

UNIVERSITE DE LILLE NORD DE FRANCE  
LABORATOIRE INSERM U908

## **THESE DE DOCTORAT**

Pour l'obtention du titre de  
**DOCTEUR DE L'UNIVERSITE DE LILLE**  
Ecole Doctorale Biologie-Santé de Lille (ED-446)

**Etude des mécanismes impliqués dans la régulation de la  
tumorigenèse mammaire par le long ARN non codant *H19***

Présentée et soutenue publiquement par

**Jordan COLLETTE**

Le 16 Septembre 2019

### **Jury**

**Président :** Pr. Xuefen Le Bourhis

**Rapporteurs :** Pr. Pascale Cohen  
Dr. Evelyne Ségal-Bendirdjian

**Examineurs :** Dr. Thierry Forné  
Pr. Fabien Gosselet

**Directeur :** Pr. Eric Adriaenssens



## Remerciements

Je remercie la région Hauts-de-France ainsi que l'Université de Lille pour le financement de mes travaux de thèse.

Je remercie les membres de mon jury qui me font l'honneur d'évaluer mes travaux de thèse : Madame le Professeur **Pascale Cohen**, Madame le Docteur **Evelyne Ségal-Bendirdjian**, Monsieur le Professeur **Fabien Gosselet**, Monsieur le Docteur **Tierry Forné**. Je vous remercie pour avoir pris le temps de relire mon manuscrit et de me consacrer une journée pour assister à ma soutenance de thèse. Je remercie particulièrement le Professeur **Fabien Gosselet** qui m'a permis de découvrir le monde de la recherche au cours de mes stages de licence, au sein de son laboratoire.

Par la suite, je souhaite remercier la directrice de notre laboratoire, Madame le Professeur **Xuefen Le Bourhis**, de m'avoir accueillie dans son unité pour la réalisation de ma thèse et les discussions passionnantes que nous avons eu tout au long de celle-ci.

Je souhaite remercier mon directeur de thèse, Monsieur le Professeur **Eric Adriaenssens**, pour m'avoir encadré pendant mon Master II et ma thèse. Merci de m'avoir fait confiance pendant ses trois années de thèse qui m'ont permis de me développer au niveau personnel et professionnel. Merci également de m'avoir laissé une grande liberté, que ce soit au niveau scientifique où j'ai eu l'occasion de réaliser les expériences que je désirais mais également au niveau personnel lorsque j'ai pu choisir de partir en congrès aux Etats-Unis.

Un grand merci à **Constance Vennin** pour ta formation technique et scientifique au cours de mon master II. Merci du temps que tu m'as accordé lors de ta présence au laboratoire et de ta patience pour répondre à toutes mes questions ! Ta grande rigueur scientifique a été un exemple pour moi.

Je remercie également les membres de mon comité de suivi de thèse pour leurs réflexions très intéressantes qui ont permis d'améliorer la qualité de mes travaux. Monsieur le Docteur **Thierry Forné**, qui fait également parti de mon jury de thèse et qui a malheureusement assisté

à mes CST par Skype au cours de ma thèse. Ma soutenance de thèse sera l'occasion de discuter de vive voix. Le Docteur **Bruno Lefebvre**, qui a également participé à notre projet sur p53. J'en profite pour remercier son ancienne étudiante, le Docteur **Mélissa Giles**, qui m'a formé aux joies des expérimentations *in vitro* ! Merci au Docteur **Roland Bourette** pour ses précieux conseils.

Merci à **Valérie Cendré** pour son aide dans la gestion des problèmes administratifs et les nombreuses commandes UGAP réalisées ensemble ! Ta joie de vivre me manquera.

Je remercie l'ensemble des statutaires de l'équipe. Merci **Robert-Alain Toillon** pour tous les conseils sur les problèmes techniques et les protocoles partagés, **Chann Lagadec** pour les conseils sur l'étude des cellules souches cancéreuses et ses connaissances sur la Californie ! Merci **Pierre Olivier Angrand** pour ses discussions intéressantes sur les nouvelles technologies de biologie moléculaire, **Samuel Meignan** et **Valérie Chopin** pour vos suggestions lors des réunions hebdomadaires du laboratoire.

Merci aux anciens du laboratoire, **Jérémy Duval** pour sa bonne humeur et nos discussions quotidiennes en salle de culture, **Barbara Dupret** pour sa joie de vivre et son aide précieuse lorsque j'en ai eu besoin, **Justine Bailleul** pour ses nombreux conseils et son aide avec notre capricieux cytomètre en flux. Merci aussi pour cette visite express du parc de UCLA ! **Romain Lévêque** pour son aide pendant ces trois ans et ses conseils pour l'écriture de ma thèse !

Je remercie **Nadège Bidan**, ma collègue de promotion, pour son aide et son soutien ! Ces trois années sont passées si vite... Tu es la prochaine ! J'en profite pour remercier ta successeuse, **Marie Denoulet**, qui a repris mon poste de gestionnaire du stock des plastiques du laboratoire ! Bon courage pour ta thèse ! Merci également aux autres personnes que j'ai côtoyé au sein du laboratoire : Pamela, Bernadette, Isabelle, Ludivine, Mathilde, Marie W et les personnes de passage.

Merci aux membres de la **Team H19**, **Clément Lecerf** pour son aide technique dans la réalisation du projet p53 et toutes les fois où tu t'es occupé de mes précieuses cellules. Toutes nos conversations sur nos groupes de musique préférés, les derniers films et cette fameuse

dernière saison de Game of Thrones vont me manquer ! Bon courage pour la fin de ta thèse. Merci à **Evodie Peperstraete** pour ta bonne humeur quotidienne ! Je te souhaite un bon courage pour cette nouvelle aventure que sera ta thèse, tu seras bien encadrée. Merci à **Véronique Fafeur**, qui partage le bureau de la Team H19, pour ses discussions sur le monde de la recherche et le monde tout court.

Finalement je voudrais remercier toutes les personnes de mon entourage qui ont participé à la réussite de ma thèse ! Merci à ma famille pour son immense soutien dans les moments difficiles. **Mes grands-parents et mes parents** pour avoir fait de moi ce que je suis aujourd'hui. Je n'aurais jamais réussi ce doctorat sans vous. Un grand merci à **Mathilde** qui partage ma vie depuis bientôt trois ans et qui l'enjolive depuis neuf ! Tu as été essentielle à ma réussite et je te remercie pour ton aide dans la correction de ce manuscrit. Merci de m'avoir suivi jusque Los Angeles à la suite de mon congrès et pour cet inoubliable road trip dans l'Ouest Américain ! Je remercie également **Timo**, notre chat, qui m'a accompagné pendant ces longues soirées de rédaction !



# Etude des mécanismes impliqués dans la régulation de la tumorigenèse mammaire par le long ARN non codant *H19*

## Résumé

Le gène *H19* est soumis à l’empreinte génomique parentale et ne code aucune protéine. Le produit de ce gène, le long ARN non codant (lncRNA) *H19*, agit en tant qu’ARN et est impliqué dans le développement embryonnaire ainsi que dans la tumorigenèse. Le lncRNA *H19* est le précurseur du miR-675. Mes travaux de thèse ont permis d’identifier de nouveaux mécanismes impliqués dans la régulation de la tumorigenèse mammaire par le gène *H19*.

