNAi mutant armitage. *Cell* 116, 831–841.
- Tomari, Y.**, Du, T., and Zamore, P.D. (2007). Sorting of *Drosophila* small silencing RNAs. *Cell* 130, 299–308.
- Tran, V.G.**, Court, F., Duputié, A., Antoine, E., Aptel, N., Milligan, L., Carbonell, F., Lelay-Taha, M.-N., Piette, J., Weber, M., et al. (2012). H19 antisense RNA can up-regulate Igf2 transcription by activation of a novel promoter in mouse myoblasts. *PLoS ONE* 7, e37923.
- Trang, P.**, Wiggins, J.F., Daige, C.L., Cho, C., Omotola, M., Brown, D., Weidhaas, J.B., Bader, A.G., and Slack, F.J. (2011). Systemic delivery of tumor suppressor microRNA mimics using a neutral lipid emulsion inhibits lung tumors in mice. *Mol. Ther.* 19, 1116–1122.
- 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, 4322–4329.
- Tripathi, V.**, Ellis, J.D., Shen, Z., Song, D.Y., Pan, Q., Watt, A.T., Freier, S.M., Bennett, C.F., Sharma, A., Bubulya, P.A., et al. (2010). The nuclear-retained noncoding RNA MALAT1

regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell* 39, 925–938.

Tsai, M.-C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J.K., Lan, F., Shi, Y., Segal, E., and Chang, H.Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 329, 689–693.

Tsang, W.P., Wong, T.W.L., Cheung, A.H.H., Co, C.N.N., and Kwok, T.T. (2007). Induction of drug resistance and transformation in human cancer cells by the noncoding RNA CUDR. *RNA* 13, 890–898.

Tsang, W.P., Ng, E.K.O., Ng, S.S.M., Jin, H., Yu, J., Sung, J.J.Y., and Kwok, T.T. (2010). Oncofetal H19-derived miR-675 regulates tumor suppressor RB in human colorectal cancer. *Carcinogenesis* 31, 350–358.

Tsutsumi, A., Kawamata, T., Izumi, N., Seitz, H., and Tomari, Y. (2011). Recognition of the pre-miRNA structure by *Drosophila* Dicer-1. *Nat. Struct. Mol. Biol.* 18, 1153–1158.

Vanhecke, E., Adriaenssens, E., Verbeke, S., Meignan, S., Germain, E., Berteaux, N., Nurcombe, V., Le Bourhis, X., and Hondermarck, H. (2011). Brain-derived neurotrophic factor and neurotrophin-4/5 are expressed in breast cancer and can be targeted to inhibit tumor cell survival. *Clin. Cancer Res.* 17, 1741–1752.

Vasudevan, S., Tong, Y., and Steitz, J.A. (2007). Switching from repression to activation: microRNAs can up-regulate translation. *Science* 318, 1931–1934.

Venkatraman, A., He, X.C., Thorvaldsen, J.L., Sugimura, R., Perry, J.M., Tao, F., Zhao, M., Christenson, M.K., Sanchez, R., Yu, J.Y., et al. (2013). Maternal imprinting at the H19-Igf2 locus maintains adult haematopoietic stem cell quiescence. *Nature* 500, 345–349.

Venkitaraman, A.R. (2002). Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell* 108, 171–182.

Vennin, C., Dahmani, F., Spruyt, N., and Adriaenssens, E. (2013). Role of long non-coding RNA in cells: Example of the H19/IGF2 locus. *Advances in Bioscience and Biotechnology* 04, 34–44.

Vennin, C., Spruyt, N., Dahmani, F., Julien, S., Bertucci, F., Finetti, P., Chassat, T., Bourette, R.P., Le Bourhis, X., and Adriaenssens, E. (2015). H19 non coding RNA-derived miR-675 enhances tumorigenesis and metastasis of breast cancer cells by downregulating c-Cbl and Cbl-b. *Oncotarget* 6, 29209–29223.

- 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.
- Veronese, A.**, Lupini, L., Consiglio, J., Visone, R., Ferracin, M., Fornari, F., Zanesi, N., Alder, H., D’Elia, G., Gramantieri, L., et al. (2010). Oncogenic role of miR-483-3p at the IGF2/483 locus. *Cancer Res.* *70*, 3140–3149.
- Vilmont, V.**, Filhol, O., Hesse, A.-M., Couté, Y., Hue, C., Rémy-Tourneur, L., Mistou, S., Cochet, C., and Chiocchia, G. (2015). Modulatory role of the anti-apoptotic protein kinase CK2 in the sub-cellular localization of Fas associated death domain protein (FADD). *Biochim. Biophys. Acta* *1853*, 2885–2896.
- Vinall, R.L.**, Ripoll, A.Z., Wang, S., Pan, C.-X., and deVere White, R.W. (2012). MiR-34a chemosensitizes bladder cancer cells to cisplatin treatment regardless of p53-Rb pathway status. *Int. J. Cancer* *130*, 2526–2538.
- Viré, E.**, Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J.-M., et al. (2006). The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* *439*, 871–874.
- Visvader, J.E.** (2011). Cells of origin in cancer. *Nature* *469*, 314–322.
- Visvader, J.E.**, and Stingl, J. (2014). Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes Dev.* *28*, 1143–1158.
- Volinia, S.**, Calin, G.A., Liu, C.-G., Ambs, S., Cimmino, A., Petrocca, F., Visone, R., Iorio, M., Roldo, C., Ferracin, M., et al. (2006). A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc. Natl. Acad. Sci. U.S.A.* *103*, 2257–2261.
- Völkel, P.**, and Angrand, P.-O. (2007). The control of histone lysine methylation in epigenetic regulation. *Biochimie* *89*, 1–20.
- Vona-Davis, L.**, and Rose, D.P. (2007). Adipokines as endocrine, paracrine, and autocrine factors in breast cancer risk and progression. *Endocr. Relat. Cancer* *14*, 189–206.
- Voutilainen, R.**, Ilvesmäki, V., Ariel, I., Rachmilewitz, J., de Groot, N., and Hochberg, A. (1994). Parallel regulation of parentally imprinted H19 and insulin-like growth factor-II genes in cultured human fetal adrenal cells. *Endocrinology* *134*, 2051–2056.
- Wagschal, A.**, Sutherland, H.G., Woodfine, K., Henckel, A., Chebli, K., Schulz, R., Oakey, R.J., Bickmore, W.A., and Feil, R. (2008). G9a histone methyltransferase contributes to imprinting in the mouse placenta. *Mol. Cell. Biol.* *28*, 1104–1113.

- Walter, M.**, Liang, S., Ghosh, S., Hornsby, P.J., and Li, R. (2009). Interleukin 6 secreted from adipose stromal cells promotes migration and invasion of breast cancer cells. *Oncogene* 28, 2745–2755.
- Wan, G.**, Zhang, X., Langley, R.R., Liu, Y., Hu, X., Han, C., Peng, G., Ellis, L.M., Jones, S.N., and Lu, X. (2013). DNA-damage-induced nuclear export of precursor microRNAs is regulated by the ATM-AKT pathway. *Cell Rep* 3, 2100–2112.
- Wang, K.C.**, Yang, Y.W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., Lajoie, B.R., Protacio, A., Flynn, R.A., Gupta, R.A., et al. (2011). A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 472, 120–124.
- Wang, L.-H.**, Chan, J.L.-K., and Li, W. (2007a). Rapamycin together with herceptin significantly increased anti-tumor efficacy compared to either alone in ErbB2 over expressing breast cancer cells. *Int. J. Cancer* 121, 157–164.
- Wang, X.**, Tang, S., Le, S.-Y., Lu, R., Rader, J.S., Meyers, C., and Zheng, Z.-M. (2008a). Aberrant expression of oncogenic and tumor-suppressive microRNAs in cervical cancer is required for cancer cell growth. *PLoS ONE* 3, e2557.
- Wang, X.-S.**, Zhang, Z., Wang, H.-C., Cai, J.-L., Xu, Q.-W., Li, M.-Q., Chen, Y.-C., Qian, X.-P., Lu, T.-J., Yu, L.-Z., et al. (2006). Rapid identification of UCA1 as a very sensitive and specific unique marker for human bladder carcinoma. *Clin. Cancer Res.* 12, 4851–4858.
- Wang, Y.**, Medvid, R., Melton, C., Jaenisch, R., and Blelloch, R. (2007b). DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nat. Genet.* 39, 380–385.
- Wang, Y.**, Sheng, G., Juranek, S., Tuschl, T., and Patel, D.J. (2008b). Structure of the guide-strand-containing argonaute silencing complex. *Nature* 456, 209–213.
- Wang, Y.**, Xu, Z., Jiang, J., Xu, C., Kang, J., Xiao, L., Wu, M., Xiong, J., Guo, X., and Liu, H. (2013). Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal. *Dev. Cell* 25, 69–80.
- Weber, M.**, Hagège, H., Murrell, A., Brunel, C., Reik, W., Cathala, G., and Forné, T. (2003). Genomic imprinting controls matrix attachment regions in the Igf2 gene. *Mol. Cell. Biol.* 23, 8953–8959.
- Wendt, K.S.**, Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T., et al. (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* 451, 796–801.

- Wiggins, J.F.**, Ruffino, L., Kelnar, K., Omotola, M., Patrawala, L., Brown, D., and Bader, A.G. (2010). Development of a lung cancer therapeutic based on the tumor suppressor microRNA-34. *Cancer Res.* *70*, 5923–5930.
- Wightman, B.**, Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* *75*, 855–862.
- Wilkin, F.**, Paquette, J., Ledru, E., Hamelin, C., Pollak, M., Deal, C.L., and Mamelin, C. (2000). H19 sense and antisense transgenes modify insulin-like growth factor-II mRNA levels. *Eur. J. Biochem.* *267*, 4020–4027.
- Wilusz, J.E.**, Freier, S.M., and Spector, D.L. (2008). 3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. *Cell* *135*, 919–932.
- Wu, H.K.**, Weksberg, R., Minden, M.D., and Squire, J.A. (1997). Loss of imprinting of human insulin-like growth factor II gene, *IGF2*, in acute myeloid leukemia. *Biochem. Biophys. Res. Commun.* *231*, 466–472.
- Wu, L.**, Fan, J., and Belasco, J.G. (2006). MicroRNAs direct rapid deadenylation of mRNA. *Proc. Natl. Acad. Sci. U.S.A.* *103*, 4034–4039.
- Wyckoff, J.**, Wang, W., Lin, E.Y., Wang, Y., Pixley, F., Stanley, E.R., Graf, T., Pollard, J.W., Segall, J., and Condeelis, J. (2004). A paracrine loop between tumor cells and macrophages is required for tumor cell migration in mammary tumors. *Cancer Res.* *64*, 7022–7029.
- Xie, M.**, Li, M., Vilborg, A., Lee, N., Shu, M.-D., Yartseva, V., Šestan, N., and Steitz, J.A. (2013). Mammalian 5'-capped microRNA precursors that generate a single microRNA. *Cell* *155*, 1568–1580.
- Xu, B.**, Gerin, I., Miao, H., Vu-Phan, D., Johnson, C.N., Xu, R., Chen, X.-W., Cawthorn, W.P., MacDougald, O.A., and Koenig, R.J. (2010a). Multiple roles for the non-coding RNA SRA in regulation of adipogenesis and insulin sensitivity. *PLoS ONE* *5*, e14199.
- Xu, J.**, Chen, Y., and Olopade, O.I. (2010b). MYC and Breast Cancer. *Genes Cancer* *1*, 629–640.
- Xu, N.**, Papagiannakopoulos, T., Pan, G., Thomson, J.A., and Kosik, K.S. (2009). MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. *Cell* *137*, 647–658.

- Xue, W.**, Dahlman, J.E., Tammela, T., Khan, O.F., Sood, S., Dave, A., Cai, W., Chirino, L.M., Yang, G.R., Bronson, R., et al. (2014). Small RNA combination therapy for lung cancer. *Proc. Natl. Acad. Sci. U.S.A.* *111*, E3553–E3561.
- Yang, J.-S.**, and Lai, E.C. (2011). Alternative miRNA biogenesis pathways and the interpretation of core miRNA pathway mutants. *Mol. Cell* *43*, 892–903.
- Yang, F.**, Bi, J., Xue, X., Zheng, L., Zhi, K., Hua, J., and Fang, G. (2012). Up-regulated long non-coding RNA H19 contributes to proliferation of gastric cancer cells. *FEBS J.* *279*, 3159–3165.
- Yang, P.**, Du, C.W., Kwan, M., Liang, S.X., and Zhang, G.J. (2013). The impact of p53 in predicting clinical outcome of breast cancer patients with visceral metastasis. *Sci Rep* *3*, 2246.
- Yang, J.-S.**, Maurin, T., Robine, N., Rasmussen, K.D., Jeffrey, K.L., Chandwani, R., Papapetrou, E.P., Sadelain, M., O’Carroll, D., and Lai, E.C. (2010). Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. *Proc. Natl. Acad. Sci. U.S.A.* *107*, 15163–15168.
- Ye, X.**, Huang, N., Liu, Y., Paroo, Z., Huerta, C., Li, P., Chen, S., Liu, Q., and Zhang, H. (2011). Structure of C3PO and mechanism of human RISC activation. *Nat. Struct. Mol. Biol.* *18*, 650–657.
- Yeom, K.-H.**, Lee, Y., Han, J., Suh, M.R., and Kim, V.N. (2006). Characterization of DGCR8/Pasha, the essential cofactor for Drosha in primary miRNA processing. *Nucleic Acids Res.* *34*, 4622–4629.
- Yoda, M.**, Kawamata, T., Paroo, Z., Ye, X., Iwasaki, S., Liu, Q., and Tomari, Y. (2010). ATP-dependent human RISC assembly pathways. *Nat. Struct. Mol. Biol.* *17*, 17–23.
- Yoda, M.**, Cifuentes, D., Izumi, N., Sakaguchi, Y., Suzuki, T., Giraldez, A.J., and Tomari, Y. (2013). Poly(A)-specific ribonuclease mediates 3’-end trimming of Argonaute2-cleaved precursor microRNAs. *Cell Rep* *5*, 715–726.
- Yoshimizu, T.**, Miroglio, A., Ripoché, M.-A., Gabory, A., Vernucci, M., Riccio, A., Colnot, S., Godard, C., Terris, B., Jammes, H., et al. (2008). The H19 locus acts in vivo as a tumor suppressor. *Proc. Natl. Acad. Sci. U.S.A.* *105*, 12417–12422.
- Yu, F.**, Yao, H., Zhu, P., Zhang, X., Pan, Q., Gong, C., Huang, Y., Hu, X., Su, F., Lieberman, J., et al. (2007). let-7 regulates self renewal and tumorigenicity of breast cancer cells. *Cell* *131*, 1109–1123.

Zeira, E., Abramovitch, R., Meir, K., Even Ram, S., Gil, Y., Bulvik, B., Bromberg, Z., Levkovitch, O., Nahmansson, N., Adar, R., et al. (2015). The knockdown of H19lncRNA reveals its regulatory role in pluripotency and tumorigenesis of human embryonic carcinoma cells. *Oncotarget* 6, 34691–34703.

Zeng, Y., Yi, R., and Cullen, B.R. (2005). Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. *EMBO J.* 24, 138–148.

Zhai, L.-L., Wang, P., Zhou, L.-Y., Yin, J.-Y., Tang, Q., Zhang, T.-J., Wang, Y.-X., Yang, D.-Q., Lin, J., and Deng, Z.-Q. (2015). Over-expression of miR-675 in formalin-fixed paraffin-embedded (FFPE) tissues of breast cancer patients. *Int J Clin Exp Med* 8, 11195–11201.

Zhang, A., Zhou, N., Huang, J., Liu, Q., Fukuda, K., Ma, D., Lu, Z., Bai, C., Watabe, K., and Mo, Y.-Y. (2013a). The human long non-coding RNA-RoR is a p53 repressor in response to DNA damage. *Cell Res.* 23, 340–350.

Zhang, B., Arun, G., Mao, Y.S., Lazar, Z., Hung, G., Bhattacharjee, G., Xiao, X., Booth, C.J., Wu, J., Zhang, C., et al. (2012). The lncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. *Cell Rep* 2, 111–123.

Zhang, E., Li, W., Yin, D., De, W., Sun, S., and Han, L. (2015). c-Myc-regulated long non-coding RNA H19 indicates a poor prognosis and affects cell proliferation in non-small-cell lung cancer. *Tumour Biol.*

Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E., and Filipowicz, W. (2004a). Single processing center models for human Dicer and bacterial RNase III. *Cell* 118, 57–68.

Zhang, H., Niu, B., Hu, J.-F., Ge, S., Wang, H., Li, T., Ling, J., Steelman, B.N., Qian, G., and Hoffman, A.R. (2011). Interruption of intrachromosomal looping by CCCTC binding factor decoy proteins abrogates genomic imprinting of human insulin-like growth factor II. *J. Cell Biol.* 193, 475–487.

Zhang, J., Zhang, D., and Hua, Z. (2004b). FADD and its phosphorylation. *IUBMB Life* 56, 395–401.

Zhang, S., Shan, C., Kong, G., Du, Y., Ye, L., and Zhang, X. (2012b). MicroRNA-520e suppresses growth of hepatoma cells by targeting the NF- κ B-inducing kinase (NIK). *Oncogene* 31, 3607–3620.

Zhang, Y., and Tycko, B. (1992). Monoallelic expression of the human H19 gene. *Nat. Genet.* 1, 40–44.

Zhang, Y., Zhang, X.-O., Chen, T., Xiang, J.-F., Yin, Q.-F., Xing, Y.-H., Zhu, S., Yang, L., and Chen, L.-L. (2013b). Circular intronic long noncoding RNAs. *Mol. Cell* 51, 792–806.

Zhao, J., Ohsumi, T.K., Kung, J.T., Ogawa, Y., Grau, D.J., Sarma, K., Song, J.J., Kingston, R.E., Borowsky, M., and Lee, J.T. (2010). Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol. Cell* 40, 939–953.

Zhao, R., DeCoteau, J.F., Geyer, C.R., Gao, M., Cui, H., and Casson, A.G. (2009). Loss of imprinting of the insulin-like growth factor II (IGF2) gene in esophageal normal and adenocarcinoma tissues. *Carcinogenesis* 30, 2117–2122.

Zhao, W.-G., Yu, S.-N., Lu, Z.-H., Ma, Y.-H., Gu, Y.-M., and Chen, J. (2010b). The miR-217 microRNA functions as a potential tumor suppressor in pancreatic ductal adenocarcinoma by targeting KRAS. *Carcinogenesis* 31, 1726–1733.

Zhou, W., Shi, G., Zhang, Q., Wu, Q., Li, B., and Zhang, Z. (2014a). MicroRNA-20b promotes cell growth of breast cancer cells partly via targeting phosphatase and tensin homologue (PTEN). *Cell Biosci* 4, 62.

Zhou, X., Gao, Q., Wang, J., Zhang, X., Liu, K., and Duan, Z. (2014b). Linc-RNA-RoR acts as a “sponge” against mediation of the differentiation of endometrial cancer stem cells by microRNA-145. *Gynecol. Oncol.* 133, 333–339.

Zhou, X., Ye, F., Yin, C., Zhuang, Y., Yue, G., and Zhang, G. (2015). The Interaction Between MiR-141 and lncRNA-H19 in Regulating Cell Proliferation and Migration in Gastric Cancer. *Cell. Physiol. Biochem.* 36, 1440–1452.

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

Zhu, M., Chen, Q., Liu, X., Sun, Q., Zhao, X., Deng, R., Wang, Y., Huang, J., Xu, M., Yan, J., et al. (2014). lncRNA H19/miR-675 axis represses prostate cancer metastasis by targeting TGFBI. *FEBS J.* 281, 3766–3775.

Zhuang, M., Gao, W., Xu, J., Wang, P., and Shu, Y. (2014). The long non-coding RNA H19-derived miR-675 modulates human gastric cancer cell proliferation by targeting tumor suppressor RUNX1. *Biochem. Biophys. Res. Commun.* 448, 315–322.

Zimmerman, D.L., Boddy, C.S., and Schoenherr, C.S. (2013). Oct4/Sox2 binding sites contribute to maintaining hypomethylation of the maternal igf2/h19 imprinting control region. *PLoS ONE* 8, e81962.

ANNEXE

Annexe 1

Constance Vennin, Fatima Dahmani, Nathalie Spruyt, Eric Adriaenssens. (2013) Role of long non coding RNA in cells: example of the *H19/IGF2* locus. *Advances in Bioscience and biotechnology*, 04, 34-44.

Role of long non coding RNA in cells: example of the *H19/IGF2* locus.

Constance Vennin, Fatima Dahmani, Nathalie Spruyt, Eric Adriaenssens.

UMR 8161 CNRS, Institut Biologie de Lille, University of Lille1, Lille, France.

Corresponding author : Eric Adriaenssens, UMR 8161 CNRS, Institut Biologie de Lille, 1 rue Pr Calmette, 59021 Lille Cedex. Tel : + 33 3 20 87 12 43, mail : eric.adriaenssens@ibl.fr

Key Words: H19 gene, Genomic imprinting, non coding RNA, Cell cycle

Abstract

In the past decade, studies of non coding RNAs increase. Non coding RNAs are divided in two classes: small and long non coding RNA. It was shown that long non coding RNAs regulate expression of 70% of genes. Long non coding RNAs are involved in several cellular processes like epigenetic regulation, dosage compensation, alternative splicing and stem cells maintenance for example. Misregulations of their expression induce diseases such as syndrome or cancer. Here, we describe some functions of long non coding RNA in cells. Furthermore, we study the *H19/IGF2* cluster: an imprinted genomic locus located on chromosome 11p15.5. Genomic imprinting allows gene expression from a single allele in a parent-origin-dependent manner. These cluster encode for the first long non coding RNA identified: *H19*. In 1990, it was established that *H19* functions as a *riboregulator*. Recently, it was shown that *H19* is a precursor of microRNA (hsa-miR-675), and several new transcripts were identified at the *H19/IGF2* locus. So, the complexity of this locus increasing, in this review, we summarize our current understanding about the *H19/IGF2* cluster both in terms of transcription as well as in terms of functions in cells. We highlight the involvement of *H19*, its new antisense transcript *91H* and its microRNA, in the regulation of IGF1 receptor and ligand and cell cycle progression.

Recently, the ENCODE project have proved that approximately 90% of the genome is transcribed. In cells, about 2% of sequences coding for protein, the others transcripts act as introns or non-coding RNA. These non-coding transcripts are composed of non-protein coding gene and background of transcription. This transcriptional noise would allow maintaining chromatin in a conformation favourable to the transcription of protein-coding gene. In the past decade, studies of non-protein coding gene, also called non-coding RNAs (ncRNAs), have increased and revealed that they are important in cells. NcRNAs are highly regulated and have several functions for normal development of tissues to tumorigenesis. Non-coding RNAs are divided in two classes of RNAs: the small and the long non-coding RNAs.

Small non-coding RNA (<200nts), notably microRNAs (miRNAs), are involved in post-transcriptional regulation of RNA. MiRNAs are the most studied of the small ncRNAs because they regulate 60% of protein-coding gene. Mostly, they are encoded by the genome, transcribed by RNA polymerase II, cleaved by Drosha in the nucleus and by Dicer in the cytosol. Then, they interact with Argonaut protein family to form the RISC complex (RNA-induced silencing complex). This complex allows the function of miRNA [1]. They induce silencing of gene by degrading or inhibiting the translation of mRNA into protein. They are involved in all process of the cell: proliferation, cell cycle progression, differentiation, and apoptosis. Two others small non-coding RNAs are described: piRNA (PIWI-interacting RNA) and snoRNA (small nucleolar RNAs). piRNAs have recently been discovered. They are important for germline development and for the suppression of transposon activity in germline cells. SnoRNA are components of ribonucleoproteins (snoRNPs). They are responsible for post-transcriptional modifications of rRNA that take place in the nucleolus (the nucleus compartment where ribosomes are formed) [1].

Long non coding RNAs (LncRNAs (>200 nts)) can have different origins: mRNA, antisense transcripts or long-intergenic non-coding RNA (lincRNA). LincRNAs are produced by sequence between two genes or set of genes. In Homo sapiens, 4500 lincRNAs have discovered but only 108 lincRNAs are studied and published [2]. LncRNAs have an important place in cell as they regulate expression of 70% of genes. They are involved, for example, in epigenetic regulation, alternative splice or stem cells maintenance. Misregulations of their expression induce diseases such as cancer.

The first imprinting lincRNA identified is the *H19* RNA. *H19* is transcript from the *H19/IGF2* cluster located on the chromosome 11p15.5. Every day, the complexity of this locus increases: recently, several news transcripts have been identified.

In this review, firstly, we describe some functions of long non coding RNA in cells. Then, to illustrate our point, we summarize our current understanding about the *H19/IGF2* cluster, its complexity and its function in cells.

I. Long non coding RNA functions

1. Dosage compensation

The lncRNA *Xist* (X-inactive-specific-transcript), a lncRNA of 17.000 nucleotides, is the most studied. This lncRNA is implicated in the X chromosome inactivation in dosage compensation [3]. *Xist* is encoded by the X chromosome and acts in *cis*. Thanks to its conserved repeat motif RepA, *Xist* interacts with the Polycomb repressive complex (PRC2), the complex responsible for trimethylation of histone H3 at Lys²⁷, and targets this complex to the XIC (X-inactivation center). PRC2 complex induces histone modifications, heterochromatin formation and silencing of the targeted X chromosome. *Xist* is regulated by two others lncRNAs, one acting negatively *Tsix*, and the other positively *Jpg*. *Tsix*, antisense RNA to *Xist*, is expressed from the X-active chromosome and inhibit *Xist* expression in *cis*. When expressed, *Tsix* recruits DNA methyltransferases (Dnmt3a) to repress the expression of *Xist*, and blocks the interaction between *Xist* and the PRC2 complex.

2. Epigenetic regulation

The lncRNA *Air* is submitting to the genomic imprinting. It consists of a 108 kb-long transcript. *Air* promoter is localized in the imprinting centre within the *IGF2r* gene and it is necessary for the paternal repression of the gene of the locus [4, 5]. However, the molecular mechanism remains unclear and authors propose hypothesis of methylation propagation from the *IGF2R* gene or of repressive ARN/protein complexes formation.

The third lncRNA well studied is located in the cluster *Kcnq1/Kcnq1ot1* on the chromosome 11 in position 15.5. *Kcnq1ot1* RNA is a 91kb transcript which is expressed in antisense orientation from a highly conserved and differentially methylated region *Kcnq1* ICR or ICR2 present in intron 10 of *Kcnq1* gene. Expression of this transcript is exclusively paternal. Indeed, the *Kcnq1ot1* promoter shows a maternal specific methylation. This differential epigenetic mark is lost in patients affected by Beckwith-Wiedemann syndrome with RNA biallelic expression [6, 7, 8]. More recently, Pandey and colleagues (2004) have documented that the *Kcnq1ot* transcript has a key role in silencing of genes contained in the *Kcnq1* gene imprinted region and that it participates directly or indirectly to the methylation but without RNA interference mechanisms [9]. Furthermore, interruption of *Kcnq1ot1* RNA production by the insertion of a polyadenylation sequence downstream of the promoter also caused a loss of both silencing

activity and methylation spreading. Thus, the antisense RNA plays a key role in the silencing function of the ICR [10].

3. Regulation of alternative splicing

The lncRNA *Malat1* (Metastasis-associated lung adenocarcinoma transcript 1), also known as *NEAT2* (Nuclear-enriched abundant transcript 2), is polyadenylated and overexpressed in various cancers. It is a conserved transcript among mammals of 6-7 kb, localized in nuclear. RNA-fish studies have shown that *Malat1* is localized in sub-compartment of nuclear: nuclear speckles [11, 12]. Contrary to *NEAT1* (a lncRNA essential for nuclear paraspeckle formation), *Malat1* is not essential to nuclear speckle integrity. This compartment is composed in majority of factor involved in pre-mRNA splicing, like SR family protein and protein implicated in RNA transport for example. Bernard et al, have shown that *Malat1* controls of SR family protein (SF2/ASF) of splicing factor to transcription site. Tripathi and co-workers have established that *Malat1* regulates expression levels, localization and activity of SR protein. Targets genes of *Malat1* are tissue-dependant. In neuronal cells, *Malat1* regulates preferentially splicing of genes involved in synaptogenesis like Neuroigin gene (*Nlgn1*) and synaptic cell adhesion molecule 1 (*SynCAM1*).

4. Stem cells maintenance

Recently, it was observed that 133 lincRNAs were overexpressed and 104 down-regulated in ESC (Embryonic Stem Cell) or iPSC (induced Pluripotent Stem Cells) compared with fibroblast [13]. They have shown that twenty-eight lincRNAs upregulated in iPSC, notably lincRNA-RoR, could be regulated by pluripotency transcription factors OCT4, SOX2 or NANOG. Depletion of lincRNA-RoR inhibits iPSC colony formation. They have proved that lincRNA-RoR promotes survival of iPSC and ESC by preventing the activation of stress pathways like p53 response. This RNA is important to reprogramming stem cells whence its name 'Regulator of Reprogramming'. So, they identified the first functional lincRNA in establishing iPSC.

Furthermore, a study has identified several ncRNAs implicated in stem cells differentiation [14]. They have demonstrated that lncRNAs are associated with trimethylated H3K4 histones and histone methyltransferase MLL1. These suggest that lncRNAs have a role in epigenetic regulation during ES cell differentiation.

5. LncRNA in cancer

In human, HOX transcription factor are encoded by four HOX cluster on four different chromosomes: HOXA to D. From HOX cluster only 39 transcription factors are expressed but 231 ncRNAs are transcript [15, 16]. The well HOX ncRNA studied is *HOTAIR* (Hox antisense transcript RNA). It is antisense RNA of 2,2 kb, transcript from the HOXC cluster. Studies have

shown that *HOTAIR* regulates expression of genes on HOXD cluster, so acts in *trans*. Indeed, when expressed, due to its 5' domain, *HOTAIR* interacts with PRC2 complex, notably Suz12 and EZH2 protein [15]. PRC2 complex induces trimethylation of Histone H3 lysine 27 on HOXD cluster (an inactive methylation). *HOTAIR* can also interact with the LSD1/CoREST/REST complex: a complex involved in trimethylation of Histone H3 lysine 4 (active chromatin). *HOTAIR* regulates chromatin conformation from active chromatin to inactive. So, it is scaffold RNA [17]. In several cancers, notably breast cancer, *HOTAIR* expression is associated to metastasis [18]. Authors have shown that *HOTAIR* overexpression increases cells invasion and metastasis in mice. They established that *HOTAIR* invasion is PRC2 complex dependent. So, *HOTAIR* expression is associated to poor prognosis.

LncRNAs are implicated in several cellular processes (epigenetic regulation, dosage compensation, stem cells self-renewal and differentiation). In some cases, lncRNA expression allows maintains of stem cells pool for example, whereas, sometimes, lncRNA expression is responsible of cancers. So, expression of lncRNA must be well regulated. They can act *per se* but can also acts as precursor of small ncRNA such as microRNA. The lncRNA *H19*, the first imprinting ncRNA discover is the precursor of the microRNA: miR-675 [19]. Recently, several groups have identified targets of this miRNA in several cell lines [20, 21]. To illustrate our point, we summarize our current understanding about the *H19/IGF2* cluster, its transcription complexity and its function in cells.

II. The *H19/IGF2* cluster

1. Genomic imprinting at the *H19/IGF2* cluster

The *H19/IGF2* cluster is submitted to genomic imprinting. Genomic imprinting is a form of epigenetic gene regulation that results in expression of a single allele in a parent-of-origin-dependent manner. This form of monoallelic expression is essential for normal development. Despite extensive studies, the molecular mechanisms of genomic imprinting remain unclear. However, some hallmarks of this phenomenon have been identified and we can note that:

- Gene expression is allele-specific and tissue or stage-specific.
- Many of imprinted genes are found in clusters throughout the genome. The clusters contain two or more imprinted genes over a region that can span 1 Mb or more.
- Within each cluster, a common regulating region which are called “imprinting control region” (ICR, also called IC for imprinting Centre or ICE for imprinting control element) controls the imprinting of all genes in the cluster and can act over hundreds of kilobases. ICRs are designed as differentially methylated regions with parental-specific modifications that determine their activity. Deletions of this region lead to the loss of imprinting of multiple genes of the cluster [22, 23].

- More recently, it has been reported that non coding RNA were associated with imprinted clusters and have an essential role in regulating gene expression.

The *H19/IGF2* cluster is located on the human chromosome 11 in position p15.5. This 1Mbp domain contains 9 imprinted genes and 2 independent imprinting center. The first imprinting center (ICR) regulates the cluster *H19/IGF2* and the second (ICR2), the cluster *Kcnq1/Kcnq1ot1*.

The *H19* gene is one of the first genes proven to be imprinted. This gene is co-regulated negatively with the *IGF2* gene located 200kb upstream of the transcription site of the *H19* gene. Indeed, *H19* is expressed only from the maternal allele whereas *IGF2* is expressed from the paternal allele [24]. The paternal allele exhibits several characteristics that explain the silencing of the *H19* gene: it is hypermethylated in the promoter region and the promoter shows a compact chromatin structure [25, 26]. Moreover, the histone acetylation rate is lower than the one of the maternal allele [27].

Surprisingly, the *IGF2* promoter region is not methylated and its chromatin structure is favourable to a biallelic transcription [28]. However, two other differentially methylated regions (DMR) on the expressed paternal allele have been identified within the gene: the DMR1 located 3 kbp upstream the P1 promoter acts as a silencer on the maternal allele when it is unmethylated, and the DMR2, located within exons 5 and 6 is an activator on the paternal allele when it is methylated [29, 30, 31].

However, DNA methylation is not sufficient to explain the mono-allelic expression. Indeed, the ICR is the key of the genomic imprinting: it controls the chromatin structure and regulates the effect of enhancers located downstream of the *H19* gene [32, 33]. This region is located 2 to 4 kbp upstream of the transcription site of the *H19* gene. In human, it contains seven binding site of zinc-finger protein named CTCF (CCTC-binding factor) but only the sixth is differentially methylated [34]. On the maternal allele, the CTCF protein interacts with non-methylated ICR due to four consensus site (figure 1) [35]. On the ICR, this protein has a chromatin insulator function as it prevents the action of enhancers on the promoter of *IGF2*. On the paternal allele, methylation of the ICR represses the *H19* expression and prevents the attachment of the CTCF protein [36]. So enhancers can activate the *IGF2* expression from this allele. Thus, *H19* is expressed from the maternal allele and *IGF2* from the paternal allele (figure 1).

Chromosome conformation capture (3C) analysis shows interaction between different chromosomal regions and suggests that the CTCF protein has a critical role in the epigenetic regulation of the cluster *H19/IGF2*. Kurukuti and al. 2006 demonstrated that on the maternal allele, ICR interact with the DMR1 and the Matrix Attachment Region (MAR3) at the *IGF2* locus to generate a tight loop around the *IGF2* gene [37, 38]. This interaction creates an inactive

domain where *IGF2* is far away from the enhancers. Therefore, this gene is in inactive domain so it cannot be expressed from this allele (figure 2).

On the paternal allele, the methylated ICR interacts with methylated *IGF2* DMR2 moving *IGF2* into the active chromatin domain [39].

So genomic imprinting of the *H19/IGF2* cluster is allowed by DNA methylation, chromatin composition, organization and conformation.

2. News transcripts at the *H19/IGF2* locus

In 1991, an antisense transcript of the *IGF2* gene in chicken was identified [40]. Others studies have identified antisense *IGF2* transcripts of 3-4 kb in mouse and human (figure 3) [41, 42]. This transcript is expressed only from paternal allele and no open reading frame (ORF) was identified. Its function remains unclear, but it is a good marker for Wilm's tumor where it is overexpressed [42]. Recently, it was shown that *IGF2as* is exported in the cytoplasm and associated with polysomes [43]. So, it is not impossible that *IGF2as* is a protein coding transcript.

We have identified a non-coding transcript, antisense to *H19*, that we named *91H* (figure 3) [44]. This transcript is a lncRNA of 120kb expressed only in human from the maternal allele. It is known that lncRNAs can mediate epigenetic regulation transcription. Indeed, the lncRNA *Xist* induce X-chromosome inactivation in dosage compensation. So, we have studied effects of *91H* expression at the *H19/IGF2* locus. By invalidation of *91H* with si-RNA, a reduction of *IGF2* expression was observed. However, today, the molecular mechanism remains unclear. Recently, it was shown that *91HRNA* and its function are conserved among mammals, notably in mice [45]. By *91H* overexpression, they have shown that *91H* regulates positively *IGF2* translation from a novel promoter. More recently, a group identified antisense transcript of *H19*, expressed from the maternal allele, encoding for a protein named HOTS (*H19* opposite tumour suppressor) [46]. But it is not excluded that this protein is encoded by the *91H* transcript. Thus, today the function of the *91H* transcript remains unclear.

A new paternal transcript was identified in mice. This transcript, *PIHit* (Paternally-expressed *IGF2/H19* intergenic transcript), is coding by intergenic sequence, between *IGF2* and *H19*, and expressed, in mice, principally 8 days after birth (figure 3) [47]. Then, its expression decreases rapidly during the third post-natal week. It is expressed at similar level to mRNA (*IGF2*), capped but no polyadenylated. Neither ORF was identified, so it is supposed that it is a lncRNA. Authors have identified transcription start site but not the 3'end, which is why it is a transcript of 5 to 6 kb. Neither function has been associated to *PIHit* RNA. By 3C, they observed two chromatin conformation of paternal allele. They supposed that there is a dynamic system

permitting *IGF2* or *PIHit* expression. However, it cannot exclude that there is chromatin conformation cell lines specific.

III. The *H19* RNA and its function in cells

1. The *H19* RNA

The *H19* gene was discovered in the mouse in 1984 and in the human in 1992 [48, 49, 50]. This gene is composed of five exons and encoded an mRNA of 2,3kb. This RNA is transcribed by the RNA polymerase II, polyadenylated, capped and spliced with conserved secondary RNA structure. But, no conserved open reading frame was identified. Even if deletion and/or mutation produce a 26kDa protein, no endogenous translation has so far been identified [51]. So, in 1990, Brannan et al. have proposed that *H19* RNA functions as a *riboregulateur* of which expression is developmentally regulated [52].

It is well established that a ncRNA can be precursor of microRNA. There are different biogenesis pathways of microRNA, but generally stem-loop structure RNA are recognized by protein like DGCR8, cleaved by Drosha and Dicer to generate the duplex miR-5p/miR-3p. Then, the duplex interacts with Argonaut protein family and is incorporated in the RISC complex. MicroRNA can also be generated in Drosha or Dicer-independent pathways [53]. Introns from the splicing or tRNA (tRNA-Ile for example) can be directly recognized by Dicer, cleaved by this enzyme and incorporated in the RISC. There is a microRNA (miR-451) cleaved by Drosha which is directly recognized by Ago and incorporated in the RISC.