Nous avons mis en évidence que le lncRNA *H19* régule négativement la protéine p53 dans les cellules cancéreuses mammaires. Mes travaux ont démontré que le lncRNA *H19* interagit physiquement avec la protéine p53 et MDM2 afin d’induire sa dégradation et empêcher sa translocation dans le noyau. Ce nouveau mode d’action d’*H19* dans les cancers du sein pourrait expliquer le manque de pertinence clinique de l’étude du statut mutationnel de p53 par immunohistochimie dans ce cancer.

J’ai également mis en évidence que non seulement le lncRNA *H19* est impliqué dans la régulation des cellules souches cancéreuses, mais également le miR-675-5p. En effet, les tumeurs exprimant des signatures géniques associées aux marqueurs de cellules souches cancéreuses sont des tumeurs qui surexpriment le gène *H19*. De plus, la modulation de l’expression du lncRNA *H19* et de son microARN régule les capacités fonctionnelles associées aux cellules souches cancéreuses mammaires. Pour finir, j’ai initié un projet permettant l’identification, sans *a priori*, des gènes cibles du lncRNA *H19* et de son microARN dans les cellules cancéreuses mammaires.

Pour conclure, j’ai mis en évidence l’implication du lncRNA *H19* et du miR-675-5p dans différents processus impliqués dans la tumorigenèse mammaire.

Mots clefs : *H19*, cancer du sein, long ARN non codant, miR-675, p53, cellules souches cancéreuses.

Laboratoire INSERM U908. Adresse : INSERM U908, Batiment SN3 3<sup>ème</sup> étage, Cité Scientifique, 59650 Villeneuve d’Ascq, France.

# Implication of the long non coding RNA H19 in the regulation of breast tumorigenesis

## Abstract

The *H19* gene is subject to genomic imprinting and does not encode protein. The product of this gene, the long non coding RNA (lncRNA) *H19*, act as an RNA and is involved in development and the tumorigenesis. The *H19* RNA is the precursor of miR-675. My thesis work identified new mechanism involved in the regulation of breast tumorigenesis by *H19*.

We have demonstrated that the lncRNA *H19* negatively regulates the p53 protein in breast cancer cell lines. My work revealed that *H19* interacts with p53 and MDM2 to induce the degradation of p53 and impedes its nuclear localization. This new mechanism of *H19* in breast cancer could explain the lack of clinical relevance of the p53 mutational state measured by immunohistochemistry in breast cancer.

My work also revealed that not only the lncRNA *H19* is involved in the regulation of breast cancer stem cells but also the miR-675-5p. Indeed, we have shown a correlation between overexpression of *H19* and expression of a cancer stem cell phenotype in patient tumors. Furthermore, the modulation of *H19* or miR-675 expression regulates the functional capacities associated with breast cancer stem cells. I also initiated a project that will allow the identification of *H19* and miR-675 target genes in breast cancer cell lines.

To conclude, I highlighted the implication of the lncRNA *H19* and miR-675 in different process involved in breast cancer tumorigenesis.

Keywords: *H19*, breast cancer, long non coding RNA, miR-675, p53, cancer stem cells.

Inserm U908 laboratory. Adress : INSERM U908, Bâtiment SN3 3<sup>ème</sup> étage, Cité Scientifique, 59650 Villeneuve d'Ascq, France.



## SOMMAIRE

INTRODUCTION .....	3
I. Le cancer du sein .....	3
I.1. Généralités .....	3
<i>I.1.1. Epidémiologie du cancer du sein .....</i>	<i>3</i>
<i>I.1.2. Les facteurs de risque .....</i>	<i>4</i>
I.2. Histoire naturelle du cancer du sein .....	7
<i>I.2.1. Développement du cancer du sein .....</i>	<i>7</i>
<i>I.2.2. Origine du cancer : Importance des cellules souches cancéreuses .....</i>	<i>10</i>
I.3. Classifications des cancers du sein .....	15
<i>I.3.1. Classification clinique et grade des tumeurs .....</i>	<i>15</i>
<i>I.3.2. Classification moléculaire .....</i>	<i>16</i>
I.4. Stratégies thérapeutiques .....	17
<i>I.4.1. Les chimiothérapies .....</i>	<i>17</i>
<i>I.4.2. La radiothérapie .....</i>	<i>18</i>
<i>I.4.3. Les thérapies personnalisées .....</i>	<i>18</i>
<i>I.4.4 Résistances aux traitements .....</i>	<i>19</i>
II. Les ARN non codants .....	21
II.1. De la découverte de l'ADN au rôle primordiale des ARNs .....	21
II.2. De la poubelle à la lumière : la révolution des ARN non codants .....	24
<i>II.2.1. Concept d'ADN poubelle .....</i>	<i>24</i>
<i>II.2.2. Révolution du séquençage du génome Humain .....</i>	<i>24</i>
II.3. Classification des ARN non codants .....	26
<i>II.3.1. Les microARN .....</i>	<i>26</i>
<i>II.3.2. Les longs ARN non codants .....</i>	<i>34</i>
III. Le long ARN non codant <i>H19</i> .....	41
III.1. Historique du gène <i>H19</i> .....	41
<i>III.1.1. Découverte du gène H19 et de son action en tant qu'ARN non codant .....</i>	<i>41</i>
<i>III.1.2. Informations sur le locus H19/IGF2 .....</i>	<i>42</i>
<i>III.1.3. Le gène H19 produit un microARN : le miR-675 .....</i>	<i>44</i>
III.2. Modes d'action du long ARN non codant <i>H19</i> : exemple du cancer du sein .....	44
<i>III.2.1. Interaction avec des protéines .....</i>	<i>45</i>
<i>III.2.2. H19 est un précurseur de microARN .....</i>	<i>46</i>

III.2.3. <i>H19</i> est une éponge à microARN.....	48
III.2.4. Intérêts cliniques des niveaux d'expression du lncRNA <i>H19</i> et du miR-675.....	50
III.2.5. Rôle suppresseur de tumeur du gène <i>H19</i> .....	51
Objectifs .....	57
I. Le long ARN non codant <i>H19</i> inhibe p53 et favorise l'instabilité génomique .....	61
II. Le long ARN non codant <i>H19</i> et le miR-675 régule la population de cellules souches cancéreuses mammaires.....	91
III. Mise en évidence des gènes cibles du lncRNA <i>H19</i> et du miR-675 .....	107
III.1. Stratégie utilisée pour la mise en évidence des gènes cibles du lncRNA <i>H19</i> et de son microARN .....	107
III.2. Résultats préliminaires.....	109
III.2.1. Confirmation de la validité du séquençage.....	109
III.2.2. Identifications des cibles du miR-675 dans nos cellules.....	110
III.2.3. Perspectives.....	111
IV. Identification d'un nouveau mode d'action des microARN : le miR-675 augmente l'expression de son ARNm cible FADD. ....	115
Discussion.....	129
I. Le long ARN non codant <i>H19</i> régule négativement la protéine p53 et participe à l'instabilité génomique. ....	131
I.2. Développement de nouvelles thérapies .....	132
I.3. Rôle ambigu du gène <i>TP53</i> dans les cancers : gain de fonction oncogénique de p53 .	134
I.4. Pertinence du statut mutationnel de <i>TP53</i> dans les cancers du sein .....	136
I.5. Rôle central de <i>H19</i> dans la régulation du cycle cellulaire .....	139
II. Le lncRNA <i>H19</i> a des actions dépendantes et indépendantes de son microARN dans les cancers .....	141
II.1. Le lncRNA <i>H19</i> et son microARN sont impliqués dans l'émergence et le maintien des cellules souches cancéreuses mammaires.....	141
II.2. Identification de la part relative entre le lncRNA <i>H19</i> et son microARN.....	142
II.3. Identifications des gènes cibles du lncRNA <i>H19</i> et du miR-675 .....	144
BIBLIOGRAPHIE.....	147
Annexe.....	155