In 2007, Cai and Cullen have demonstrated that *H19* is precursor of microRNAs: miR-675-5p and miR-675-3p [19]. They are generated by the exon1 of the gene. Today, few targets of the miR-675 have identified. Due to its microRNA, it was shown that *H19* can regulate placental growth and cell cycle (see below).

2. *H19/91H* regulate *IGF* associated phenotypes

We and others groups have established that *H19* regulates *IGF2* ligand and receptor expression. Expression of *H19* and *IGF2* are regulated by enhancers located downstream of *H19*. Actions of enhancers are regulated by imprinting control region (ICR), located between *H19* and *IGF2*. Expression of *H19* and *IGF2* are allele-dependent. On the maternal allele, the CTCF protein interacts with ICR non-methylated, DMR1 of *IGF2* and the MAR3 domain [35, 37, 38]. This interaction creates a loop containing *IGF2* gene. Enhancers cannot active transcription of *IGF2* gene when chromatin is in this conformation. So, only *H19* is expressed from the maternal allele. On the paternal allele, ICR is methylated, so the protein CTCF is absent on this allele. ICR methylated interacts with *IGF2* DMR2 methylated too. This interaction allows action of enhancers on *IGF2* promoter and then *IGF2* expression [30]. *H19*

and *IGF2* are in competition for enhancers. Furthermore, it was shown that deletion of *H19* and its flanking region affect expression of *IGF2*. So, *H19* and region flanking regulate *IGF2* expression in *cis*.

Moreover, *H19* is a RNA polyadenylated, spliced and exported in the cytosol. In cytosol, a group have shown that *H19* is associated to polysomes [54]. These polysomes have similar size to those associated to *IGF2* mRNA. Then, they have found an inverse co-regulation between *H19* expression and *IGF2* translation in cytosol. In Wilm's tumor, *H19*-negative cells show overexpression of *IGF2* 3 fold higher than control. Inversely, in *H19*-positive cells, *IGF2* expression protein was reduced. So, they hypothesized, that *H19* regulates translation of *IGF2* mRNA in *trans*. Moreover, it was shown a co-regulation between *H19* and *IGF2* transcription. In breast cancer cells, when *H19* is overexpressed, *IGF2* expression decreases severely [55, 56]. *H19* regulates negatively transcription of *IGF2* in *trans*. So, it was supposed that *H19* acts as a *trans-riboregulateur*.

In mouse placental cells, the expression of miR-675, from *H19* gene, is regulated negatively by HuR protein [21]. They observed a relation between miR-675 expression and size of placenta. Indeed, when miR-675-3p is expressed (from E11.5 until term), a reduction size of placenta is observed. This reduction is due to a decrease of cells proliferation but not an increase of apoptosis. They established that miR-675-3p interacts with two seed on 3'UTR *Igf1r* and inhibits its translation. So, *H19* is a key regulator of IGF ligand and receptor expression (figure 4).

Intriguingly, we have shown that the lncRNA *91H*, transcript antisense to *H19*, affects a little *H19* expression but regulates positively *IGF2* [44]. However, the mechanism remains unclear. We supposed that *91H* interacts with proteins that modulate expression of genes. More recently, a group have shown that *91H* overexpression, in mouse, upregulates *IGF2* expression [45]. They supposed that *91H* activate a novel promoter of *IGF2*.

3. Cell cycle regulation by *H19* promotes cancer

H19 is implicated in embryonic development. It is expressed in blastocyst stage of development and accumulated at high level in tissues of endodermal and mesodermal origins as well as ectodermal origin [57, 58, 59, 60]. After birth, the gene is repressed in all tissues except skeletal muscle [61]. Misregulations of *H19* expression during development induce developmental syndrome like Silver-Russel syndrome or Beckwith-Wiedemann syndrome [62, 63]. In adulthood, function of *H19* is controversial: it was supposed that *H19* act as tumor suppressor or oncogene. Nevertheless, several data show that *H19* act as oncogene in various cancer tissues: breast [61, 64, 65], uterus [66], bladder [67, 68] and gastric [69]. Indeed, today,

it was clearly established that *H19*, *per se* or through its microRNA, regulates different check-point of the cell cycle.

H19 mRNA generates two microRNAs: miR-675-5p and miR-675-3p [19]. The miR-675-5p is most studied but few targets have been identified. The first target identified is 3'UTR of Retinoblastoma (*RB*) mRNA in colon cancer cells [20]. Authors have shown a negative co-regulation between *H19* and *RB* in human colorectal tumour. They demonstrated a reduced expression of *RB* in tumour whereas *H19* and miR-675-5p are overexpressed (figure 5). Owing to a reporter luciferase vector, they established the interaction and the negative effect of miR-675-5p on 3'UTR of *RB*. They have shown that miR-675-5p increases clonogenicity in soft agar of human colon cancer cells. So, in colon cancer cells lines, *H19* and miR-675-5p increase proliferation of cells.

In 1998, it was shown that the *H19* expression is regulated by p53 protein [70]. Indeed, the *H19* promoter contains consensus site to interaction with p53. A negative regulation of p53 on this promoter was observed. In parallel, in human breast tumour, *H19* is overexpressed in 70% of tumour independent of p53 expression [71]. However, it was shown that *H19* is located in stromal cells whereas p53 is located in epithelial cells.

Recently, a group studying gastric cancer shows that *H19* expression is increased in this disease [69]. They observed an increase cell proliferation and a reduction of apoptosis when *H19* is overexpressed. So, they studied the effect of *H19* overexpression on a protein inhibiting cell cycle proliferation and inducing apoptosis: the p53 protein. Thanks to a RNA-immunoprecipitation (RIP), they have shown that *H19* RNA can interact physically with the p53 protein. By a luciferase reporter system, they demonstrated that *H19* RNA regulates negatively the p53 protein may be by blocking this phosphorylation. So, the *H19* overexpression in gastric cancer cells contributes to tumorigenesis by regulating p53 activation.

In tumour, some cells are in hypoxic condition. So, a team has studied effect of hypoxia on *H19* expression. Upon hypoxia, they observed an increase rate of *H19* RNA [72, 73]. So, they have verified that the activation is due to the activation of the HIF1- α pathway (pathway activated during hypoxia). Invalidation of *HIF1- α* by RNA interference induces a diminution of *H19* overexpression upon hypoxia. Furthermore, they have observed that *H19* is overexpressed only when p53 is mutated or absent. If, p53 is not mutated, they have observed a decrease *H19* expression. It has previously been reported that p53 inhibits action of HIF1- α by increasing its ubiquitination and degradation [74]. So, upon hypoxia, *H19* is overexpressed by activation of HIF1- α pathway and p53-dependent manner.

Surprisingly, in breast cancer cells (MCF-7), they have observed an overexpression of *H19* upon hypoxia although p53 is present. In this cell, the p53 protein is principally in the

cytoplasm. So, to repress activation of transcription by HIF1- α , p53 must be in the nucleus. Taken these results together, we can hypothesize that, in MCF-7 cells, *H19* interacts with the p53 protein and inhibits its activation by sequestering p53 protein in the cytoplasm.

Furthermore, it was shown that *H19* facilitates cell cycle transition G1/S [75]. This check-point is regulated particularly by the E2F1 protein. The *H19* promoter contains two consensus sites for this protein. It was studied the potential role of E2F1 on the *H19* promoter. Using luciferase system, it was reported that E2F1 induced *H19* expression through these two sites. Moreover, the RB protein and E2F6 factor inhibit the activation of E2F1. So, E2F1 is negatively regulated by RB and E2F6. Recently, it was shown that *H19*, thanks to its microRNA, regulates negatively RB expression [20]. So, these studies have demonstrated a positive feedback loop between *H19* and E2F1 (figure 5). Then, in breast cancer cells lines as BT20, T47D and MCF-7, E2F1 and *H19* are overexpressed. It was shown that *H19* overexpression conferred a growth advantage on cells. Indeed, an increase S-phase entry was observed when cells overexpress this gene [75]. So, *H19*, through a positive regulation by E2F1, active cell cycle progression and promotes growth of breast cancer cells.

To resume, it was shown that, in cancer cell lines, *H19* control p53 activity, reduced translation of the *RB* mRNA and promotes the G1/S cell cycle transition (figure 5). So, *H19* have a key role in the regulation of cell cycle and could be implicated in cancer progression.

Conclusion

Recently, the ENCODE project have proved that 90% of the genome is transcribed, but protein-coding genes represent only 2% of transcripts. So, the genome encode for ncRNAs which have an important place in function of cells. Today, small non coding RNAs are most studied than lncRNAs. However, lncRNAs regulate the expression of 70% of genes. So general, it is accepted that lncRNAs regulate gene expression by interacting with complex chromatin modifications, including PCR2. In this review, we show that lncRNA, notably *H19* RNA, may have several functions in cell without involving the complex. So, we show that lncRNAs can act *per se* and/or as precursor of microRNA. Furthermore, we show that lncRNAs can be involved in different cellular process. For example, the *H19* RNA and its microRNA regulate negatively expression of *IGF2* ligand and IGF1 receptor; and control positively cell cycle progression. So, today, it is clear that lncRNAs have function as important as protein within cellular processes.

References

- [1] Esteller, M. (2011) Non-coding RNAs in human disease. *Nature Reviews. Genetics*, 12, 861–874.
- [2] Knowling, S., and Morris, K.V. (2011) Non-coding RNA and antisense RNA. Nature's trash or treasure? *Biochimie*, 93, 1922–1927.
- [3] Lee, J.T. (2012) Epigenetic regulation by long noncoding RNAs. *Science*, 338, 1435–1439.
- [4] Rougeulle, C., and Heard, E. (2002) Antisense RNA in imprinting: spreading silence through Air. *Trends in Genetics*, 18, 434–437.
- [5] Sleutels, F., Zwart, R., and Barlow, D.P. (2002) The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature*, 415, 810–813.
- [6] Mitsuya, K., Meguro, M., Lee, M.P., Katoh, M., Schulz, T.C., Kugoh, H., Yoshida, M.A., Niikawa, N., Feinberg, A.P., and Oshimura, M. (1999) LIT1, an imprinted antisense RNA in the human KvLQT1 locus identified by screening for differentially expressed transcripts using monochromosomal hybrids. *Human Molecular Genetics*, 8, 1209–1217.
- [7] Lee, M.P., DeBaun, M.R., Mitsuya, K., Galonek, H.L., Brandenburg, S., Oshimura, M., and Feinberg, A.P. (1999) Loss of imprinting of a paternally expressed transcript, with antisense orientation to KVLQT1, occurs frequently in Beckwith-Wiedemann syndrome and is independent of insulin-like growth factor II imprinting. *Proceedings of the National Academy of Sciences USA*, 96, 5203–5208.
- [8] Du, M., Zhou, W., Beatty, L.G., Weksberg, R., and Sadowski, P.D. (2004) The KCNQ1OT1 promoter, a key regulator of genomic imprinting in human chromosome 11p15.5. *Genomics*, 84, 288–300.
- [9] Pandey, R.R., Ceribelli, M., Singh, P.B., Ericsson, J., Mantovani, R., and Kanduri, C. (2004) NF-Y regulates the antisense promoter, bidirectional silencing, and differential epigenetic marks of the Kcnq1 imprinting control region. *Journal of Biological Chemistry*, 279, 52685–52693.
- [10] Thakur, N., Tiwari, V.K., Thomassin, H., Pandey, R.R., Kanduri, M., Göndör, A., Grange, T., Ohlsson, R., and Kanduri, C. (2004) An antisense RNA regulates the bidirectional silencing property of the Kcnq1 imprinting control region. *Molecular and Cellular Biology*, 24, 7855–7862.
- [11] Bernard, D., Prasanth, K.V., Tripathi, V., Colasse, S., Nakamura, T., Xuan, Z., Zhang, M.Q., Sedel, F., Jourdain, L., Couplier, F., et al. (2010) A long nuclear-retained non-coding

RNA regulates synaptogenesis by modulating gene expression. *EMBO Journal*, 29, 3082–3093.

[12] Tripathi, V., Ellis, J.D., Shen, Z., Song, D.Y., Pan, Q., Watt, A.T., Freier, S.M., Bennett, C.F., Sharma, A., Bubulya, P.A., et al. (2010) The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Molecular Cell*, 39, 925–938.

[13] Loewer, S., Cabili, M.N., Guttman, M., Loh, Y.-H., Thomas, K., Park, I.H., Garber, M., Curran, M., Onder, T., Agarwal, S., et al. (2010) Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nature Genetics*, 42, 1113–1117.

[14] Dinger, M.E., Amaral, P.P., Mercer, T.R., Pang, K.C., Bruce, S.J., Gardiner, B.B., Askarian-Amiri, M.E., Ru, K., Soldà, G., Simons, C., et al. (2008) Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. *Genome Research*, 18, 1433–1445.

[15] Rinn, J.L., Kertesz, M., Wang, J.K., Squazzo, S.L., Xu, X., Bruggmann, S.A., Goodnough, L.H., Helms, J.A., Farnham, P.J., Segal, E., et al. (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*, 129, 1311–1323.

[16] Tano, K., and Akimitsu, N. (2012) Long non-coding RNAs in cancer progression. *Frontiers in Genetics*, 3, 219.

[17] Shore, A.N., Herschkowitz, J.I., and Rosen, J.M. (2012) Noncoding RNAs Involved in Mammary Gland Development and Tumorigenesis: There's a Long Way to Go. *Journal of Mammary Gland Biology and Neoplasia*, 17, 43–58.

[18] Gupta, R.A., Shah, N., Wang, K.C., Kim, J., Horlings, H.M., Wong, D.J., Tsai, M.-C., Hung, T., Argani, P., Rinn, J.L., et al. (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*, 464, 1071–1076.

[19] Cai, X., and Cullen, B.R. (2007) The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA*, 13, 313–316.

[20] Tsang, W.P., Ng, E.K.O., Ng, S.S.M., Jin, H., Yu, J., Sung, J.J.Y., and Kwok, T.T. (2010) Oncofetal H19-derived miR-675 regulates tumor suppressor RB in human colorectal cancer. *Carcinogenesis*, 31, 350–358.

[21] Keniry, A., Oxley, D., Monnier, P., Kyba, M., Dandolo, L., Smits, G., and Reik, W. (2012) The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. *Nature Cell Biology*.

- [22] 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, 34–39.
- [23] 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 and Development*, 11, 1596–1604.
- [24] DeChiara, T.M., Robertson, E.J., and Efstratiadis, A. (1991) Parental imprinting of the mouse insulin-like growth factor II gene. *Cell*, 64, 849–859.
- [25] 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 and Development*, 7, 1663–1673.
- [26] Ferguson-Smith, A.C., Sasaki, H., Cattanaach, B.M., and Surani, M.A. (1993) Parental-origin-specific epigenetic modification of the mouse H19 gene. *Nature*, 362, 751–755.
- [27] Grandjean, V., O'Neill, L., Sado, T., Turner, B., and Ferguson-Smith, A. (2001) Relationship between DNA methylation, histone H4 acetylation and gene expression in the mouse imprinted Igf2-H19 domain. *FEBS Letters*, 488, 165–169.
- [28] Sasaki, H., Jones, P.A., Chaillet, J.R., Ferguson-Smith, A.C., Barton, S.C., Reik, W., and Surani, M.A. (1992) Parental imprinting: potentially active chromatin of the repressed maternal allele of the mouse insulin-like growth factor II (Igf2) gene. *Genes and Development*, 6, 1843–1856.
- [29] Feil, R., Walter, J., Allen, N.D., and Reik, W. (1994) Developmental control of allelic methylation in the imprinted mouse Igf2 and H19 genes. *Development*, 120, 2933–2943.
- [30] Murrell, A., Heeson, S., Bowden, L., Constância, M., Dean, W., Kelsey, G., and Reik, W. (2001) An intragenic methylated region in the imprinted Igf2 gene augments transcription. *EMBO Reports*, 2, 1101–1106.
- [31] Constância, 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. *Nature Genetics*, 26, 203–206.
- [32] 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. *Molecular and Cellular Biology*, 17, 4322–4329.

- [33] 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, 1205–1213.
- [34] Takai, D., Gonzales, F.A., Tsai, Y.C., Thayer, M.J., and Jones, P.A. (2001) Large scale mapping of methylcytosines in CTCF-binding sites in the human H19 promoter and aberrant hypomethylation in human bladder cancer. *Human Molecular Genetics*, 10, 2619–2626.
- [35] Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature*, 405, 486–489.
- [36] 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 and Development*, 14, 1186–1195.
- [37] 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. *Nature Genetics*, 36, 889–893.
- [38] 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. *Proceedings of the National Academy of Sciences USA*, 103, 10684–10689.
- [39] Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002) Capturing chromosome conformation. *Science*, 295, 1306–1311.
- [40] Taylor, E.R., Seleiro, E.A., and Brickell, P.M. (1991) Identification of antisense transcripts of the chicken insulin-like growth factor-II gene. *Journal of Molecular Endocrinology*, 7, 145–154.
- [41] 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. *Proceedings of the National Academy of Sciences USA*, 94, 12509–12514.
- [42] Okutsu, T., Kuroiwa, Y., Kagitani, F., Kai, M., Aisaka, K., Tsutsumi, O., Kaneko, Y., Yokomori, K., Surani, M.A., Kohda, T., et al. (2000) Expression and imprinting status of human PEG8/IGF2AS, a paternally expressed antisense transcript from the IGF2 locus, in Wilms' tumors. *Journal of Biochemistry*, 127, 475–483.

- [43] Duart-Garcia, C., and Braunschweig, M.H. (2013) The Igf2as Transcript is Exported into Cytoplasm and Associated with Polysomes. *Biochemical Genetics*, 51, 119–130.
- [44] Berteaux, N., Aptel, N., Cathala, G., Genton, C., Coll, J., Daccache, A., Spruyt, N., Hondermarck, H., Dugimont, T., Cury, J.-J., et al. (2008) A novel H19 antisense RNA overexpressed in breast cancer contributes to paternal IGF2 expression. *Molecular and Cellular Biology*, 28, 6731–6745.
- [45] Tran, V.G., Court, F., Duputié, A., Antoine, E., Aptel, N., Milligan, L., Carbonell, F., Lelay-Taha, M.-N., Piette, J., Weber, M., et al. (2012) H19 antisense RNA can up-regulate Igf2 transcription by activation of a novel promoter in mouse myoblasts. *PLoS ONE*, 7, e37923.
- [46] Onyango, P., and Feinberg, A.P. (2011) A nucleolar protein, H19 opposite tumor suppressor (HOTS), is a tumor growth inhibitor encoded by a human imprinted H19 antisense transcript. *Proceedings of the National Academy of Sciences USA*, 108, 16759–16764.
- [47] Court, F., Baniol, M., Hagege, H., Petit, J.S., Lelay-Taha, M.-N., Carbonell, F., Weber, M., Cathala, G., and Forne, T. (2011) Long-range chromatin interactions at the mouse Igf2/H19 locus reveal a novel paternally expressed long non-coding RNA. *Nucleic Acids Research*, 39, 5893–5906.
- [48] Pachnis, V., Belayew, A., and Tilghman, S.M. (1984) Locus unlinked to alpha-fetoprotein under the control of the murine raf and Rif genes. *Proceedings of the National Academy of Sciences USA*, 81, 5523–5527.
- [49] Pachnis, V., Brannan, C.I., and Tilghman, S.M. (1988) The structure and expression of a novel gene activated in early mouse embryogenesis. *EMBO Journal*, 7, 673–681.
- [50] Zhang, Y., and Tycko, B. (1992) Monoallelic expression of the human H19 gene. *Nature Genetics*, 1, 40–44.
- [51] Joubel, A., Cury, J.J., Pelczar, H., Begue, A., Lagrou, C., Stehelin, D., and Coll, J. (1996) The 5' part of the human H19 RNA contains cis-acting elements hampering its translatability. *Cellular and Molecular Biology*, 42, 1159–1172.
- [52] 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. *Molecular and Cellular Biology*, 10, 28–36.
- [53] Yang, J.-S., and Lai, E.C. (2011) Alternative miRNA biogenesis pathways and the interpretation of core miRNA pathway mutants. *Molecular Cell*, 43, 892–903.

- [54] 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. *Journal of Biological Chemistry*, 273, 28247–28252.
- [55] Wilkin, F., Paquette, J., Ledru, E., Hamelin, C., Pollak, M., Deal, C.L., and Mamelin, C. (2000) H19 sense and antisense transgenes modify insulin-like growth factor-II mRNA levels. *European Journal of Biochemistry*, 267, 4020–4027.
- [56] Lottin, S., Adriaenssens, E., Dupressoir, T., Berteaux, N., Montpellier, C., Coll, J., Dugimont, T., and Cury, J.J. (2002) Overexpression of an ectopic H19 gene enhances the tumorigenic properties of breast cancer cells. *Carcinogenesis*, 23, 1885–1895.
- [57] 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, 1105–1114.
- [58] Lustig, O., Ariel, I., Ilan, J., Lev-Lehman, E., De-Groot, N., and Hochberg, A. (1994) Expression of the imprinted gene H19 in the human fetus. *Molecular Reproduction and Development*, 38, 239–246.
- [59] Ohlsson, R., Hedborg, F., Holmgren, L., Walsh, C., and Ekström, T.J. (1994) Overlapping patterns of IGF2 and H19 expression during human development: biallelic IGF2 expression correlates with a lack of H19 expression. *Development*, 120, 361–368.
- [60] Hemberger, M., Redies, C., Krause, R., Oswald, J., Walter, J., and Fundele, R.H. (1998) H19 and Igf2 are expressed and differentially imprinted in neuroectoderm-derived cells in the mouse brain. *Development Genes and Evolution*, 208, 393–402.
- [61] Douc-Rasy, S., Coll, J., Barrois, M., Joubel, A., Prost, S., Dozier, C., Stéhelin, D, Riou, G. (1993) Expression of the human fetal BAC/H19 gene in invasive cancer. *International Journal of Oncology*, 2, 753-758.
- [62] Okamoto, K., Morison, I.M., Taniguchi, T., and Reeve, A.E. (1997) Epigenetic changes at the insulin-like growth factor II/H19 locus in developing kidney is an early event in Wilms tumorigenesis. *Proceedings of the National Academy of Sciences USA*, 94, 5367–5371.
- [63] Steenman, M.J., Rainier, S., Dobry, C.J., Grundy, P., Horon, I.L., and Feinberg, A.P. (1994) Loss of imprinting of IGF2 is linked to reduced expression and abnormal methylation of H19 in Wilms' tumour. *Nature Genetics*, 7, 433–439.
- [64] Dugimont, T., Cury, J.J., Wernert, N., Delobelle, A., Raes, M.B., Joubel, A., Stehelin, D., and Coll, J. (1995) The H19 gene is expressed within both epithelial and stromal components of human invasive adenocarcinomas. *Biology of the Cell*, 85, 117–124.

- [65] Adriaenssens, E., Lottin, S., Dugimont, T., Fauquette, W., Coll, J., Dupouy, J.P., Boilly, B., and Cury, J.J. (1999) Steroid hormones modulate H19 gene expression in both mammary gland and uterus. *Oncogene*, 18, 4460–4473.
- [66] Ariel, I., Ayesh, S., Perlman, E.J., Pizov, G., Tanos, V., Schneider, T., Erdmann, V.A., Podeh, D., Komitowski, D., Quasem, A.S., et al. (1997) The product of the imprinted H19 gene is an oncofetal RNA. *MP, Molecular Pathology*, 50, 34–44.
- [67] Ariel, I., Lustig, O., Schneider, T., Pizov, G., Sappir, M., De-Groot, N., and Hochberg, A. (1995) The imprinted H19 gene as a tumor marker in bladder carcinoma. *Urology*, 45, 335–338.
- [68] Elkin, M., Shevelev, A., Schulze, E., Tykocinsky, M., Cooper, M., Ariel, I., Pode, D., Kopf, E., De Groot, N., and Hochberg, A. (1995). The expression of the imprinted H19 and IGF-2 genes in human bladder carcinoma. *FEBS Letters*, 374, 57–61.
- [69] Yang, F., Bi, J., Xue, X., Zheng, L., Zhi, K., Hua, J., and Fang, G. (2012) Up-regulated long non-coding RNA H19 contributes to proliferation of gastric cancer cells. *FEBS Journal*, 279, 3159–3165.
- [70] Dugimont, T., Montpellier, C., Adriaenssens, E., Lottin, S., Dumont, L., Iotsova, V., Lagrou, C., Stéhelin, D., Coll, J., and Cury, J.J. (1998) The H19 TATA-less promoter is efficiently repressed by wild-type tumor suppressor gene product p53. *Oncogene*, 16, 2395–2401.
- [71] Adriaenssens, E., Dumont, L., Lottin, S., Bolle, D., Leprêtre, A., Delobelle, A., Bouali, F., Dugimont, T., Coll, J., and Cury, J.J. (1998) H19 overexpression in breast adenocarcinoma stromal cells is associated with tumor values and steroid receptor status but independent of p53 and Ki-67 expression. *American Journal of Pathology*, 153, 1597–1607.
- [72] 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.
- [73] Matouk, I.J., Mezan, S., Mizrahi, A., Ohana, P., Abu-Lail, R., Fellig, Y., Degroot, N., Galun, E., and Hochberg, A. (2010) The oncofetal H19 RNA connection: hypoxia, p53 and cancer. *Biochimica et Biophysica ACTA/General Subjects*, 1803, 443–451.
- [74] Ravi, R., Mookerjee, B., Bhujwala, Z.M., Sutter, C.H., Artemov, D., Zeng, Q., Dillehay, L.E., Madan, A., Semenza, G.L., and Bedi, A. (2000) Regulation of tumor angiogenesis by p53-induced degradation of hypoxia-inducible factor 1alpha. *Genes and Development*, 14, 34–44.
- [75] Berteaux, N., Lottin, S., Monté, D., Pinte, S., Quatannens, B., Coll, J., Hondermarck, H., Cury, J.-J., Dugimont, T., and Adriaenssens, E. (2005) H19 mRNA-like noncoding RNA

promotes breast cancer cell proliferation through positive control by E2F1. *Journal of Biological Chemistry*, 280, 29625–29636.

Legends

Figure 1. Genomic imprinting at the *H19/IGF2* locus

Activation of gene transcription is represented by (+), repression by (-) and inhibition of enhancers function by ($^{-}$). Relative positions are expressed in kilobase pairs relatively to the *H19* transcription start site. Gene expression is regulated by three mechanisms: methylation, enhancers activity and insulator activity. Three DNA region are differentially methylated : DMR 1 and 2 of the *IGF2* gene (violet diamond) and ICR (blue oval). On the maternal allele, the CTCF protein interacts with the non-methylated ICR (blue oval). This interaction prohibits enhancers access to the *IGF2* gene. Furthermore, the DMR1 non-methylated of *IGF2* gene acts as a silencer. On the paternal allele, the methylated ICR repress *H19* and inhibits CTCF interaction. So, enhancers can activate *IGF2* transcription. Moreover, the methylated DMR2 of *IGF2* gene activate also *IGF2* transcription.

Figure 2. Chromatin loop structure at the *H19/IGF2* locus

Chromosome conformation capture revealed that the CTCF protein orchestrate chromatin structure. On the maternal allele, CTCF interact with ICR but also with DMR1 of *IGF2* and matrix attachment region (MAR3). These interactions create a loop around the *IGF2* gene. Then, enhancers can interact only with *H19* promoter and activate its expression. On the paternal allele, the methylated DMR2 of *IGF2* interact with the methylated ICR allowing thus moving of *IGF2* into the active chromatin domain.

Figure 3. News transcripts at the *H19/IGF2* locus

Gene expressed are in blue rectangle with arrow whereas inactive gene are in white rectangle. Protein coding gene are represented by (▣) and non protein coding gene by (▣). Direction of arrows indicates direction of gene transcription. Relative positions are expressed in kilobase pairs relatively to the *H19* transcription start site. ICR is represented by circle and DMR by diamond. Methylation of ICR and DMR is in black. Enhancers are represented by orange oval. The first antisense transcript described is *IGF2as* (Taylor, 1991). It is ncRNA overexpressed in Wilm's tumour. Its function remains unclear. Recently, news transcripts have been discovered at the *H19/IGF2* locus. On maternal allele, we have identified long antisense transcript: *91H* (Berteaux, 2008). This ncRNA and its function are conserved in mice (Tran, 2012). Neither protein was identified, so it is supposed that *91H* acts as lncRNA. Recently,

from the same allele, a protein coding gene imprinting has discovered: HOTS (Onyango, 2011). This transcript is antisense to *H19*. It is possible that HOTS protein is coding by *91H* RNA. The last transcript discovered from *H19/IGF2* locus is *PIHit* (Court, 2011). It is ncRNA expressed after birth from paternal allele.

Figure 4. Regulation of IGF by *H19* and its antisense transcript *91H*

Arrows represent positive regulation (→) and vertical bar negative regulation (⊣). *H19* overexpression regulates negatively *IGF2* transcription and translation. The lncRNA *91H* regulates positively transcription of *H19* and *IGF2* but the mechanism is unclear. *H19* is precursor of microRNA: miR-675. MiR-675 inhibits translation of *Igf1r* mRNA. So, *H19* is a key regulator of IGF ligand and receptor.

Figure 5. Regulation of cell cycle by *H19*

Proteins are represented by circle, RNA by diamond, cellular stress by rectangle and cell cycle by circle with arrow. Activation of gene expression is represented by arrow (→) and repression by vertical bar (⊣). MiR-675 is processed from *H19* transcript and represses translation of an inhibitor cell cycle progression mRNA: *Rb* (*Retinoblastoma*). So, E2F1 protein is activated and promotes *H19* overexpression and expression of genes involved in cell cycle progression. *H19* overexpression facilitates G1/S transition. Hypoxia, by activating HIF1- α , induces *H19* expression in p53-dependent manner. DNA damage activates the p53 protein. P53 repress *H19* expression and inhibits cell cycle progression by activating p21. Recently, it was shown that *H19* overexpression inhibits p53 activity. So, *H19* is a key regulator of cell cycle progression.