## Liste des Figures et Tableaux

Figures	Pages
Figure 1 : Représentation de la distribution des nouveaux cas et des décès des 10 cancers les plus communs en 2018	3
Figure 2 : Evolution de l'incidence et de la mortalité par cancer du sein de 1980 à 2012 en France métropolitaine	3
Figure 3 : Anatomie de la glande mammaire	7
Figure 4 : Etapes de la cancérisation d'un épithélium canalaire mammaire	8
Figure 5 : Représentation schématique des Hallmarks du cancer	9
Figure 6 : Représentation des modèles de développement tumoral	11
Figure 7 : Dogme central de la biologie moléculaire selon Crick	22
Figure 8 : Proportion des gènes codants pour des protéines et des séquences non codantes dans le génome humain	25
Figure 9 : Voie de biosynthèse canonique des microARN	28
Figure 10 : Voies de biosynthèse non canoniques des microARN	30
Figure 11: Différentes classes de lncRNAs chez les mammifères	34
Figure 12 Classification des long ARN non codants basée sur leurs localisations génomiques	35
Figure 13 : Représentation des principaux modes d'action des longs ARN non codants	36
Figure 14 : Estimation des niveaux d'ARNs dans une cellule mammaire	40
Figure 15 : Découvertes majeurs sur H19 et son implication dans le cancer du sein	41
Figure 16: Transcrits identifiés sur le locus H19/IGF2	43
Figure 17: H19 interagit avec EZH2 pour permettre le ciblage du complexe PRC2 sur des gènes cibles dans le cancer du sein	45
Figure 18 : Représentation du complexe ternaire formé par l'interaction entre H19, P53 et MDM2	46
Figure 19: Le miR-675-5p réprime les protéines Cbl-b et C-cbl dans le cancer du sein	46
Figure 20 : Représentation du rôle d'éponge à microARN de H19 dans le cancer du sein	48
Article p53 - Figure 1 H19 down-regulated P53 protein expression in breast cancer cells	82
Article p53 - Figure 2 Trimeric complex formation of H19/P53/MDM2	83-84
Article p53 - Figure 3 H19 prevented P53 nuclear translocation upon doxorubicin treatment	85
Article p53 - Figure 4 H19 repressed P53 activities	86
Article p53 - Figure 5 H19 increased mutation frequency measured by HPRT assay	87
Article p53 - Figure 6 Graphical representation of H19 effects on p53 activities	88
Figure 21: La surexpression du gène H19 augmente les capacités souches des cellules	100
Figure 22 : L'activité du promoteur du gène H19 permet d'identifier une population enrichie en cellules souches cancéreuses	101
Figure 23: Le lncRNA H19 est nécessaire au maintien des cellules souches cancéreuses mammaires	102
Figure 24 : Le miR-675-5p est nécessaire au maintien des cellules souches cancéreuses mammaires	103
Figure 25 : La surexpression du lncRNA H19 et du miR-675-5p accroît les capacités de migration et d'invasion des cellules can	104
Figure 26: Identification de 1342 cibles du miR-675-5p in silico dans la base de données TargetScan	107
Figure 27 : Représentation graphique du miR-CLIP	108
Figure 28 : Stratégie d'analyse des résultats obtenus par le RNAseq	109
Figure 29 : Analyse de l'expression de certains gènes HOX par qRT-PCR	110
Figure 30 : Cibles potentielles du miR-675 identifiées par notre RNAseq	110
Figure 32 : Nouveau mécanisme d'action des microARN	121
Figure 31 : H19, via le miR-675, augmente l'expression de l'ARNm de FADD	124
Figure 33: La protéine FXR1 régule le recrutement du miR-675 et sa fonction sur l'ARNm de FADD	125
Figure 34: H19, via la surexpression de FADD, promeut la résistance à l'apoptose et la prolifération	126
Figure 35 : H19 promeut la croissance tumorale et l'expression de FADD dans les tumeurs	127
Figure 36 : H19 promeut la tumorigenèse en augmentant l'expression génique	128
Figure 37: Spectre de mutation des gènes fréquemment mutés dans les cancers	135
Figure 38: Hétérogénéité des tumeurs « p53 négatif / p53 positif » identifiées par immunohistochimie	138
Figure 39 : Régulation du cycle cellulaire par H19	139
Figure 40 : Génération d'un allèle H19 déficient en miR-675	143

Tableaux	Page
Tableau 1 : Classification TNM et stade des tumeurs du sein	15
Tableau 2 : Caractéristiques des différents sous-types moléculaire de cancers du sein	17
Tableau 3 : Caractérisation fonctionnelle des lncRNAs dans le cancer du sein	38
Tableau 4: Cibles des miR-675-5p et miR-675-3p	47
Tableau 5: Liste des microARN cibler par H19	49
Tableau 6: Actions connues de certains mutants gain de fonctions de p53 dans les cancers	136

## Annexe

- Revue de Collette et al. 2017 au format publié dans la revue *International Journal of Molecular Sciences* (Open acces Creative Commons CC BY 4.0 license). Page 155.

## Liste des abréviations

ALDH : Aldehyde dehydrogenase

CCI : Carcinome canalaire invasif

CCIS : Carcinome canalaire *in situ*

CE model : modèle d'évolution clonale

ceRNA : Competing endogenous RNA

CSC : Cellules souches cancéreuses

EMT : Transition épithélio-mésenchymateuse

ENCODE: Encyclopedia of DNA Elements

ER: Estrogen receptor

FXR1 : Fragile X mental retardation syndrome-related protein 1

GOF: Gain-of-function

HER2: Human epidermal growth factor receptor 2

HGP: Human Genome Project

IHC: Immunohistochimie

INCA : Institut National du Cancer

Junk DNA : ADN poubelle

lncRNA: Long ARN non codant

LOF: Loss-of-function

MET : Transition mésenchymateuse-épithéliale

miR: microRNA

PR : Progesterone receptor

PCG : Gènes codants pour des protéines

RISC : RNA-induced silencing complex

ROS : Dérivés réactifs de l'oxygène (Reactive Oxygen Species)

siRNA: Small interfering RNA

TNBC: Cancer du sein triple négatif (Triple Negative Breast Cancer)

TP53 : Tumor Protein P53



# Introduction





# INTRODUCTION

## I. Le cancer du sein

### I.1. Généralités

#### I.1.1. Epidémiologie du cancer du sein

D'après l'étude GLOBOCAN 2018, près de 2,1 millions de nouveaux cas de cancer du sein seront recensés cette année, ce qui constituera environ  $\frac{1}{4}$  des cancers chez la femme. Avec 620.000 décès estimés, le cancer du sein est également la première cause de mortalité par cancer chez la femme, suivi du cancer du poumon (Globocan 2018, Figure1).

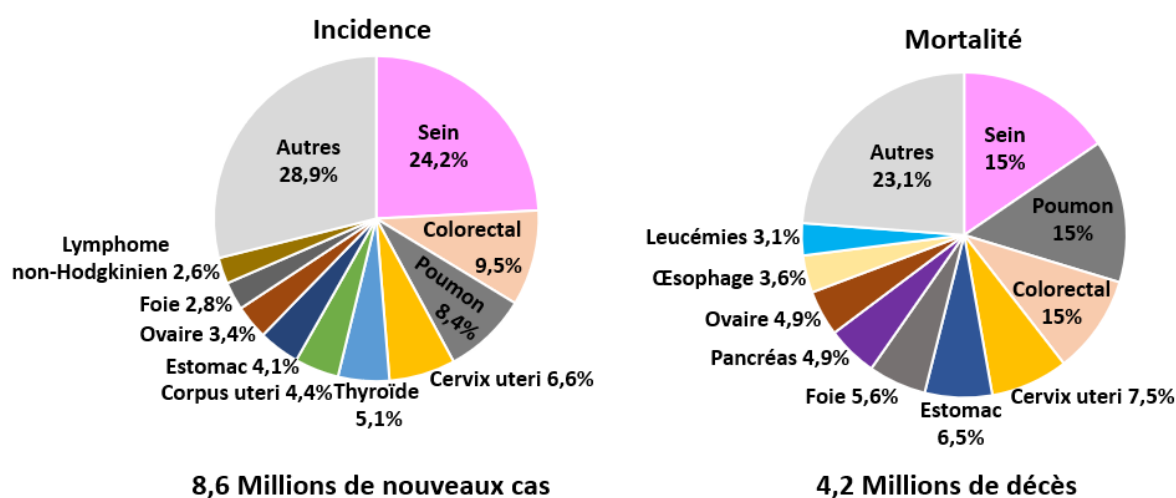


Figure 1 : Représentation de la distribution des nouveaux cas et des décès des 10 cancers les plus communs en 2018.

En France, près de 60 000 nouveaux cas ont été diagnostiqués et environ 12 000 décès sont à déplorer en 2017 selon les chiffres de l'INCA. L'incidence du cancer du sein a augmenté de manière considérable entre 1980 et 2000 mais le nombre de cas dépistés chaque année a tendance à diminuer d'environ 1,5% par an depuis 2005 (Figures 2, données INCA).

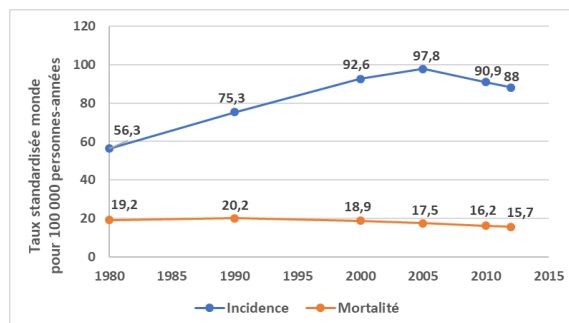


Figure 2 : Evolution de l'incidence et de la mortalité par cancer du sein de 1980 à 2012 en France métropolitaine. Adapté de (Binder-Foucard et al.,2013).

Depuis 2005, nous constatons une diminution du taux de mortalité qui peut s'expliquer par une meilleure prévention de cette maladie associée à plusieurs campagnes de communications et à un dépistage systématique proposée aux femmes dès 50 ans. Cette diminution peut également s'expliquer par l'évolution des thérapies et l'apparition des thérapies personnalisées.

Le cancer du sein est considéré comme un cancer à bon pronostic puisque la survie nette des femmes diagnostiquées entre 1989 et 2010 est de 87% à 5 ans et de 76% à 10 ans (INCA). Plus de la moitié des cancers du sein sont diagnostiqués après 50 ans et l'âge médian au diagnostic est de 63 ans.

### ***1.1.2. Les facteurs de risque***

Un facteur de risque est défini par l'organisation mondiale de la santé (OMS) comme étant « tout attribut, caractéristique ou exposition d'un sujet qui augmente la probabilité de développer une maladie ou de souffrir d'un traumatisme ».

- Certains risques majeurs sont intrinsèques à l'individu et ne peuvent être modulés. Par exemple, le premier facteur de risque est le sexe de l'individu, en effet, 99% des cancers du sein vont toucher la femme. Le vieillissement est également un facteur qui augmente le risque d'apparition d'un cancer du sein (Feng et al.,2018).

- Des altérations génétiques somatiques sont la cause de forme héréditaire de cancers du sein, les plus connus étant des mutations sur les gènes de prédispositions aux cancers du sein : *BRCA1* et *BRCA2*. Environ 5 à 10% des cancers du sein sont liés à une mutation génétique héritée des parents. Une femme qui possède une mutation *BRCA1* a un risque de 55-65% de développer un cancer du sein au cours de sa vie, 45% pour une mutation *BRCA2* et en moyenne une femme portant une de ces mutations aura 70% de chance d'avoir un cancer du sein avant 80 ans (Antoniou et al.,2003). D'autres mutations germinales dans de nombreux autres gènes sont capables d'amener au développement d'une forme familiale de cancer du sein, quelques exemples sont décrits dans les points suivants :

- **ATM (Ataxia telangiectasia mutated)** : Des mutations dans le gène ATM (gène responsable notamment de l'ataxie télangiectasie) favorisent l'apparition d'un cancer du sein. La protéine ATM est une protéine activée à la suite de cassures doubles brins de l'ADN afin d'induire des mécanismes permettant l'arrêt du cycle et la réparation de ces dommages à l'ADN.

- **CHEK2 (Checkpoint kinase 2)** : La présence de mutations germinales dans le gène *CHEK2* a été découverte dans le cas de syndrome Li Fraumeni et Li Fraumeni like en 1999 et associée à un risque plus élevé de développer un cancer du sein (Bell Daphne W. et al., 1999). La protéine CHK2 a un rôle suppresseur de tumeur, elle est activée à la suite de dommage à l'ADN par les protéines ATM et ATR. L'activation de CHK2 via sa phosphorylation va lui permettre d'activer la protéine p53 afin d'induire une signalisation de réparation de l'ADN ou d'apoptose.
- **TP53 (Tumor Protein p53)** : Des mutations somatiques dans le gène *TP53* sont à l'origine du syndrome de Li Fraumeni qui est associé à un risque accru de cancer du sein chez l'individu, mais également d'autres cancers comme des leucémies. Le gène *TP53* est un gène suppresseur de tumeur, la protéine p53 est obtenue à partir de ce gène. Cette protéine possède un rôle très important dans la cellule puisqu'elle va réguler de nombreux processus essentiels au bon fonctionnement cellulaire. La protéine p53 va être activée à la suite de nombreux stimuli (Dommages à l'ADN, stress oxydatifs, hypoxie...etc.) afin d'agir sur différents processus cellulaires pour assurer le bon fonctionnement et l'intégrité de la cellule. Par exemple, lors de dommages à l'ADN, l'activation de la protéine p53 va induire un arrêt dans le cycle cellulaire afin de pouvoir procéder à la réparation de l'ADN et de conserver l'intégrité de la cellule. Si les dommages sont trop importants, la protéine p53 peut induire l'activation de l'apoptose afin d'éliminer une cellule déficiente.
- **PALB2 (Partner and localizer of BRCA2)** : PALB2 est une protéine qui interagit avec BRCA2 et permet l'action nucléaire de BRCA2. Les individus porteurs de mutations dans le gène *PALB2* ont un risque plus important de développer un cancer du sein. ( Rahman et al.,2007)
- **PTEN (Phosphatase and tensin homolog)** : Des mutations héréditaires dans le gène *PTEN* sont responsables du syndrome de Cowden qui s'accompagne par un risque accru de développement de cancer du sein. La protéine PTEN agit notamment comme un antagoniste de la voie de signalisation PI3-K (Phosphatidylinositol-3-phosphate kinase) qui joue un rôle dans différents processus tels que le contrôle du cycle cellulaire, l'apoptose ou la migration cellulaire.