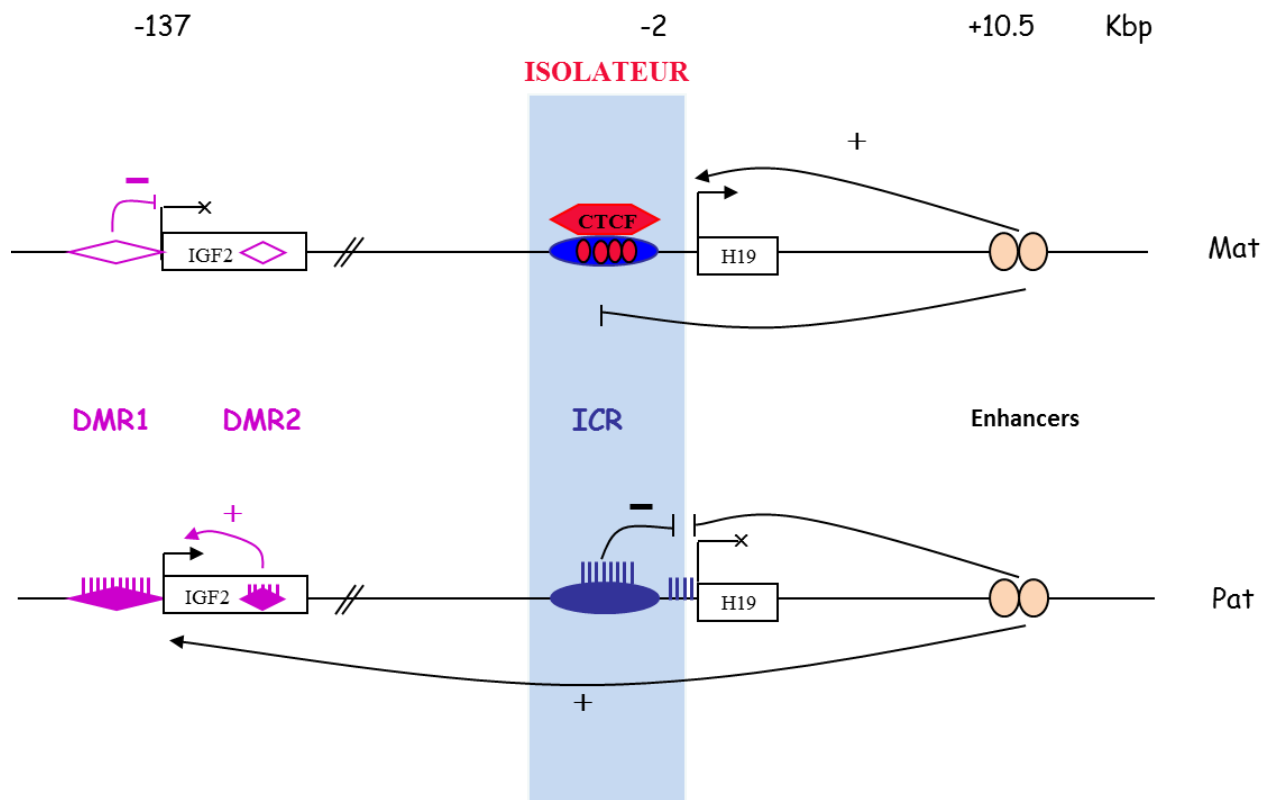


Figure 1

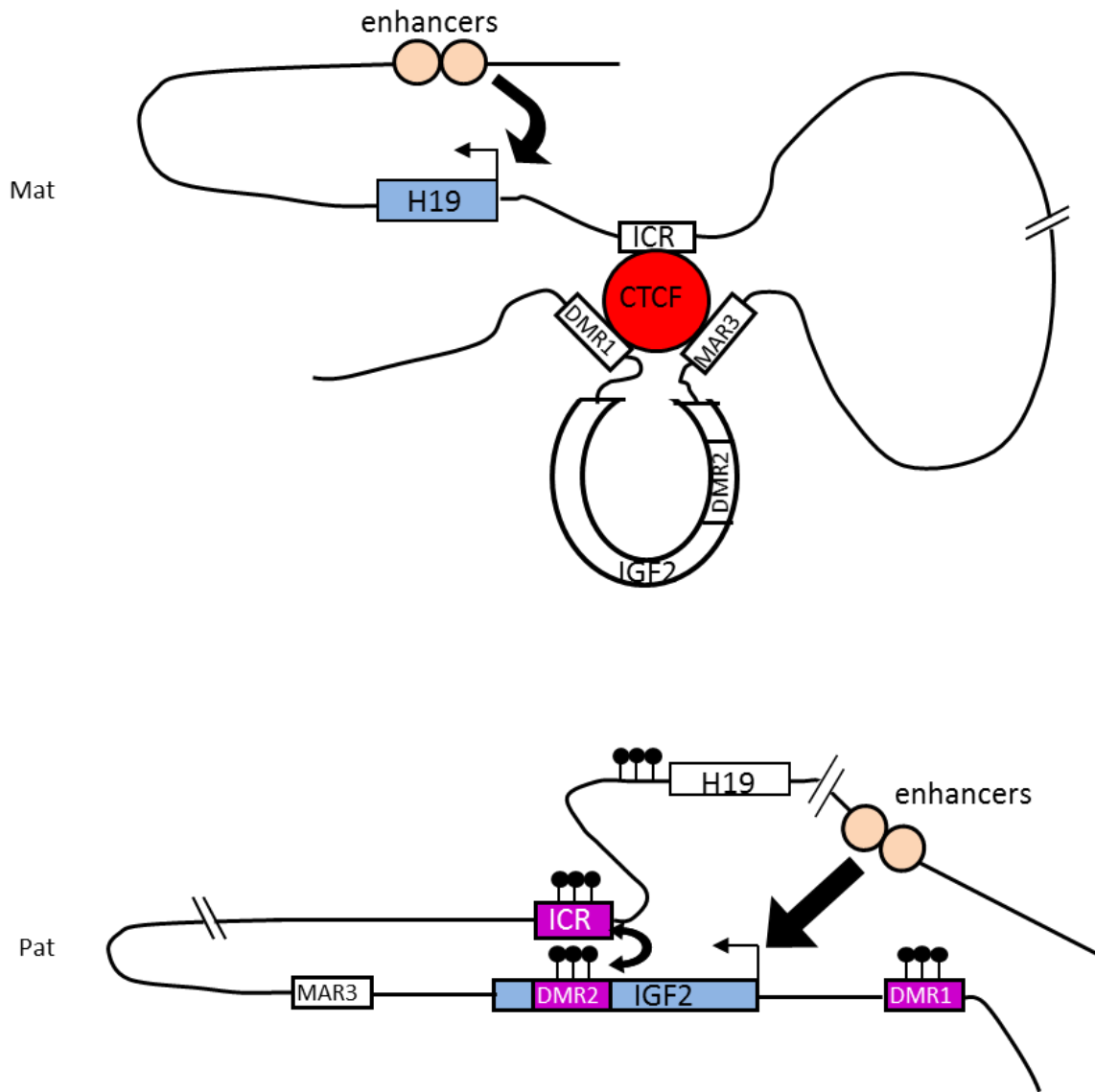


Figure 2

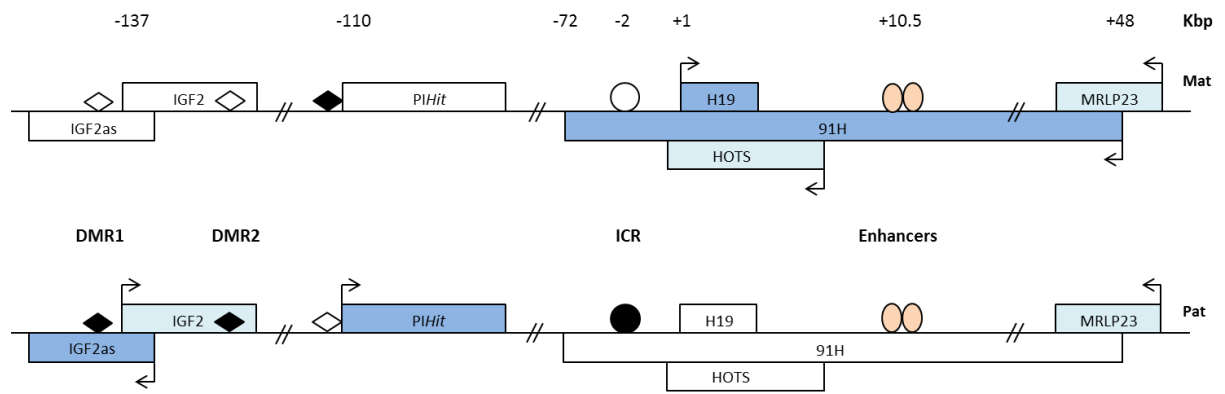


Figure 3

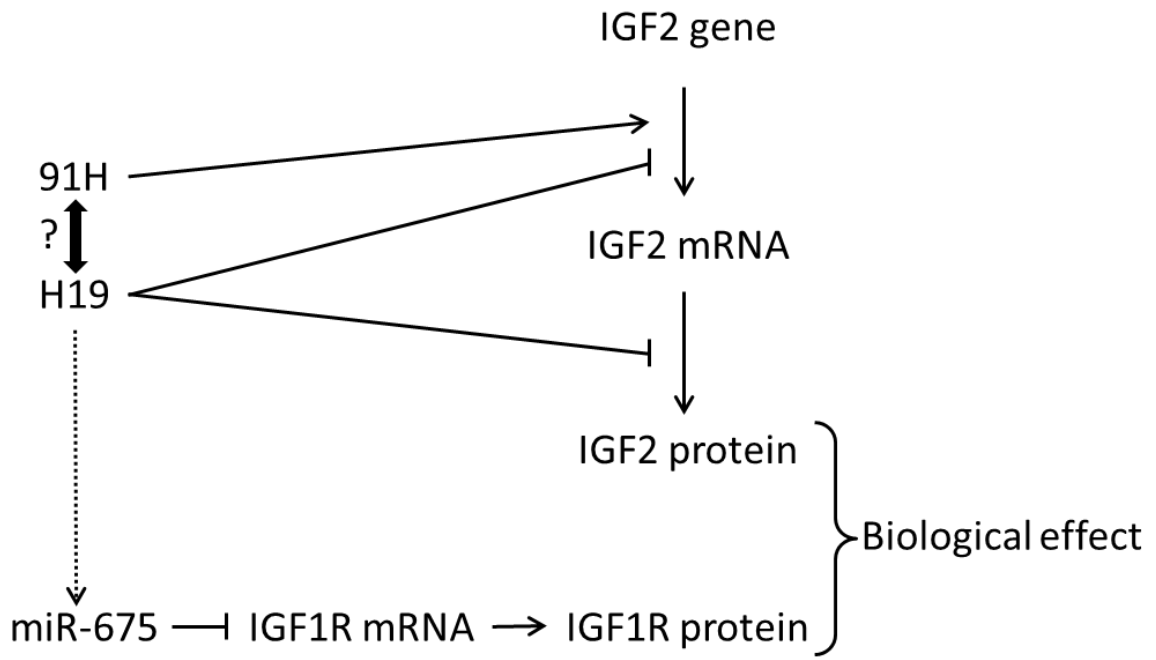


Figure 4

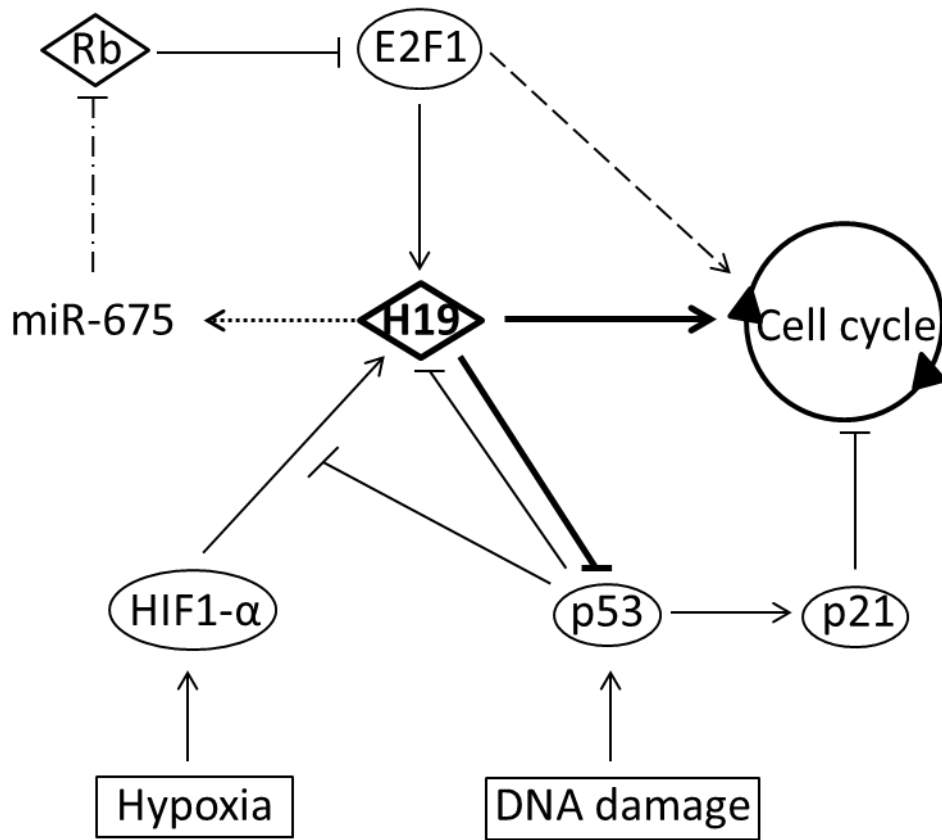


Figure 5

Annexe 2

Pierre-Olivier Angrand, **Constance Vennin**, Xuefen Le Bourhis and Eric Adriaenssens. (2015)
The role of long non-coding RNAs in genome formatting and transcription. *Frontier in genetics*,
2015.00165

The role of long non-coding RNAs in genome formatting and expression

Pierre-Olivier Angrand, Constance Vennin, Xuefen Le Bourhis and Eric Adriaenssens*

Cell Plasticity & Cancer - Inserm U908 / University of Lille, Bâtiment SN3, Cité Scientifique, F-59655 Villeneuve d'Ascq, France.

* Corresponding author: Eric ADRIAENSSENS
Tel.: +33 3 20 33 64 06
Email: eric.adriaenssens@univ-lille1.fr

ABSTRACT

Long non-coding RNAs (lncRNAs) are transcripts without protein-coding potential but having a pivotal role in numerous biological functions. Long non-coding RNAs act as regulators at different levels of gene expression including chromatin organization, transcriptional regulation and post-transcriptional control. Misregulation of lncRNAs expression has been found to be associated to cancer and other human disorders. Here, we review the different types of lncRNAs, their mechanisms of action on genome formatting and expression and emphasized on the multifaceted action of the H19 lncRNA.

KEYWORDS: lncRNAs, H19, Chromatin organization, Transcriptional regulation, Post-transcriptional control.

RUNNING TITLE: lncRNAs in Genome Formatting and Expression

The advent of DNA tiling arrays and deep sequencing technologies has revealed that a much larger part of the genome is transcribed into RNAs than previously assumed. It is estimated that up to 70% of the genome is transcribed but only 2% of the human genome codes for proteins (Bertone et al., 2004; ENCODE Project Consortium, Birney et al., 2007; Krapanov et al., 2007) and RNAs without coding potential are collectively referred as non-coding RNAs (ncRNAs).

Non-coding RNAs include the well-known ribosomal (r) RNAs, ribozymes, transfer (t) RNAs, small nuclear (sn) RNAs, telomere-associated RNAs (TERRA, TERC), as well as a plethora of far less characterized RNAs. Based on their size, these ncRNAs are subdivided into two groups: small ncRNAs (< 200 nt) and long ncRNAs [lncRNA (> 200 nt)]. Small ncRNAs, such as microRNAs (miRs), small interfering RNAs (siRNAs) or PIWI-interacting RNAs (piRNAs) received much attention and were shown to mainly act as negative regulators of gene expression. In contrast, lncRNAs represent a more functionally diverse class of transcripts. LncRNAs are found in a large diversity of animals species (Guttman et al., 2009; Jia et al., 2010; Pauli et al., 2012), but also in plants (Swiezewski et al., 2009), yeast (Houseley et al., 2008), and even in prokaryotes (Bernstein et al., 1993) and viruses (Reeves et al., 2007). LncRNAs remains poorly conserved among species (Pang et al., 2006; Derrien et al., 2012). However, accumulating evidences indicate that this RNA class plays an important role in a variety of biological processes and may be involved in cancer and other human diseases (Wapinski and Chang, 2011; Tano and Akimitsu, 2012).

Majority of lncRNAs are 5' capped, 3' polyadenylated, multi-exonic and are subjected to transcriptional regulation as coding mRNAs (Carninci et al., 2005; Guttman et al., 2010; Cabili et al., 2011; Derrien et al., 2012). Some of the lncRNAs such as XIST, MALAT1 or NEAT1 are almost exclusively localized in the nucleus (Brown et al., 1992; Hutchinson et al., 2007), whereas others are mostly found in the cytoplasm (Coccia et al., 1992; Yoon et al., 2012). In term of genomic organization, lncRNAs can be classified according to their proximity to protein coding genes into 5 categories: sense, when overlapping one or more exons of another transcript; antisense, when overlapping one or more exons of another transcript on the opposite strand; bidirectional, when its expression and the expression of the neighboring coding transcript on the opposite strand are initiated in close proximity; intronic, when arising from an intron of another transcript; or intergenic, when produced from an independent transcription unit in the interval between two protein coding genes. This crude classification illustrates that lncRNA expression may be controlled by different molecular mechanisms, but it does account neither for their modes of action nor for their cellular functions.

While only a limited number of lncRNAs has been studied, numerous evidences indicate that lncRNAs interact with a plethora of proteins. Furthermore, homologous Watson-Crick base pairing provides an efficient way by which lncRNAs may selectively interact with other nucleic

acid species. It is believed that lncRNAs are involved in a diversity of cellular functions through gene expression regulation at different levels including chromatin organization, transcriptional regulation and post-transcriptional mRNA processing (Mercer et al., 2009; Wilusz et al., 2009). To complicate matters further, Anderson et al. (2015) recently described that a conserved micropeptide is encoded by a skeletal muscle-specific RNA previously annotated as a putative long non-coding RNA. This finding leads to the proposal that several lncRNAs could also have a biological function through the production of micropeptides.

LncRNAs in the control of mRNA processing. The ability of lncRNAs to recognize complementary sequences allows the regulation of mRNA processing at various steps, including degradation, splicing, translation or transport (Figure 1).

Base pairing between defined regions of the human site APP-cleaving enzyme 1 (BACE1) transcript and its antisense lncRNA BACE1-AS induces the mRNA stabilization and consequently the increase in BACE1 protein abundance (Faghihi et al., 2008). Similarly, the lncRNA TINCR (terminal differentiation-induced ncRNA) interacts with a range of differentiation mRNAs including FLG, LOR, ALOXE3, ALOX12B, ABCA12, CASP14 or ELOVL3, to increase their stability (Kretz et al., 2013). In contrast, the recognition of mRNAs by other lncRNAs, such as half-STAU1-binding site RNAs (1/2sbsRNAs) decrease target mRNA stability by inducing STAU1 recruitment and the STAU1-mediated mRNA decay pathway (Gong and Maquat, 2011).

The translational process may also be modulated positively or negatively by lncRNA-mRNA pairing. For example, the antisense lncRNA ULCH-AS1 (ubiquitin carboxy-terminal hydrolase L1 antisense RNA 1) enhances ULCH mRNA translation (Carrieri et al., 2012), whereas lincRNA-p21 or pseudo-NOS suppress target mRNA translation (Yoon et al., 2012; Korneev et al., 1999).

The lncRNA MALAT1 (metastasis associated lung adenocarcinoma transcript 1) regulates pre-mRNA alternative splicing by modulating active serine/arginine splicing factors levels (Tripathi et al., 2010). In this case, the modulation of the mRNA processing is not achieved by a lncRNA-mRNA pairing mechanism but rather by the MALAT1-mediated modulation of the distribution of various splicing factors in nuclear speckle domains. However, antisense transcripts may also affect alternative splicing of their sense transcripts by virtue of masking splice sites by base complementarity (Krystal et al., 1990; Khochbin et al., 1992; Beltran et al., 2008). For example, a specific isoform of the lncRNA NPPA-AS is capable of down-regulating the intron-retained NPPA (atriuretic peptide precursor A) mRNA variant through RNA duplex formation between the sense and antisense transcripts (Annilo et al., 2009).

LncRNAs and the connection with the microRNA world. Some lncRNAs act on post-transcriptional regulation through the modulation of the microRNA (miR) pathways. MiRs, a large class of small ncRNA, function by annealing to complementary sites in the coding sequences or 3'-untranslated regions (UTRs) of target mRNAs where they favor the recruitment of protein factors that impair translation and/or promote transcript degradation leading to a decrease in protein abundance (Beak et al., 2008; Bartel, 2009). Specifically, one mechanism by which the BACE1-AS lncRNA enhances BACE1 sense mRNA stability could be by masking the binding site for miR-485-5p (Faghihi et al., 2010). Rather than competing for miR-binding sites, a number of lncRNAs contain miR-binding sites in their sequence and therefore act as “sponges” to sequester miRs away from their mRNA targets. The pseudogene PTENP1 previously considered as biologically inactive was found to sequester miRs, consequently affecting their action on target gene regulation (Poliseno et al., 2010). In particular, the 3'-UTR of the PTENP1 lncRNA binds the same set of miRs targeting the tumor suppressor gene PTEN, then reducing the downregulation of this transcript and thus enhancing PTEN protein abundance. A number of other lncRNAs, including KRASP1, linc-MD1, HULC or linc-ROR were shown to control mRNA activity through a miR sponge mechanism (Poliseno et al., 2010; Wang et al., 2010; Cesana et al., 2011; Wang et al., 2013). These examples illustrate that lncRNAs could counteract miR actions, but lncRNAs can themselves give rise to miRs and thus favor post-translational control by miR pathways as it is the case for the mouse Dlk1-Dio3 cluster or the BIC lncRNA (Eis et al., 2005; Hagan et al., 2009). Within the Dlk1-Dio3 cluster, *Meg3/Gtl2* contains in its last intron the evolutionarily conserved microRNA miR-770 whereas *Meg8* transcripts have the intron-encoded miR-341, miR-1188, and miR-370. Similarly, miR-155 is processed from sequences present in BIC lncRNA that accumulates in lymphoma cells.

LncRNAs in the transcriptional control. A number of evidences indicate that lncRNAs can act at the level of transcription either negatively or positively through a variety of molecular mechanisms (Figure 2). The dihydrofolate reductase (DHFR) gene contains a major and a minor promoter. The minor promoter gives rise to a lncRNA that forms a stable triplex lncRNA-DNA association at the major DHFR promoter and interacts with the general transcription factor II B (TFIIB) leading to the dissociation of the transcriptional preinitiation complex at this major promoter and then reducing DHFR expression (Martianov et al., 2007). Other lncRNAs act as decoys to negatively control transcription by titrating transcription factors away from their cognate promoters. The lncRNA PANDAR (promoter of CDKN1A antisense DNA damage activated RNA) is induced in a TP53-dependent manner and inhibits apoptotic gene expression to favor cell-cycle arrest through direct interaction with, and sequestration of NFYA, a transcription factor controlling the apoptotic program upon DNA damage (Hung et al.,

2011). Similarly, the lncRNA GAS5 (growth arrest-specific 5) contains an RNA motif derived from a stem-loop structure mimicking a DNA motif corresponding to the glucocorticoid response element. GAS5 binds to the DNA-binding domain of the glucocorticoid receptor, acts as a decoy glucocorticoid response element and is thus competing with DNA sites for binding to the glucocorticoid receptor (Kino et al., 2010).

Rather than acting as molecular decoys, lncRNA could modulate transcription by recruiting factors at target gene promoters or acting as transcription factor co-activators. For example, a lncRNA produced at the 5' regulatory region of the cyclin D1 (CCND1) gene in response to genotoxic stress tethers and modulates the activity of the RNA-binding protein TLS (translocated in liposarcoma) which in turn inhibits the activity of the histone acetyltransferases CBP (CREB binding protein) and EP300, leading to CCND1 transcriptional repression (Wang et al., 2008). The lncRNA Evf-2 (DLX6-AS1) forms a stable complex with the homeodomain-containing protein DLX2 to induce expression of the adjacent genes at the DLX5/6 locus (Feng et al., 2006). In this later case, the Evf-2 lncRNA functions as a co-factor regulating transcription factor activity.

Other lncRNAs regulate transcription by controlling transcription factor trafficking. As such, the lncRNA NRON (non-protein coding RNA, repressor of NFAT) interacts with importin-beta family members to inhibit nuclear translocation of the inactive dephosphorylated nuclear factor of activated T cells (NFAT) trans-activator (Willingham et al., 2005).

LncRNAs and epigenetics. LncRNAs have been implicated in the control of gene expression through the recruitment of epigenetic modifiers at specific genomic loci. In eukaryotic chromatin, epigenetic regulation is conveyed by covalent modifications of DNA (methylation, hydroxymethylation), modifications of histone tails (acetylation, methylation, phosphorylation, ubiquitylation) and the incorporation of various histone variants. These modifications locally change chromatin organization and regulate gene expression without changes in the DNA sequence. A number of evidences indicate that lncRNAs, acting as guides targeting enzymes involved in chromatin modifications, are part of this picture (Figure 3).

The lncRNA HOTAIR (HOX transcript antisense RNA) is transcribed from the HOXC locus and targets Polycomb Repressive Complex 2 (PRC2) to silence distantly located genes, including genes at the HOXD locus and hundreds of other genes on various chromosomes (Rinn et al., 2007; Zhang et al., 2015). Components of PRC2 trimethylate lysine 27 of histone H3 (H3K27me3) establishing the silent chromatin state (Völkel and Angrand, 2007; Völkel et al., 2015). Interestingly, HOTAIR also binds the LSD1-CoREST complex which possesses a lysine 4 of histone H3 demethylase activity, thus removing an active H3K4me2 chromatin mark (Tsai et al., 2010). Furthermore, deletion analysis of HOTAIR revealed that distinct parts of the lncRNA interact with PRC2 and LSD1 indicating that HOTAIR is able to bridge two independent

chromatin modifying activities at a target locus. Indeed, the knockdown of HOTAIR is responsible for the concomitant loss of occupancy of PRC2 and LSD1, and concurrent loss of H3K27me3 and gain of H3K4me2 at target loci. Then, HOTAIR acts as an RNA scaffold targeting two different histone modification activities involved in heterochromatin formation.

The interplay between one lncRNA and different chromatin modifying complexes is also found at the INK4A tumor-suppressor locus. The antisense lncRNA ANRIL (antisense noncoding RNA in the INK4 locus, CDKN2B-AS) which is produced by the INK4B/ARF/INK4A locus binds specifically two Polycomb proteins, CBX7 (PRC1) and SUZ12 (PRC2). Disruption of interaction with both PRC1 and PRC2 proteins impacts the transcriptional repression at the INK4B locus in cis (Yap et al., 2010; Kotake et al., 2011). As another example, the lncRNA KCNQ1OT1 (KCNQ1 opposite strand/antisense transcript 1) mediates bidirectional silencing by interacting with chromatin and recruiting the PRC2 complex, as well as the histone methyltransferase G9a (EHMT2), resulting in an increase in the repressive histone modifications H3K27me3 and H3K9me3 at the KCNQ1 domain (Panday et al., 2008). Thus, similar to HOTAIR and ANRIL, KCNQ1OT1 represents a prototype of a scaffold RNA recruiting multiple sets of chromatin modifying activities involved in target gene silencing. Approximately 20% of lncRNAs, including HOTAIR, ANRIL, KCNQ1OT1, but also XIST, RepA, HEIH, PCAT-1, H19 or linc-UBC1 (Zhao et al., 2008; Maenner et al., 2010; Yang et al., 2011; Prensner et al., 2011; Luo et al., 2013; He et al., 2013), are believed to guide PRC2 activity to target genes, indicating that lncRNA-mediated targeting of PRC2 at chromatin is a widely used strategy to repress gene expression through a chromatin reorganization mechanism (Khalil et al., 2009).