**La plupart des mutations germinales qui causent des cancers du sein héréditaires sont portées par des gènes régulant les voies de réparation de l'ADN.**

- De nombreux facteurs de risques sont cependant modulables au cours de la vie d'un individu, afin d'augmenter ou de diminuer le risque de cancer du sein.

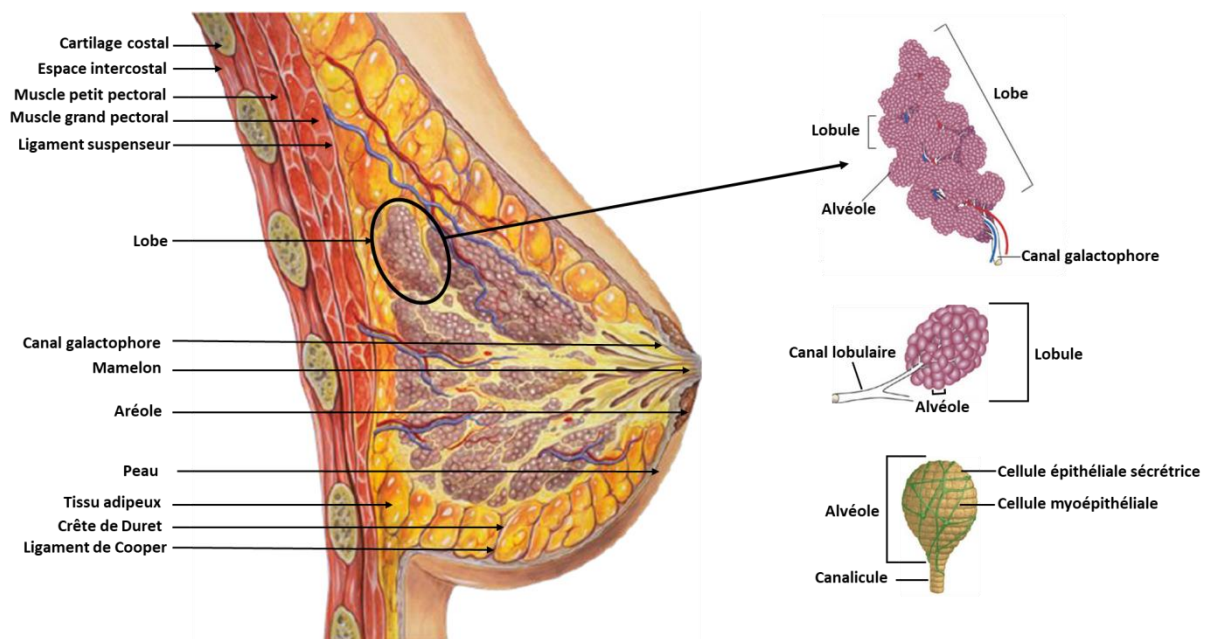
- **Expositions à des toxiques** : De nombreuses études ont démontré que la consommation excessive d'alcool augmente le risque de développement du cancer du sein. Par exemple, une femme qui boit 2 à 3 verres d'alcool par jour, a environ 20% de risques supplémentaires de développer un cancer du sein qu'une femme qui ne boit pas (données INCA). Des études récentes ont également démontré que les femmes exposées à du tabagisme actif ou passif ont un risque plus élevé que les femmes jamais exposées au tabac (Inserm, rapport de l'Agence française de sécurité sanitaire de l'environnement et du travail ; afsset).
- **Surpoids, obésité et activité physique** : Le surpoids et l'obésité augmentent le risque de cancer du sein chez la femme, surtout après la ménopause. Être en surpoids augmenterait également le risque de récurrence du cancer. Cependant, le lien entre surpoids et cancer du sein est compliqué et affecté par d'autres facteurs comme par exemple la localisation du surplus graisseux. A contrario, une activité physique régulière est associée à une diminution du risque de cancer du sein après ménopause.
- **Consommation hormonale au cours de la vie** : La prise de contraceptifs hormonaux augmente le risque de cancer du sein. Les femmes qui ont pris un contraceptif oral ont un risque plus élevé que celles qui n'en ont jamais utilisé. Une thérapie hormonale peut être proposée à une femme pour soulager les symptômes de la ménopause. Cependant, des études ont démontré que certaines thérapies hormonales augmentent le risque du cancer du sein.
- **Procréation** : Les femmes qui n'ont pas d'enfants ou qui ont eu un enfant après 30 ans ont un risque plus important de développer un cancer du sein. A contrario, avoir plusieurs enfants ou être mère plus jeune diminue ce risque. Il a également été suggéré que l'allaitement pourrait diminuer le risque de cancer du sein, en particulier s'il a été prolongé jusque 1,5 à 2 ans. Ce qui s'expliquerait par le fait que l'allaitement diminue le nombre total de cycle menstruel.

## I.2. Histoire naturelle du cancer du sein

### I.2.1. Développement du cancer du sein

#### I.2.1.1. Anatomie de la glande mammaire normale

Le sein est une glande exocrine qui est destinée à assurer la sécrétion du lait chez la femme après l'accouchement. Elle reste immature et non fonctionnelle chez l'homme. La glande mammaire est entourée d'un tissu adipeux reposant sur les muscles pectoraux grâce à des crêtes fibro-glandulaires de Duret, des ligaments suspenseurs et des ligaments de Cooper (Figure 3).



**Figure 3 : Anatomie de la glande mammaire**

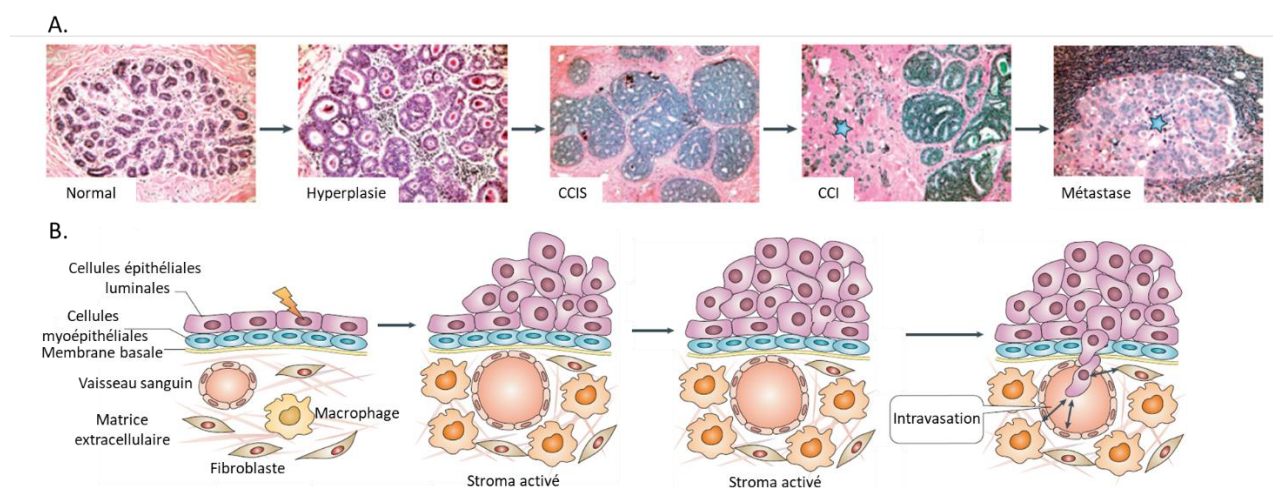
Représentation schématisique d'une coupe sagittale de sein et de la paroi thoracique antérieure. La glande mammaire est formée de tissu fibro-glandulaire et de tissu adipeux qui repose sur les muscles pectoraux. Elle est constituée d'une dizaine de canaux galactophores aboutissant dans 10 à 20 lobes et permet la sécrétion de lait au niveau du mamelon. Chaque lobe est composé de 20 à 40 lobules, eux-mêmes sous-divisés en 10 à 100 alvéoles. L'intégrité de la glande mammaire est préservée grâce aux crêtes de Duret, liées au derme via les ligaments de Cooper. Adapté de « Medical Illustrations by Patrick J. Lynch » et d'illustrations de <http://www.santeallaitemental.com>.

L'aréole est la zone de peau qui entoure le mamelon, constituée d'un épiderme fin et pigmenté. Cet endroit contient des glandes sébacées, appelées glandes de Morgagni, qui s'hypertrophient lors de la grossesse et prennent alors le nom de tubercules de Montgomery. La glande mammaire est constituée d'une vingtaine de lobes irréguliers qui sont enfouis dans le tissu graisseux et séparés par une couche de tissu conjonctif interlobulaire dense. Chaque lobe est sous-divisé en 20 à 40 lobules, qui sont également séparés entre-deux par du tissu conjonctif interlobulaire et divisés en 10 à 100 alvéoles (ou acini). Les canaux galactophores se

ramifient progressivement en canaux de plus en plus étroits, jusqu'à l'unité terminale ducto-lobulaire (UTDL). L'UTDL est formée par les canaux extra et intra lobulaires qui se terminent par les alvéoles qui assurent la production de lait lors de la grossesse et la lactation. Un canal galactophore se ramifie en canaux lobulaires qui aboutissent dans les lobules puis en canicule au niveau des acini. Les UTDL sont formées de deux principaux types cellulaires : luminal et basal. La couche basale est formée par les cellules myoépithéliales qui sont des cellules contractiles permettant l'éjection du lait et de cellules souches assurant le renouvellement des différents types cellulaires. La partie luminale est composée de cellules épithéliales sensibles aux hormones et sécrétrices de lait. L'ensemble se situe dans un stroma composé de matrice extracellulaire, de vaisseaux sanguins et de différents types cellulaires (fibroblastes, adipocytes, macrophages).

### 1.2.1.2. Différentes étapes de cancérisation du sein

Le développement type d'un cancer du sein s'effectue en plusieurs étapes (Figure 4).



**Figure 4 : Etapes de la cancérisation d'un épithélium canalaire mammaire**

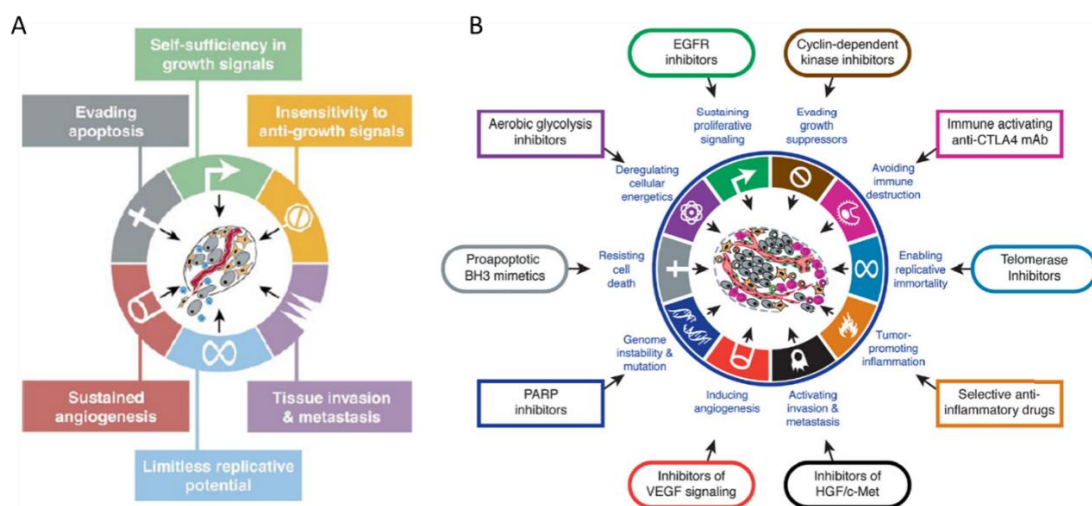
Les cellules d'un épithélium normal peuvent subir des altérations et proliférer de manière anarchique jusqu'au stade d'hyperplasie atypique. Les cellules cancéreuses vont se différencier et proliférer de manière locale pour donner un carcinome in situ (CCIS). Les modifications phénotypiques des cellules vont leur permettre de traverser la membrane basale, on parle de carcinome canalaire invasif (CCI). Une fois que les cellules ont franchies la membrane basale, elles vont être capable de pénétrer les vaisseaux sanguins (intravasation) et de se disséminer dans l'organisme pour former des métastases. A. Coupes histologiques d'un tissu mammaire normale, d'un CCIS, d'un CCI et d'une métastase dans un nœud lymphatique. B. Représentation schématique du développement d'une tumeur mammaire (Vargo-Gogola & Rosen, 2007).

La première étape correspond au développement d'une hyperplasie avec l'apparition d'un épaississement de l'épithélium canalaire, une obstruction partielle du canal galactophore peut être observée. Cet épaississement est dû à la prolifération accrue des cellules épithéliales canalaire basales ou luminales. Dans le cas d'une prolifération accrue uniquement des cellules

luminales, on parle d'un stade hyperplasique atypique. Celle-ci augmente de manière significative les risques de cancer du sein, contrairement au stade d'hyperplasie classique. Les cellules prolifératives peuvent acquérir des altérations géniques, épigénétiques et architecturales qui conduisent à la formation de lésions pré-cancéreuses de type « carcinome canalaire *in situ* (CCIS) ». Au stade CCIS, les cellules cancéreuses restent au même endroit et la membrane basale est intacte. A la suite d'accumulation d'anomalies génétiques et de transformations oncogéniques, les CCIS peuvent devenir des carcinomes canauxaires invasifs (CCI). A ce stade, l'intégrité de la membrane basale est rompue et les cellules cancéreuses se répandent et envahissent le tissu conjonctif adjacent. Les cellules cancéreuses ayant acquis un potentiel de migration accru vont être capable de pénétrer les vaisseaux sanguins proches (intravasation) et de se disséminer dans l'organisme, jusqu'à atteindre un site de dissémination primaire.