In contrast, the lncRNA HOTTIP (HOXA transcript at the distal tip) mediates transcriptional activation by controlling chromatin modification and organization (Wang et al., 2011). HOTTIP is produced from the 5'-end of the HOXA locus, downstream of HOXA13. The knockdown of HOTTIP decreases expression of HOXA genes in cis, with an efficacy that correlates with the proximity of the HOXA genes relative to the HOTTIP transcriptional unit. At the target genes, knockdown of HOTTIP results in the loss of activating H3K4me3 and H3K4me2 epigenetic marks, together with the decreases in occupancy of the MLL1 protein complex responsible for the establishment of these histone modifications. Furthermore, chromosome conformation capture carbon copy (C5) assays revealed abundant long-range looping interactions, bridging the transcribed target HOXA genes into proximity of the HOTTIP transcriptional unit. Thus, the mechanism by which the lncRNA HOTTIP controls HOXA expression relies on its potential to guide the histone methyltransferase MLL1 at target HOXA gene promoters, and on the formation of chromatin loops that connect distantly expressed HOXA genes to HOTTIP transcripts.

A role of lncRNAs in chromatin loop formation has also been described for the lncRNA CCAT1-L (Xiang et al., 2014). Indeed, CCAT1-L, is transcribed from a locus upstream of MYC and plays a role in MYC transcriptional regulation by promoting long-range chromatin looping.

Thus, lncRNAs, through the recruitment of chromatin modifiers and/or the induction of chromatin loops will modulate the chromatin conformation and will format the genome in a particular configuration. This lncRNA-mediated genome formatting emerges as a crucial and fundamental mechanism by which lncRNA may act on gene expression programs.

H19, a prototype of a multitask lncRNA. As discussed above, lncRNAs can regulate genome expression through different molecular mechanisms. However, several lncRNAs use multiple strategies that, in combination, may be required for their biological function. The action of the lncRNA H19 on gene expression illustrates the complexity of the combinatorial mechanisms of regulation achieved by a single lncRNA. H19 was the first lncRNA discovered (Brannan et al., 1990). Furthermore, H19 and its neighboring IGF2 gene located at position 11p15.5 are subjected to genomic imprinting and the study of the gene regulation at this locus serves as a model for understanding the molecular mechanisms involved in this genomic regulation. In addition, alterations of gene expression at the H19/IGF2 locus are associated to malignancies and developmental disorders. Loss of heterozygosity including loss of imprinting could be responsible for a loss of expression or a biallelic expression of these genes. Patients suffering from Beckwith-Wiedemann syndrome (BWS, OMIM 130650; Choufani et al., 2010) exhibit a loss of H19 expression and a biallelic expression of IGF2. BWS is associated with fetal and postnatal overgrowth and increased risk of embryonic or childhood cancers such as Wilm's tumors. Loss of IGF2 expression with a biallelic H19 expression is responsible for 20 to 60% of cases of Silver-Russel syndrome (SRS, OMIM 180860; Penaherrera et al., 2010). SRS is an intrauterine growth delay associated to an altered postnatal growth with facial dysmorphism and corporal asymmetry. Numerous studies including ours indicate that H19 may play a key role in tumorigenesis and could contribute to tumor progression and aggressiveness. H19 overexpression has also been reported in various cancer tissues including breast (Adriaenssens et al., 1998; Lottin et al., 2002), bladder (Cooper et al., 1996), lung (Kondo et al., 1995), and esophageal cancers (Hibi et al., 1996). Several lines of evidence indicate that H19 could play a role in tumor invasion and angiogenesis. In breast cancer, the oncogenic role of H19 has been well established (Berteaux et al., 2005), even if the precise molecular mechanisms involved in tumorigenesis are not yet fully understood.

At the H19/IGF2 locus, both genes share a common set of enhancers located downstream of the H19 gene (Figure 4). The ICR (imprinting control region), located 2 kbp upstream of the H19 promoter, controls the monoallelic expression of H19 and IGF2 by insulating communication between the 3' enhancers and the IGF2 promoter. The chromatin insulator

property of the H19/IGF2 ICR is regulated by the insulator CTCF (CCTC-binding factor), which binds specifically to the unmethylated maternal allele. On the paternal allele, the ICR methylation does not allow CTCF binding and leads to IGF2 expression (reviewed in Lewis and Murrell, 2004). The H19/IGF2 locus contains other differentially methylated regions (DMRs), with DMR1 being a methylation-sensitive silencer and DMR2 being a methylation-sensitive activator (Constancia et al., 2001; Murrell et al., 2004). CTCF binding to the maternal ICR regulates its interaction with matrix attachment region 3 (MAR3) and DMR1 at IGF2, thus forming a tight loop around the maternal IGF2 locus which may contribute to its silencing. These interactions restrict the physical access of distal enhancers to the IGF2 promoter (Weber et al., 2003; Murrell et al., 2004; Kurukuti et al., 2006). Furthermore, several lncRNAs are produced at the H19/IGF2 locus adding further complexity to the locus regulation. The first antisense transcript at the H19/IGF2 locus is the lncRNA IGF2-AS (3-4 kb) discovered in 1991 in chicken (Rivkin et al., 1993; Moore et al., 1997). IGF2-AS and IGF2 are coregulated at the transcriptional levels but the function of this IGF2-AS lncRNA remains unclear. The lncRNA 91H (about 120 kb) is transcribed from the maternal allele (Berteaux et al., 2008). Recently, at the same position, a new protein coding gene HOTS (6 kbp) has been described (Onyango et al., 2011) but the relationship between the HOTS and 91H is still not clear. However, these two transcripts are transcribed in an antisense orientation compared to H19. An additional lncRNA produced by the H19/IGF2 locus has been identified (Court et al., 2011). This PIHit (paternally expressed IGF2/H19 intergenic transcript) lncRNA is a 5-6 kb transcript expressed from the paternal allele after birth. Thus, the genomic organization of coding and non-coding transcripts illustrates the complexity of the interleaved networks of lncRNAs expressed from the H19/IGF2 locus.

To complicate matters further, H19 lncRNA mechanisms of action appear to be extremely diverse, acting at various levels (Figure 5). H19 has been shown to guide chromatin modifying enzymes to specific loci. In particular, Luo et al. (2013) have shown that H19 binds to and recruit the histone methyltransferase EZH2 at the E-cadherin promoter, leading to an increase in H3K27me3 repressive marks and to the silencing of the E-cadherin gene in bladder cancer. PRC2 protein members are not the only chromatin modifying factors interacting with H19 since it has been shown that this lncRNA physically binds to the methyl-CpG-binding domain protein 1 (MBD1). The H19-MBD1 complex is then recruited at several imprinted genes including *IGF2*, *SLC38A4* and *PEG1* (Monnier et al., 2013). This recruitment also induces methylation at lysine 9 of histone H3 (H3K9me3), probably via the additional interaction with an H3K9 histone methyltransferase.

The multifaceted action of H19 is also illustrated by its dual interaction with miR pathways. On one hand, the lncRNA H19 acts as miR sponge to sequester miR-106a as well as the mir-let7 family members (Kallen et al., 2013; Imig et al., 2015). On the other end, H19 serves as a

precursor of miR-675 that will in turn, post-translationally regulate a number of targets involved in cell tumorigenicity, including RB, IGFR1, SMAD1, SMAD5, CDC6, NMO1 or RUNX1 (Cai and Cullen, 2007; Tsang et al., 2010; Keniry et al., 2012; Gao et al., 2012; Zhuang et al., 2014; Dey et al., 2014). The role of H19 in tumor progression could also be mediated through its interaction with the tumor-suppressor TP53 protein. This association results in partial TP53 inactivation (Yang et al., 2012).

Several evidences also indicate that the H19 lncRNA controls IGF2 expression at the translational and/or post-translational levels (Li et al., 1998), suggesting that other mechanisms by which H19 exerts its action remain to be deciphered. Similarly, the possible role of RNA duplex formation between H19 and the antisense transcripts 91H and HOTS requires investigations.

LncRNAs in human diseases. Given the wide range of molecular actions achieved by the lncRNAs and their roles in various physiological processes, it is not surprising that they have been shown to be involved in many human diseases. A number of data indicate that alterations of lncRNA expression lead to tumorigenesis through changes at the chromatin, transcriptional or post-transcriptional levels that impact target genes expression (Table I). Since lncRNAs are regulating a different cellular pathways, growing evidences suggest that they could play a role in a large number of other human disorders including metabolic diseases, neurodegenerative and psychiatric disorders, cardiovascular and immune dysfunctions (Esteller, 2011; Harries, 2012; Shi et al., 2013; Clark and Blackshaw, 2014).

Perspectives and concluding remarks. lncRNAs represent a large part of the transcriptome and a very heterogeneous class of transcripts in terms of genomic organization and modes of action. Many of them are considered as key regulators of gene expression and thus, lncRNAs constitute an additional layer controlling the cellular programs. lncRNAs regulate diverse expression steps at the levels of chromatin rearrangement, transcriptional control and/or post-transcriptional processing. By these actions, lncRNAs are involved in numerous physiological functions and in many cases lncRNA alterations are associated with human disorders.

The fact that lncRNAs can be deregulated in tumors and other human pathologies, make them attractive candidates as biomarkers and as targets for therapy. lncRNAs may be down-regulated at the RNA levels by targeting their sequence. As so, short interfering RNAs (siRNAs) designed to perfectly match exact stretches of nucleotides, guarantee a high degree of specificity leading to lncRNA degradation. The power of the siRNA approach is illustrated by the success of a number of preclinical studies where siRNAs targeted mRNAs (Kaur et al., 2014). Similar approaches can thus be envisioned to target non-coding RNAs. Indeed, siRNAs have also been used to target miRs, leading to heart regeneration in an in vivo mouse model

(Aguirre et al., 2014) and the use of siRNAs has been proposed in a therapeutic strategy targeting the lncRNA HOTAIR in endometrial carcinoma (Huang et al., 2014). Similarly, antisense oligonucleotides, single-strand DNA or RNA molecules of 8 to 50 nucleotides can be used to target lncRNA. Specifically, in vivo and in vitro experiments revealed that antisense oligonucleotides directed against the lncRNA MALAT1 inhibit its expression and drastically reduce lung cancer metastasis (Gutschner et al., 2013; Tripathi et al., 2013).

In this context, further exploration in the complexity of the lncRNA world promises the emergence of novel therapeutic opportunities.

ACKNOWLEDGEMENTS. This work is supported by Inserm, the University of Lille, and grants from INCa (PLBio 2010), le Comité du Nord de la Ligue Contre le Cancer and l'ITMO Biologie Cellulaire, Développement et Evolution (BCDE).

REFERENCES.

- Adriaenssens, E., Dumont, L., Lottin, S., Bolle, D., Leprêtre, A., Delobelle, A., et al. (1998) H19 overexpression in breast adenocarcinoma stromal cells is associated with tumor values and steroid receptor status but independent of p53 and Ki-67 expression. *Am. J. Pathol.* 153, 1597-1607.
- Anderson, D. M., Anderson, K. M., Chang, C. L., Makarewich, C. A., Nelson, B. R., McAnally, J. R., et al. (2015) A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell* 160, 595-606.
- Annilo, T., Kepp, K., and Laan, M. (2009) Natural antisense transcript of natriuretic peptide precursor A (NPPA): structural organization and modulation of NPPA expression. *BMC Mol. Biol.* 10, 81.
- Baek, D., Villén, J., Shin, C., Camargo, F. D., Gygi, S. P., and Bartel, D. P. (2008) The impact of microRNAs on protein output. *Nature* 455, 64–71.
- Barry, G., Briggs, J. A., Vanichkina, D. P., Poth, E. M., Beveridge, N. J., Ratnu, V. S., Nayler, S. P., Nones, K., Hu, J., Bredy, T. W., Nakagawa, S., Rigo, F., Taft, R. J., Cairns, M. J., Blackshaw, S., Wolvetang, E. J., and Mattick, J. S. (2014) The long non-coding RNA Gomafu is acutely regulated in response to neuronal activation and involved in schizophrenia-associated alternative splicing. *Mol. Psychiatry.* 19, 486-494.
- Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* 136, 215–233.
- Beltran, M., Puig, I., Peña, C., García, J. M., Alvarez, A. B., Peña, R., Bonilla, F., and de Herreros, A. G. (2008) A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. *Genes Dev.* 22, 756-769.
- Bernstein, H. D., Zopf, D., Freymann, D. M., and Walter, P. (1993) Functional substitution of the signal recognition particle 54-kDa subunit by its *Escherichia coli* homolog. *Proc. Natl. Acad. Sci. USA* 90, 5229-5233.
- Berteaux, N., Aptel, N., Cathala, G., Genton, C., Coll, J., Daccache, A., Spruyt, N., Hondermarck, H., Dugimont, T., Cury, J. J., Forné, T., and Adriaenssens, E. (2008) A novel H19 antisense RNA overexpressed in breast cancer contributes to paternal IGF2 expression. *Mol. Cell. Biol.* 28, 6731-6745.
- Berteaux, N., Lottin, S., Monté, D., Pinte, S., Quatannens, B., Coll, J., Hondermarck, H., Cury, 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, 29625-2936.

- Bertone, P., V. Stolc, T. E. Royce, J. S. Rozowsky, A. E. Urban et al. (2004) Global identification of human transcribed sequences with genome tiling arrays. *Science* 306, 2242–2246.
- 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.
- Brown, C.J., Hendrich, B.D., Rupert, J.L., Lafrenière, R.G., Xing, Y., Lawrence, J., and Willard HF (1992) The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. *Cell* 71, 527–542.
- Cabili, M. N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A, et al. (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 25, 1915-1927.
- Cai, X., and Cullen, B. R. (2007) The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA* 13, 313-316.
- Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M. C., Maeda, N., et al.; FANTOM Consortium; RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group). (2005) The transcriptional landscape of the mammalian genome. *Science* 309, 1559-1563.
- Carrieri, C., Cimatti, L., Biagioli, M., Beugnet, A., Zucchelli, S., et al. (2012) Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature* 491, 454-457.
- Cesana, M., Cacchiarelli, D., Legnini, I., Santini, T., Sthandier, O., Chinappi, M., et al. (2011) A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* 147, 358-369.
- Choufani, S., Shuman, C., and Weksberg, R. (2010) Beckwith-Wiedemann syndrome. *Am. J. Med. Genet. C Semin. Med. Genet.* 154C, 343-354.
- Clark, B.S., and Blackshaw S. (2014) Long non-coding RNA-dependent transcriptional regulation in neuronal development and disease. *Front. Genet.* 5, 164.
- Coccia, E. M., Cicala, C., Charlesworth, A., Ciccarelli, C., Rossi, G. B, Philipson, L., and Sorrentino, V. (1992) Regulation and expression of a growth arrest-specific gene (gas5) during growth, differentiation, and development. *Mol. Cell. Biol.* 12, 3514–3521.
- 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, 203-206.
- Cooper, M. J., Fischer, M., Komitowski, D., Shevelev, A., Schulze, E., Ariel, I., et al. (1996) Developmentally imprinted genes as markers for bladder tumor progression. *J. Urol.* 155, 2120-2127.

- Court, F., Baniol, M., Hagege, H., Petit, J. S., Lelay-Taha, M. N., Carbonell, F., et al. (2011) Long-range chromatin interactions at the mouse Igf2/H19 locus reveal a novel paternally expressed long non-coding RNA. *Nucleic Acids Res.* 39, 5893-5906.
- Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., et al. (2012) The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 22, 1775-1189.
- Dey, B.K., Pfeifer, K., and Dutta, A. (2014) The H19 long noncoding RNA gives rise to microRNAs miR-675-3p and miR-675-5p to promote skeletal muscle differentiation and regeneration. *Genes Dev.* 28, 491-501.
- Eis, P. S., Tam, W., Sun, L., Chadburn, A., Li, Z., Gomez, M. F., Lund, E., and Dahlberg J. E. (2005) Accumulation of miR-155 and BIC RNA in human B cell lymphomas. *Proc. Natl. Acad. Sci. USA* 102, 3627-3632.
- ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74.
- Esteller, M. (2011) Non-coding RNAs in human disease. *Nat. Rev. Genet.* 12, 861-874.
- Faghihi, M. A., Modarresi, F., Khalil, A. M., Wood, D. E., Sahagan, B. G., Morgan, T. E., et al. (2008) Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feedforward regulation of β -secretase. *Nature Med.* 14, 723–730.
- Faghihi, M. A., Zhang, M., Huang, J., Modarresi, F., Van der Brug, M. P., Nalls, M. A., et al. (2010) Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biol.* 11, R56.
- Feng, J., Bi, C., Clark, B. S., Mady, R., Shah, P., and Kohtz, J. D. (2006) The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes Dev.* 20, 1470-1484.
- Gao, W. L., Liu, M., Yang, Y., Yang, H., Liao, Q., Bai, Y., et al. (2012) The imprinted H19 gene regulates human placental trophoblast cell proliferation via encoding miR-675 that targets Nodal Modulator 1 (NOMO1). *RNA Biol.* 9, 1002-1010.
- Gong, C., and Maquat, L.E. (2011) lncRNAs transactivate STAU1-mediated mRNA decay by duplexing with 3' UTRs via Alu elements. *Nature.* 470, 284-288.
- Gutschner, T., Hämmerle, M., Eissmann, M., Hsu, J., Kim, Y., Hung, G., et al. (2013) The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res* 73, 1180-1189.
- Guttman, M., Amit, I., Garber, M., French, C., Lin, M. F., Feldser, D., et al. (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 458, 223-227.

- Guttman, M., Garber, M., Levin, J. Z., Donaghey, J., Robinson, J., Adiconis, X., et al. (2010) Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat. Biotechnol.* 28, 503-510.
- Hagan, J. P., O'Neill, B.L., Stewart, C.L., Kozlov, S.V., and Croce, C.M. (2009) At least ten genes define the imprinted Dlk1-Dio3 cluster on mouse chromosome 12qF1. *PLoS One* 4, e4352.
- Harries, L. W. (2012) Long non-coding RNAs and human disease. *Biochem. Soc. Trans.* 40, 902-906.
- He, W., Cai, Q., Sun, F., Zhong, G., Wang, P., Liu, H., et al. (2013) linc-UBC1 physically associates with polycomb repressive complex 2 (PRC2) and acts as a negative prognostic factor for lymph node metastasis and survival in bladder cancer. *Biochim. Biophys. Acta* 1832, 1528-1537.
- Hibi, K., Nakamura, H., Hirai, A., Fujikake, Y., Kasai, Y., Akiyama, S., et al. (1996) Loss of H19 imprinting in esophageal cancer. *Cancer Res.* 56, 480-482.
- Houseley, J., Rubbi, L., Grunstein, M., Tollervey, D., and Vogelauer, M. (2008) A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. *Mol. Cell* 32, 685-695.
- Huang, J., Ke, P., Guo, L., Wang, W., Tan, H., Liang, Y., and Yao, S. (2014) Lentivirus-mediated RNA interference targeting the long noncoding RNA HOTAIR inhibits proliferation and invasion of endometrial carcinoma cells in vitro and in vivo. *Int. J. Gynecol. Cancer* 24, 635-642.
- Hung, T., Wang, Y., Lin, M. F., Koegel, A. K., Kotake, Y., Grant, G. D., et al. (2011) Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat. Genet.* 43, 621-629.
- Hutchinson, J. N., Ensminger, A. W., Clemson, C. M., Lynch, C. R., Lawrence, J.B., and Chess, A. (2007). A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* 8, 39.
- Iacobucci, I., Sazzini, M., Garagnani, P., Ferrari, A., Boattini, A., Lonetti, A., et al. (2011) A polymorphism in the chromosome 9p21 ANRIL locus is associated to Philadelphia positive acute lymphoblastic leukemia. *Leuk Res.* 2011 Aug;35(8):1052-9.
- Imig, J., Brunschweiler, A., Brümmer, A., Guennewig, B., Mittal, N., Kishore, S., et al. (2015) miR-CLIP capture of a miRNA targetome uncovers a lincRNA H19-miR-106a interaction. *Nat. Chem. Biol.* 11, 107-114.
- Ishizuka, A., Hasegawa, Y., Ishida, K., Yanaka, K., and Nakagawa, S. (2014) Formation of nuclear bodies by the lincRNA Gomafu-associating proteins Celf3 and SF1. *Genes Cells* 19, 704-721.