### 1.2.1.3. Hallmarks of cancer : Les éléments clés de la cancérisation

Les événements induisant la transformation d'une cellule saine en cellule cancéreuse ainsi que la dissémination des métastases sont complexes et multiples. En 2000, Douglas Hanahan et Robert A. Weinberg ont publié une revue qui fait date dans le domaine de la cancérologie (Hanahan & Weinberg, 2000). Dans cette revue, les auteurs ont recensé six modifications physiologiques dont l'acquisition est essentielle au développement et à la propagation des cancers (Figure 5.A).



**Figure 5 : Représentation schématique des Hallmarks du cancer**

A. Représentation originale des 6 premières hallmarks de 2000. B. Représentation des 10 hallmarks et des pistes thérapeutiques pour les contrer (Hanahan & Weinberg, 2011).

- Autosuffisance vis-à-vis des signaux de croissance
- Perte de sensibilité aux signaux antiprolifératifs
- Capacité d'invasion tissulaire et de former des métastases
- Potentiel répliatif illimité
- Capacité à effectuer de la néo-angiogenèse
- Échappement à l'apoptose

Plus récemment, quatre nouveaux facteurs ont été ajoutés à cette liste (Figure 5.B).

- Échappement à la destruction par le système immunitaire
- Production d'inflammation tumorale
- Instabilité génomique conduisant à l'accumulation de mutations
- Dérégulation du métabolisme cellulaire au profit d'une prolifération accrue

L'ensemble de ces processus est acquis par les cellules tumorales au cours du processus de cancérisation. Chacun de ces aspects peut-être une cible thérapeutique intéressante (Figure 6.B). Toutes ces modifications sont directement liées à l'apparition d'altérations génétiques et épigénétiques. Cependant, le microenvironnement tumoral ainsi que le système immunitaire jouent également un rôle important.

## ***1.2.2. Origine du cancer : Importance des cellules souches cancéreuses***

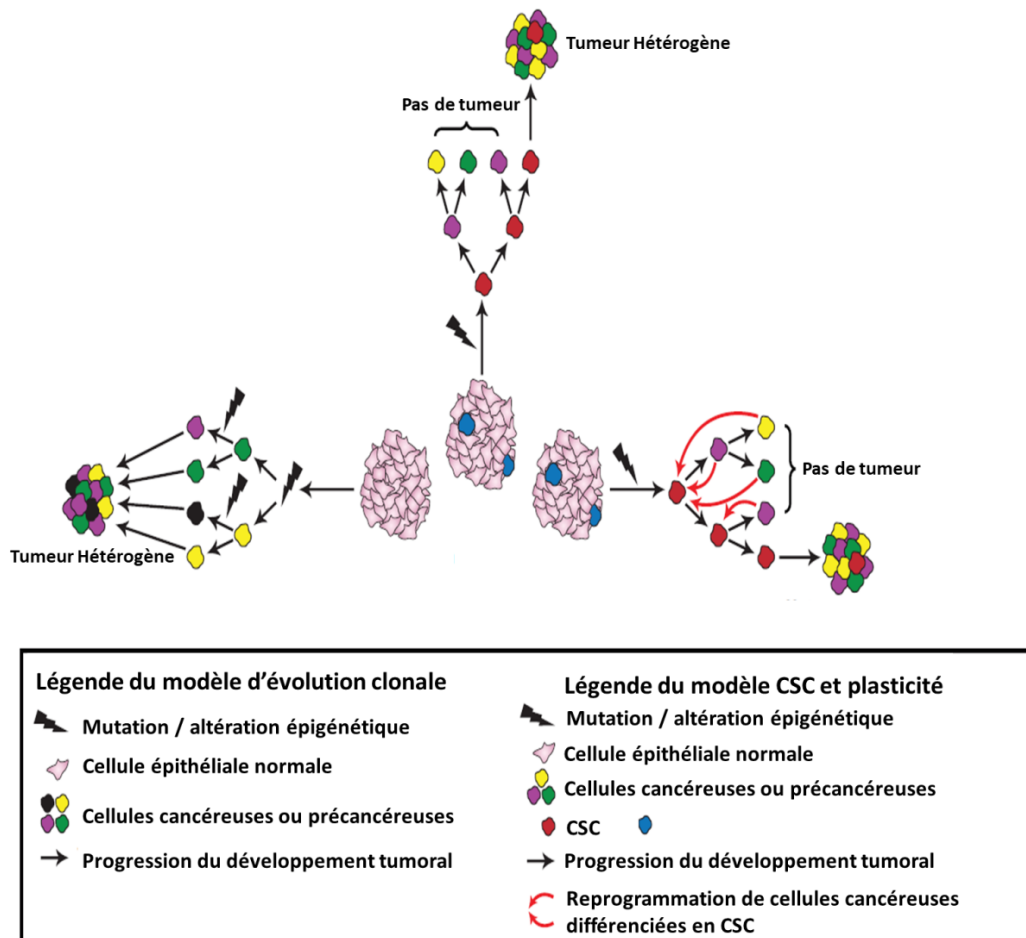
### *1.2.2.1. Modèles de développement tumoral*

La compréhension de l'hétérogénéité tumorale est un élément majeur dans la recherche de pistes thérapeutiques pour les cancers et en particulier pour le cancer du sein. Il existe une hétérogénéité intertumorale observée entre des cancers du sein issue de plusieurs individus. Mais il existe également une hétérogénéité intratumorale due à la présence de différentes populations cellulaires à l'intérieure d'une même tumeur. Ces sous-populations de cellules cancéreuses ont des spécificités en termes de prolifération cellulaire, de résistance aux traitements et de pouvoir métastatique. Plusieurs modèles d'évolution des cancers ont été développés au fur et à mesure des recherches dans le domaine (Figure 6).

Selon le **modèle d'évolution clonale**, toutes les cellules d'une tumeur sont capables de participer au développement tumoral. L'accumulation de mutations génétiques et de modifications épigénétiques entraînent l'apparition de clones distincts dans la tumeur. Ces changements permettraient la création de sous-populations possédant des capacités



tumorigènes plus importante comme une prolifération accrue, une meilleure capacité d'invasion ou une résistance à l'apoptose par exemple. Dans le cadre du cancer du sein, l'existence de différents clones au sein d'une même tumeur est bien établie (Torres et al.,2007, Nakada et al.,2017). Lorsqu'une tumeur est traitée par une thérapie, seules les clones possédant l'outillage cellulaire nécessaire à la survie survivront et pourront redonner une tumeur sur le long terme. Selon ce modèle, il est impossible de déterminer la cellule d'origine du cancer puisque toutes les cellules du tissu sont potentiellement capables de participer au développement tumoral et de former une tumeur.



**Figure 6 : Représentation des modèles de développement tumoral**

Dans le modèle stochastique (CE), l'accumulation de mutations au cours du temps dans n'importe quelle cellule peut conférer un pouvoir tumorigène. Dans le modèle hiérarchique (CSC), seule les cellules souches possèdent un potentiel tumorigène. Le modèle de la plasticité cellulaire suggère que les cellules cancéreuses différenciées peuvent se transformer en CSC, qui elles auront un pouvoir tumorigène. CE = évolution clonale. CSC = cellules souches cancéreuses. Adapté de Rich, 2016.