- Jia, H., Osak, M., Bogu, G. K., Stanton, L. W., Johnson, R., and Lipovich L. (2010) Genome-wide computational identification and manual annotation of human long noncoding RNA genes. *RNA* 16, 1478-1487.
- Kallen, A.N., Zhou, X. B., Xu, J., Qiao, C., Ma, J., Yan, L., Lu, L., et al. (2013) The imprinted H19 lncRNA antagonizes let-7 microRNAs. *Mol. Cell.* 52, 101-112.
- Kapranov, P., Willingham, A. T., and Gingeras, T. R. (2007) Genome-wide transcription and the implications for genomic organization. *Nat. Rev. Genet.* 8, 413-423.
- Kaur, I. P., Chopra, K., Rishi, P., Puri, S., and Sharma, G. (2014) Small RNAs: the qualified candidates for gene manipulation in diverse clinical pathologies. *Crit. Rev. Ther. Drug. Carrier Syst.* 31, 305-329.
- Keniry, A., Oxley, D., Monnier, P., Kyba, M., Dandolo, L., Smits, G., and Reik, W. (2012) The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. *Nat. Cell Biol.* 14, 659-665.
- Khalil, A. M., Guttman, M., Huarte, M., Garber, M., Raj, A., Rivea Morales, D., et al. (2009) Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci. USA* 106, 11667-11672.
- Khochbin, S., Brocard, M. P., Grunwald, D., and Lawrence, J. J. (1992) Antisense RNA and p53 regulation in induced murine cell differentiation. *Ann. N. Y. Acad. Sci.* 660, 77-87.
- Kino, T., Hurt, D. E. , Ichijo, T., Nader, N., and Chrousos, G. P. (2010) Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci. Signal* 3, ra8.
- Kondo, M., Suzuki, H., Ueda, R., Osada, H., Takagi, K., Takahashi, T., and Takahashi, T. (1995) Frequent loss of imprinting of the H19 gene is often associated with its overexpression in human lung cancers. *Oncogene* 10, 1193-1198.
- Korneev, S. A., Park, J. H., and O'Shea, M. (1999) Neuronal expression of neural nitric oxide synthase (nNOS) protein is suppressed by an antisense RNA transcribed from an NOS pseudogene. *J. Neurosci.* 19, 7711-7720.
- Kotake, Y., Nakagawa, T., Kitagawa, K., Suzuki, S., Liu, N., Kitagawa, M., and Xiong, Y. (2011) Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. *Oncogene* 30, 1956-1962.
- Kretz, M., Siprashvili, Z., Chu, C., Webster, D. E., Zehnder, A., Qu, K., Lee, C. S., et al., (2013) Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature* 493, 231-235.
- Krystal, G. W., Armstrong, B. C., and Battey, J. F. (1990) N-myc mRNA forms an RNA-RNA duplex with endogenous antisense transcripts. *Mol. Cell. Biol.* 10, 4180-4191.
- Kurukuti, S., Tiwari, V. K., Tavoosidana, G., Pugacheva, E., Murrell, A., Zhao, Z., et al. (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. USA* 103, 10684-10649.
- Lewis, A., and Murrell, A. (2004) Genomic imprinting: CTCF protects the boundaries. *Curr. Biol.* 14, R284-286.
- Li, Y. M., Franklin, G., Cui, H. M., Svensson, K., He, X. B., Adam, G., et al. (1998) The H19 transcript is associated with polysomes and may regulate IGF2 expression in trans. *J. Biol. Chem.* 273, 28247-28252.
- Liu, Z., Wang, W., Jiang, J., Bao, E., Xu, D., Zeng, Y., et al. (2013) Downregulation of GAS5 promotes bladder cancer cell proliferation, partly by regulating CDK6. *PLoS One* 8, e73991.
- Lottin, S., Adriaenssens, E., Dupressoir, T., Berteaux, N., Montpellier, C., Coll J., et al. (2002) Overexpression of an ectopic H19 gene enhances the tumorigenic properties of breast cancer cells. *Carcinogenesis* 23, 1885-1895.
- Luo, M., Li, Z., Wang, W., Zeng, Y., Liu, Z., and Qiu, J. (2013) Long non-coding RNA H19 increases bladder cancer metastasis by associating with EZH2 and inhibiting E-cadherin expression. *Cancer Lett.* 333, 213-221.
- Ma C, Nong K, Zhu H, Wang W, Huang X, Yuan Z, Ai K. H19 promotes pancreatic cancer metastasis by derepressing let-7's suppression on its target HMGA2-mediated EMT. *Tumour Biol.* 2014 Sep;35(9):9163-9.
- Maenner, S., Blaud, M., Fouillen, L., Savoye, A., Marchand, V., Dubois, A., et al. (2010) 2-D structure of the A region of Xist RNA and its implication for PRC2 association. *PLoS Biol.* 8, e1000276.
- Martianov, I., Ramadass, A., Serra Barros, A., Chow, N., and Akoulitchev, A. (2007) Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature* 445, 666-670.
- Mercer, T. R., Dinger, M. E., and Mattick, J. S. (2009) Long noncoding RNAs: insights into functions. *Nat. Rev. Genet.* 10, 155-159.
- Monnier, P., Martinet, C., Pontis, J., Stancheva, I., Ait-Si-Ali, S., and Dandolo, L. (2013) H19 lncRNA controls gene expression of the Imprinted Gene Network by recruiting MBD1. *Proc. Natl. Acad. Sci. USA* 110, 20693-20698.
- 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. USA* 94, 12509-12514.
- Mourtada-Maarabouni, M., Pickard, M. R., Hedge, V. L., Farzaneh, F., and Williams, G. T. (2009) GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. *Oncogene* 28, 195-208.

- 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, 889-893.
- Onyango, P., and Feinberg, A. P. (2011) A nucleolar protein, *H19* opposite tumor suppressor (*HOTS*), is a tumor growth inhibitor encoded by a human imprinted *H19* antisense transcript. *Proc. Natl. Acad. Sci. USA* 108, 16759-16764.
- Pandey, R. R., Mondal, T., Mohammad, F., Enroth, S., Redrup, L., Komorowski, J. et al. (2008) *Kcnq1ot1* antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation. *Mol. Cell.* 32, 232-246.
- Pang, K. C., Frith, M. C., and Mattick, J. S. (2006) Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. *Trends Genet.* 22, 1-5.
- Pasmant, E., Laurendeau, I., Héron, D., Vidaud, M., Vidaud, D., and Bièche, I. (2007) Characterization of a germ-line deletion, including the entire *INK4/ARF* locus, in a melanoma-neural system tumor family: identification of *ANRIL*, an antisense noncoding RNA whose expression coclusters with *ARF*. *Cancer Res.* 67, 3963-3969.
- Pasmant, E., Sabbagh, A., Masliah-Planchon, J., Ortonne, N., Laurendeau, I., Melin, L., et al.; NF France Network. (2011) Role of noncoding RNA *ANRIL* in genesis of plexiform neurofibromas in neurofibromatosis type 1. *J. Natl. Cancer Inst.* 103, 1713-1722.
- Pauli, A., Valen, E., Lin, M. F., Garber, M., Vastenhouw, N. L., Levin, J. Z., et al. (2012) Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. *Genome Res.* 22, 577-591.
- Penaherrera, M. S., Weindler, S., Van Allen, M. I., Yong, S. L., Metzger, D. L., McGillivray, B., et al. (2010) Methylation profiling in individuals with Russell-Silver syndrome. *Am. J. Med. Genet.* 152A, 347-355.
- Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W. J., and Pandolfi, P. P. (2010) A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* 465, 1033-1038.
- Prensner, J. R., Iyer, M. K., Balbin, O. A., Dhanasekaran, S. M., Cao, Q., Brenner, J. C., et al. (2011) Transcriptome sequencing across a prostate cancer cohort identifies *PCAT-1*, an unannotated lincRNA implicated in disease progression. *Nat. Biotechnol.* 29, 742-749.
- Reeves, M. B., Davies, A. A., McSharry, B. P., Wilkinson, G. W., and Sinclair, J. H. (2007) Complex I binding by a virally encoded RNA regulates mitochondria-induced cell death. *Science* 316, 1345-1348.
- Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Bruggmann, S.A., et al. (2007) Functional demarcation of active and silent chromatin domains in human *HOX* loci by noncoding RNAs. *Cell.* 129, 1311-1323.

- Rivkin, M., Rosen, K. M., and Villa-Komaroff L. (1993) Identification of an antisense transcript from the IGF-II locus in mouse. *Mol. Reprod. Dev.* 35, 394-397.
- Schmidt, L. H., Spieker, T., Koschmieder, S., Schäffers, S., Humberg, J., Jungen, D., et al. (2011) The long noncoding MALAT-1 RNA indicates a poor prognosis in non-small cell lung cancer and induces migration and tumor growth. *J. Thorac. Oncol.* 6, 1984-1992.
- Shi, X., Sun, M., Liu, H., Yao, Y., and Song, Y. (2013) Long non-coding RNAs: a new frontier in the study of human diseases. *Cancer Lett.* 339, 159-166.
- Swiezewski, S., Liu, F., Magusin, A., and Dean, C. (2009) Coldinduced silencing by long antisense transcripts of an Arabidopsis Polycomb target. *Nature* 462, 799-802.
- Taft, R. J., Pang, K. C., Mercer, T. R., Dinger, M., and Mattick J. S. (2010) Non-coding RNAs: regulators of disease. *J. Pathol.* 220, 126-139.
- Tano, K., and Akimitsu, N. (2012) Long non-coding RNAs in cancer progression. *Front. Genet.* 3, 219.
- Tripathi, V., Ellis, J. D., Shen, Z., Song, D.Y., Pan, Q., Watt, A. T., Freier, S. M. et al. (2010). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol. Cell* 39, 925–938.
- Tripathi, V., Shen, Z., Chakraborty, A., Giri, S., Freier, S. M., Wu, X., et al. (2013) Long noncoding RNA MALAT1 controls cell cycle progression by regulating the expression of oncogenic transcription factor B-MYB. *PLoS Genet.* 9, e1003368.
- Tsai, M. C., Manor, O., Wan, Y., Mosammamaparast, N., Wang, J. K., Lan, F., Shi, Y., et al., (2010) Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 329, 689-693.
- Tsang, W. P., Ng, E. K., Ng, S. S., Jin, H., Yu, J., Sung, J. J., and Kwok, T. T. (2010) Oncofetal H19-derived miR-675 regulates tumor suppressor RB in human colorectal cancer. *Carcinogenesis* 31, 350-358.
- Völkel, P., and Angrand, P. O. (2007) The control of histone lysine methylation in epigenetic regulation. *Biochimie* 89, 1-20.
- Völkel, P., Dupret, B., Le Bourhis, X. and Angrand, P. O. (2015) Diverse involvement of EZH2 in cancer epigenetics. *Am. J. Transl. Res.* 7, 175-193.
- Wang, J., Liu, X., Wu, H., Ni, P., Gu, Z., Qiao, Y., et al. (2010) CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. *Nucleic Acids Res.* 38, 5366-5383.
- Wang, K. C., Yang, Y. W., Liu, B., Sanyal, A., Corces-Zimmerman, R., Chen, Y., et al. (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 472, 120-124.

- Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., et al. (2008) Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature* 454, 126-130.
- Wang, Y., Xu, Z., Jiang, J., Xu, C., Kang, J., Xiao, L., et al., (2013) Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal. *Dev. Cell.* 25, 69-80.
- Wapinski, O., and Chang, H. Y. (2011) Long noncoding RNAs and human disease. *Trends Cell. Biol.* 21, 354-361.
- Weber, M., Hagège, H., Murrell, A., Brunel, C., Reik, W., Cathala, G., and Forné, T. (2003) Genomic imprinting controls matrix attachment regions in the Igf2 gene. *Mol. Cell. Biol.* 23, 8953-8959.
- Willingham, A. T., Orth, A. P., Batalov, S., Peters, E. C., Wen, B. G., Aza-Blanc, P., et al. (2005) A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* 309, 1570-1573.
- Wilusz, J. E., Sunwoo, H., and Spector, D. L. (2009) Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev.* 23, 1494-1504.
- Xiang, J. F., Yin, Q. F., Chen, T., Zhang, Y., Zhang, X. O., Wu, Z., et al. (2014) Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. *Cell Res.* 24, 513-531.
- Yang, F., Bi, J., Xue, X., Zheng, L., Zhi, K., Hua, J., and Fang, G. (2012) Up-regulated long non-coding RNA H19 contributes to proliferation of gastric cancer cells. *FEBS J.* 279, 3159-3165.
- Yang, F., Zhang, L., Huo, X. S., Yuan, J. H., Xu, D., Yuan, S. X., et al. (2011) Long noncoding RNA high expression in hepatocellular carcinoma facilitates tumor growth through enhancer of zeste homolog 2 in humans. *Hepatology* 54, 1679-1689.
- Yap, K. L., Li, S., Muñoz-Cabello, A. M., Raguz, S., Zeng, L., Mujtaba, S., Gil, J., et al. (2010) Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol. Cell* 38, 662-674.
- Yoon, J. H., Abdelmohsen, K., Srikantan, S., Yang, X., Martindale, J. L., De, S., et al. (2012) LincRNA-p21 suppresses target mRNA translation. *Mol. Cell* 47, 648-655.
- Zhang, K., Sun, X., Zhou, X., Han, L., Chen, L., Shi, Z., et al. (2015) Long non-coding RNA HOTAIR promotes glioblastoma cell cycle progression in an EZH2 dependent manner. *Oncotarget* 6, 537-546.
- Zhao, J., Sun, B. K., Erwin, J. A., Song, J. J., and Lee, J.T. (2008) Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* 322, 750-756.

- Zhao, X., Tang, Z., Zhang, H., Atianjoh, F. E., Zhao, J. Y., Liang, L., et al. (2013) A long noncoding RNA contributes to neuropathic pain by silencing *Kcna2* in primary afferent neurons. *Nat. Neurosci.* 16, 1024-1031.
- Zhuang, M., Gao, W., Xu, J., Wang, P., and Shu, Y. (2014) The long non-coding RNA H19-derived miR-675 modulates human gastric cancer cell proliferation by targeting tumor suppressor RUNX1. *Biochem. Biophys. Res. Commun.* 448, 315-322.

TABLE

Table I. Examples of lncRNAs associated with human disorders

lncRNA	Cancer/disease	Mechanisms of Action	References
ANRIL	Neurofibromatosis type 1, prostate cancer, melanoma, acute lymphoblastic leukemia	Chromatin modification via the recruitment of the PRC2 at the INKB/ARF/INK4A tumor suppressor locus	Pasmant et al., 2007; 2011; Yap et al., 2010; Iacobucci et al., 2011
HOTAIR	Hepatocellular carcinoma, colorectal cancer, breast cancer, glioblastomas	Chromatin modification via the recruitment of PRC2 and LSD1 in trans.	Gupta et al., 2010; Yang et al., 2011; Kogo et al., 2011; Zhang et al., 2015
H19	Colorectal, gastric, breast, lung, esophageal, bladder, pancreas, ovary cancers	Chromatin modification via the recruitment of PRC2; Decoy for miR-Let-7; source of miR-675; TP53 inactivation	Kondo et al., 1995; Hibi et al., 1996; Cooper et al., 1996; Lottin et al., 2002; Berteaux et al., 2005; Tsang et al., 2010; Yang et al., 2012; Luo et al., 2013; Ma et al., 2014; Zhuang et al., 2014
HEIH	Hepatocellular carcinoma	Chromatin modification via the recruitment of PRC2	Yang et al., 2011
PCAT-1	Prostate cancer	Chromatin modification via the recruitment of PRC2	Prensner et al., 2011
linc-UBC1	Bladder cancer	Chromatin modification via the recruitment of PRC2	He et al., 2013
BACE1-AS	Alzheimer's disease	Increase in mRNA stability	Faghihi et al., 2008
GAS5	Breast, bladder cancers	Decoy for the glucocorticoid receptor; regulation of CDK6 expression	Mourtada-Maarabouni et al. 2009; Kino et al., 2010
PTENP1	Prostate cancer	miR decoy	Poliseno et al., 2010
KCNA2-AS	Neuropathic pain	Decrease of KCNA2 expression	Zhao et al., 2013
MIAT	Schizophrenia	Component of the nuclear matrix involved in mRNA splicing	Barry et al., 2014
MALAT1	Lung cancer	Alternative splicing regulation	Schmidt et al., 2011

FIGURE LEGENDS.

Figure 1. Functional mechanisms of lncRNA action at the post-transcriptional levels. (A) mRNA stabilization. Base pairing between specific regions of a long non-coding antisense RNA and its sense transcript induces stabilization of the target mRNA and increases protein abundance. (B) mRNA degradation. Staufen double-stranded RNA-binding protein 1 (STAU1)-mediated mRNA decay is induced when base pairing is formed between the mRNA and a lncRNA. (C) Ribosome targeting. Through homologous base pairing with mRNAs and interactions with ribosomal proteins lncRNAs target transcripts to ribosomes or prevent translation. (D) Regulation of splicing. Base pairing between mRNAs and lncRNAs may prevent splicing by masking the splicing sites. In addition, lncRNAs are also implicated in the formation and maintenance of nuclear structures involved in alternative splicing of nascent transcripts. (E) miR sponge. By sequestering miRs through base pairing formations, lncRNAs affect the expression of the miR target genes. (F) Precursor of miRs. lncRNAs can serve as a source of miRs after processing.

lncRNAs are shown in red, whereas mRNAs are in blue. See text for examples.

Figure 2. Functional mechanism of action at the levels of transcriptional regulation. (A) lncRNA may regulate transcription by virtue of RNA-DNA triplex formation preventing the formation of the transcription initiation complex at promoters. (B) lncRNAs can act as decoys by titrating transcription factors away from their cognate promoters. (C) lncRNAs can regulate transcription through the targeting of transcription factors to promoters or acting as co-factors involved in transcription factor activity. (D) lncRNA can also control transcription factor trafficking.

lncRNAs are shown in red.

Figure 3. Examples of lncRNAs controlling chromatin organization. (A) HOTAIR represses transcription in trans by recruiting two different chromatin modifying activities. The PRC2 complex produces the repressive H3K27me3 marks, whereas the LSD1-CoREST complex is responsible for the removal of the active H3K4me2/3 marks. (B) The ANRIL lncRNA represses transcription in cis at the INK4B/ARF/INK4A locus by recruiting the Polycomb repressive complexes PRC1 and PRC2. (C) The HOTTIP lncRNA activates genes by recruiting the histone modifier complex WDR5-MLL which is responsible for H3K4me3 methylation, and by mediating long-range chromatin looping at one extremity of the HOXA locus.

Purple hexagons represent H3K27me3 repressive marks, whereas green hexagons illustrate H3K4me3 activating marks.

Figure 4. Schematic representation of the transcriptional complexity at the H19/IGF2 locus. Non-coding transcripts at the H19/IGF2 are shown as blue squares when they are expressed. Coding genes are in green, when expressed. The differences in gene expression between the paternal and maternal alleles are shown. The DNA methylation status of the regulatory elements ICR (imprinting control region) and DMRs (differentially methylated regions) is indicated for the paternal and maternal alleles.

Figure 5. The multifaceted action of the lncRNA H19. The lncRNA H19 controls genome expression at multiple levels. H19 acts on chromatin organization through the recruitment of chromatin modifying complex PRC2 (1) and on post-transcriptional control as a miR decoys sequestering miR-106a and miR-let7 (2) or as a precursor for miR-675-5p and miR-675-3p (3). H19 also interact with p53 (TP53) and inactivate the tumor suppressor protein action (4). Furthermore, possible base pairing between H19 and the antisense transcripts 91H and HOTS may have biological outcomes (5).

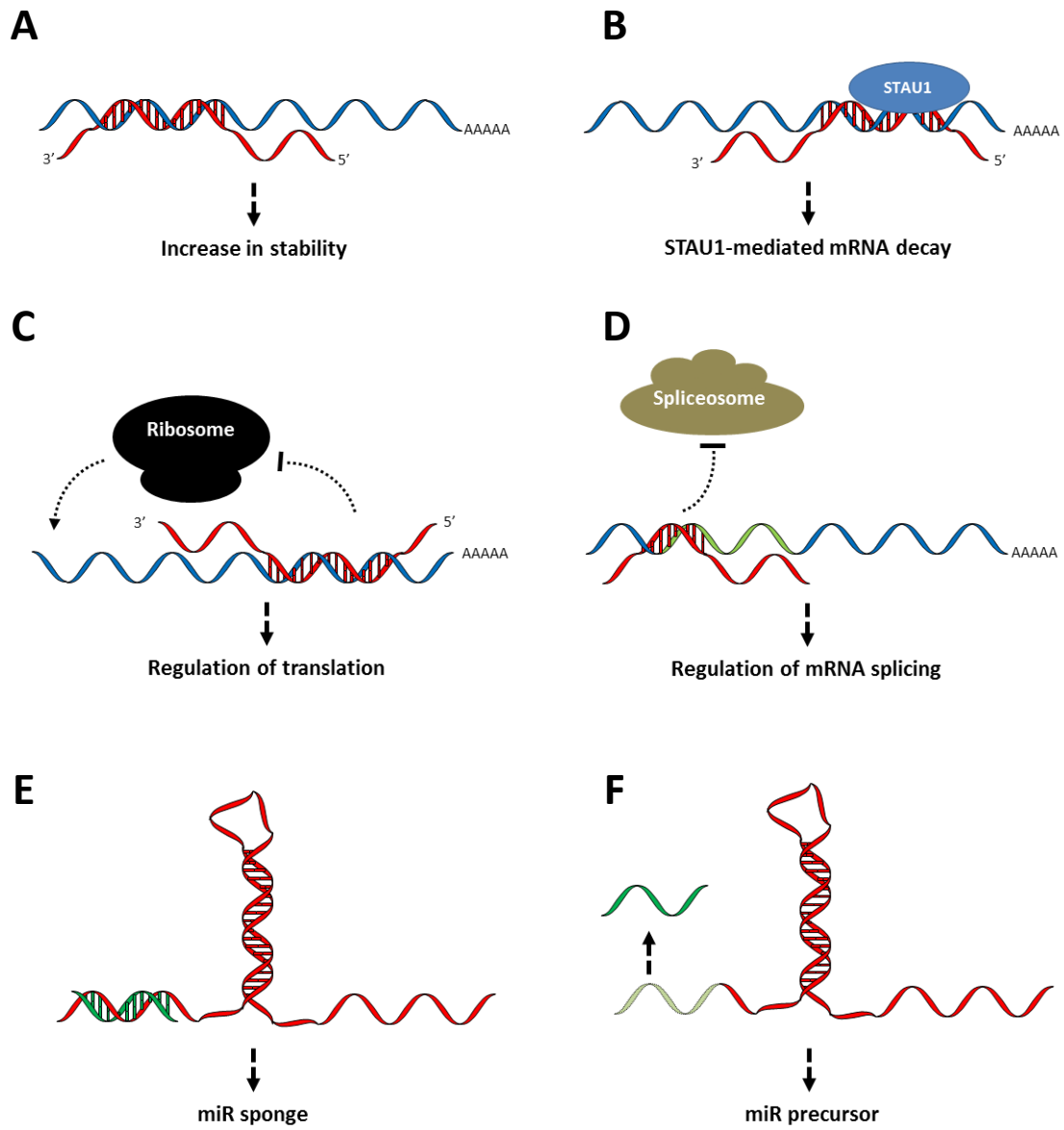


Figure 1.

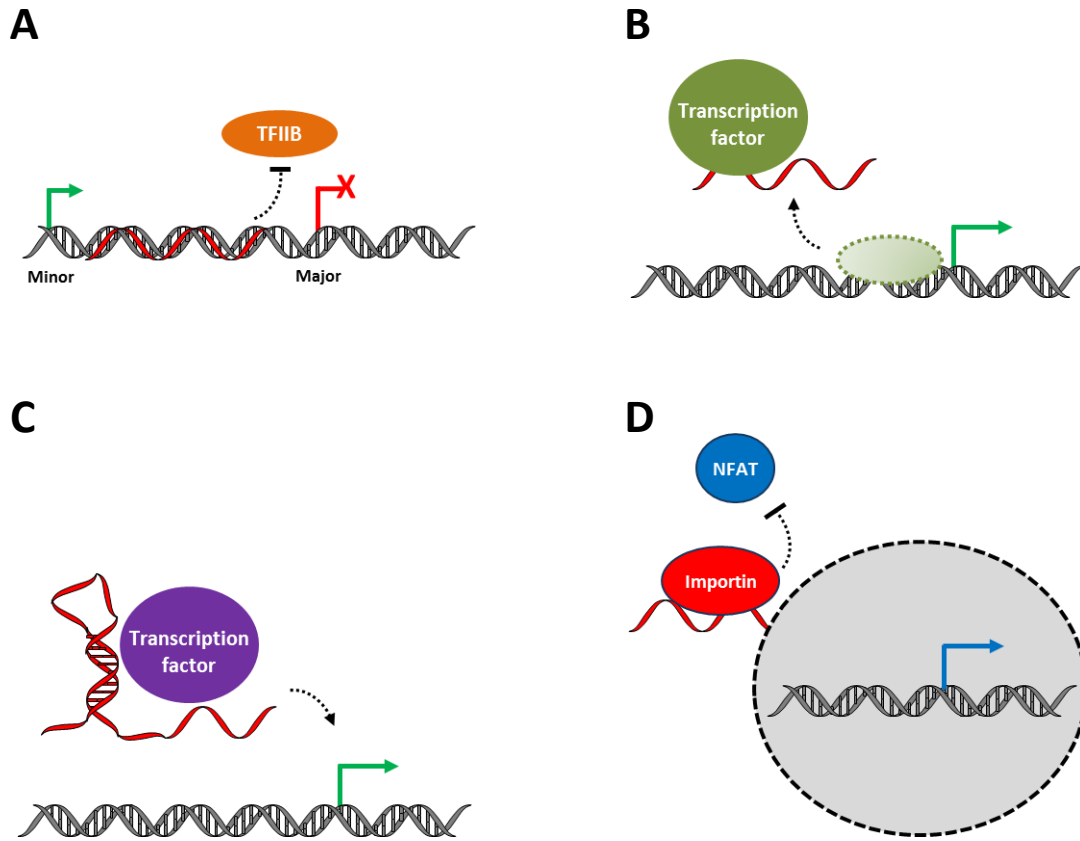


Figure 2.

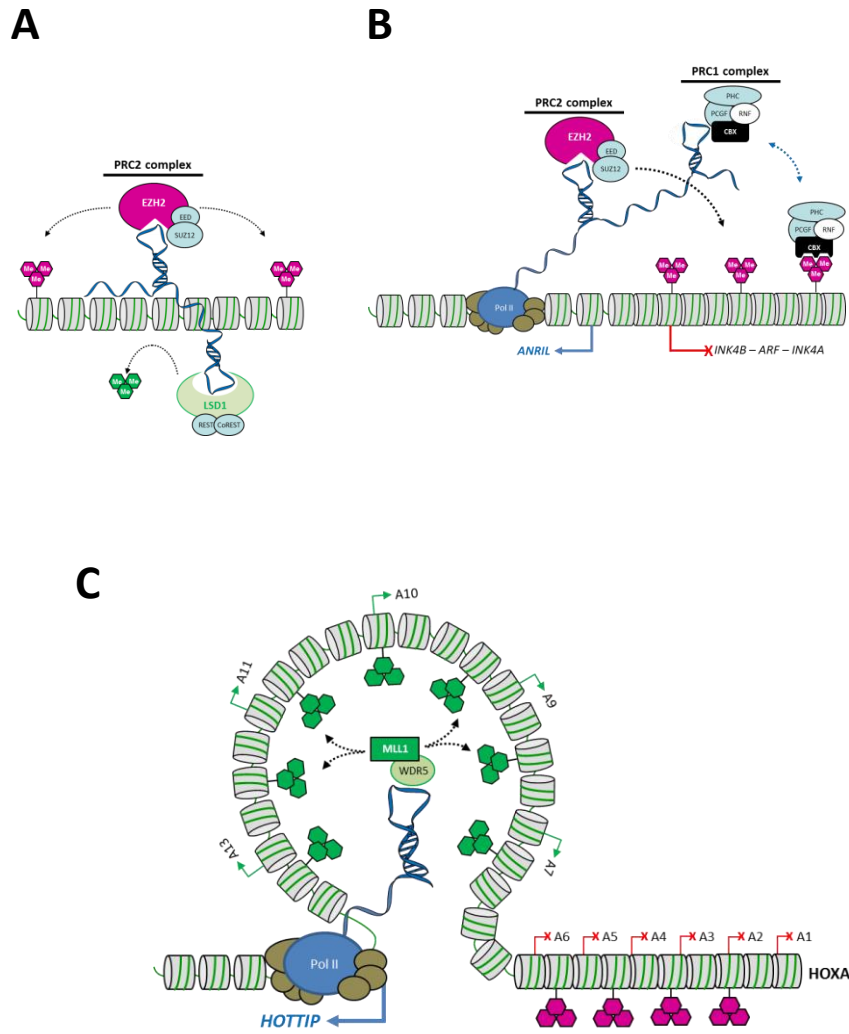


Figure 3.

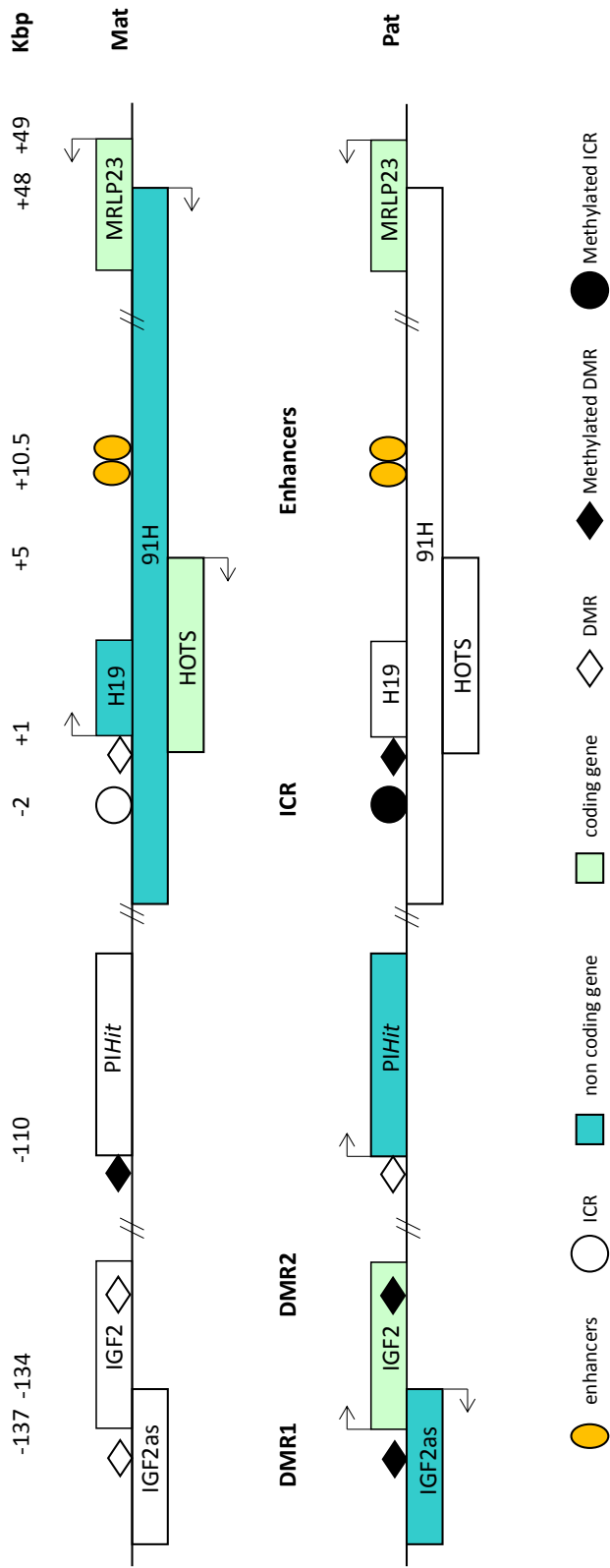


Figure 4.

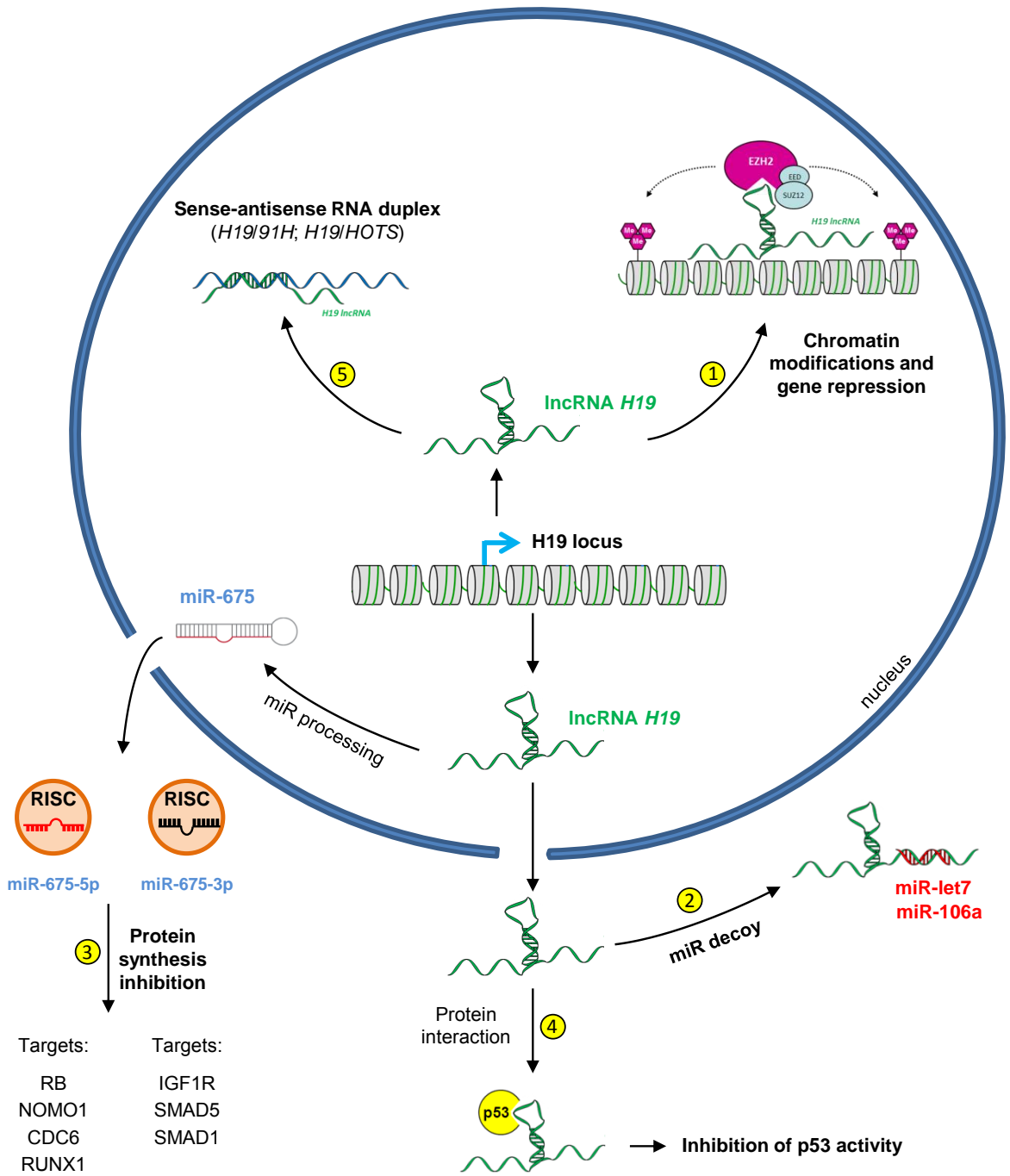


Figure 5.

RESUME

Le locus *H19/IGF2* est soumis à l’empreinte génomique. A partir de l’allèle maternel, deux ARN non codants sont exprimés : *H19* et *91H*. Ces ARN sont fréquemment surexprimés dans les cancers du sein mais leurs fonctions et/ou modes d’action lors de la tumorigenèse ne sont pas encore définis.

L’ARN *H19* est le précurseur de microARN (miR-675) dont aucune cible n’a été décrite dans la glande mammaire. Dans un premier temps, nous avons donc recherché et validé plusieurs cibles de ces microARN permettant d’expliquer les propriétés oncogéniques d’*H19*. En effet, le miR-675 régule l’expression des gènes *c-Cbl*, *Cbl-b* et *FADD* dans les cellules cancéreuses mammaires. Ces nouvelles régulations permettent l’augmentation de la prolifération cellulaire, la migration/invasion mais aussi la résistance aux apoptogènes. Dans cette étude, de manière inattendue et surprenante, nous avons également découvert un nouveau mécanisme de recrutement des microARN. Nous montrons que les protéines associées aux ARNm peuvent favoriser le recrutement du complexe RISC, mais surtout moduler l’action de celui-ci.

De plus, dans certains cancers, l’ARN *H19* peut s’associer aux protéines, notamment la protéine P53 dans les cancers gastriques, pour moduler leur fonction. Dans le cancer du sein, cette interaction empêche l’action de la protéine P53 fonctionnelle et favorise sa dégradation. Un nouveau mécanisme de régulation de P53 a donc été identifié. Ce mécanisme peut être responsable de certaines résistances tumorales aux thérapies.

Par ailleurs, nous avons mis en évidence le rôle du gène *H19* dans la formation et le maintien des cellules souches. Cette nouvelle fonction peut permettre à l’ARN *H19* de promouvoir l’initiation ou la récurrence tumorale. Néanmoins, le mécanisme associé à l’ARN *H19* dans cette population n’est pas encore établi.

Pour finir, les fonctions du long ARN non codant *91H* dans les cellules cancéreuses mammaires ont été étudiées. Cet ARN possède des propriétés oncogéniques et permet l’expression des gènes *H19* et *IGF2* en modulant la conformation de la chromatine au locus.

Pour conclure, j’ai identifié plusieurs modes d’action de l’ARN *H19* lui permettant de promouvoir l’apparition de la tumeur, sa progression ou encore ses résistances aux thérapies. J’ai également défini le rôle de l’ARN *91H* au locus *H19/IGF2*, ainsi que sa fonction lors de la tumorigenèse.

Mots clefs : *H19*, *91H*, long ARN non codant, microARN, cancer du sein, cellules souches.