Contrairement au modèle d'évolution clonale, le **modèle hiérarchique** suppose que seule une sous-population rare possède la capacité de former une tumeur. Cette sous-population appelée « Cellules souches tumorales (CSC) » dériverait des cellules souches

normales et posséderait la capacité d'auto-renouvellement et de différenciation. Les CSC seraient donc à l'origine de l'initiation et du développement tumoral. L'hétérogénéité tumorale s'expliquerait par la capacité de différenciation des CSC en cellules progénitrices puis en cellules différenciées. Expérimentalement, l'existence des cellules souches a été démontré dans de nombreux cancers (Batlle & Clevers, 2017). Contrairement au modèle stochastique, ce modèle permettrait l'identification de la sous-population de cellules cancéreuses qui possède le pouvoir tumorigène.

Ces deux modèles ne sont cependant pas exclusifs et un nouveau modèle a été développé afin de les unifier. Le **modèle de plasticité cellulaire** postule que des cellules différenciées peuvent subir un phénomène de plasticité cellulaire afin de se dédifférencier et devenir des CSC. Différents stimuli (altérations épigénétiques, mutations géniques, thérapies, stimulation par le microenvironnement) seraient capables d'induire une reprogrammation de cellules différenciées en cellules souches. Par ailleurs, il est également possible d'obtenir différents clones de CSC à la suite de l'accumulation d'évènements génétiques et épigénétiques par exemple. Ces différentes sous-populations auraient des propriétés différentes, mais seraient toutes capables de reformer une tumeur.

#### *1.2.2.2. Les cellules souches cancéreuses mammaires*

##### *1.2.2.2.1. Généralités sur les cellules souches cancéreuses*

Le terme de cellules souches cancéreuses désigne une sous-population cellulaire qui possède des propriétés d'auto-renouvellement illimité et la capacité de donner naissance à des progéniteurs et des cellules plus différenciées ayant des pouvoirs tumorigènes limités (Kreso & Dick, 2014). Les CSC ont été initialement découvertes par l'équipe du Pr. John Dick en 1997 dans les cancers hématopoïétiques (Bhatia et al.,1997). La première preuve de l'existence de CSC dans des tumeurs solides a été apportée par l'équipe de Clarke dans le cadre du cancer du sein (Al-Hajj et al.,2003). Sur la base d'un marquage  $CD44^+/CD24^{-/low}/Lin^{-}$ , l'équipe a identifié une sous-population cellulaire capable de produire des tumeurs en souris SCID immunodéficientes lors d'expériences de xénogreffes à partir de seulement 100 cellules contre au moins 10.000 pour des cellules non triées.

L'origine des CSC est encore débattue, cependant une hypothèse vraisemblable serait que les CSC dérivent des cellules souches normales. Ces cellules possèdent en effet une durée de vie dans les tissus plus grande que les cellules différenciées, elles seraient ainsi exposées

de manière importante aux risques d'accumulations de mutations ce qui pourrait aboutir à leurs transformations (Sugihara & Saya, 2013). Elles sont également moins sensibles aux thérapies conventionnelles et pourraient être à l'origine des résistances aux traitements et des récurrences.

#### *1.2.2.2.2. Résistances aux thérapies des CSC et reprogrammation*

Dans la littérature il est admis que les CSC expliquent une partie de la résistance aux thérapies anticancéreuses. Les CSC seraient donc également responsables des phénomènes de récurrences puisque ces cellules sont capables de survivre aux thérapies anticancéreuses et redonner une hétérogénéité tumorale. La résistance aux thérapies induite par les CSC s'explique notamment par plusieurs caractéristiques de ces cellules :

- **Surexpression de transporteurs d'efflux** : De nombreux transporteurs d'efflux sont surexprimés par les CSC (ex : famille des Multi Drug Resistance, MDR). Les CSC vont donc pouvoir éjecter les drogues anticancéreuses et ainsi être moins sensibles aux traitements.
- **Diminution du stress oxydant** : De nombreuses thérapies vont agir en induisant un stress oxydant via la production d'espèces réactives de l'oxygène (ROS) dans les cellules. Ces ROS vont induire un stress oxydant qui provoque des dommages à l'ADN et amène à la mort cellulaire. Certaines enzymes de détoxification sont surexprimées afin de diminuer ce stress oxydant, comme par exemple certaines Aldéhyde déshydrogénases (ALDH).
- **Quiescence** : La plupart des thérapies anticancéreuses ciblent les cellules en prolifération et nécessitent l'activation du cycle cellulaire pour induire l'apoptose. Les CSC étant majoritairement en quiescence, elles seront protégées contre ces thérapies.

#### *1.2.2.2.3. Méthodes d'identifications et d'études des CSC mammaires*

Plusieurs méthodes d'identifications et d'études des CSC mammaires ont été décrites dans la littérature, cependant, un marqueur donné n'est pas valable pour toutes les lignées cellulaires. Il est également intéressant de constater que tous les marqueurs de CSC décrits ne vont pas permettre d'isoler une seule sous-population de CSC identiques. Plusieurs marqueurs de CSC vont permettre d'isoler différentes populations de CSC distinctes.

- **Marqueurs de surface** : Les marqueurs de surface sont les premiers outils utilisés pour identifier les CSC (Bhatia et al., 1997). Dans le cancer du sein, le phénotype CD44<sup>+</sup>/CD24<sup>-/low</sup> (expression de la protéine de surface CD44 et absence ou très faible expression de la

protéine membranaire CD24) permet l'identification d'une population de CSC capable de former des tumeurs hétérogènes *in vivo*. La proportion de cellules CD44<sup>+</sup>/CD24<sup>-/low</sup> est très variable, selon la lignée étudiée : 100% dans les SUM159PT triple-négative et 0% dans les MCF-7 luminales par exemple.

- **Activité enzymatique de l'ALDH** : L'équipe de Ginestier et collaborateurs ont démontré que l'utilisation de l'activité enzymatique de l'ALDH (Aldéhyde déshydrogénase) permet d'isoler une population enrichie en CSC (Ginestier et al., 2007). Il est possible d'identifier une sous-population enrichie en CSC par cytométrie en flux en détectant l'activité ALDH grâce à l'apport d'un substrat qui devient fluorescent lorsqu'il est clivé par l'ALDH.

- **Exclusion de colorants vitaux - Side population** : Les CSC possèdent une capacité à exclure les chimiothérapies plus importantes que les cellules cancéreuses non souches grâce à une activité élevée des pompes d'efflux de la famille ABC (ATP binding cassettes ; (Zhou et al., 2001, Wu & Alman, 2008)). Il est possible d'identifier par cytométrie en flux l'activité de ces pompes d'efflux par un test d'exclusion de colorants vitaux fluorescents tels que le Hoechst 33342.

- **Capacité de formation de sphères** : Des tests fonctionnels ont été mis au point afin de définir le phénotype souche d'une population cellulaire. Le test de formation de mammosphères repose sur la capacité des cellules souches à croître en l'absence d'ancrage. Les cellules sont mises en culture à faible densité dans un milieu particulier (dépourvu de sérum et riche en facteurs de croissances) et en condition non adhérentes. La capacité des cellules à former des mammosphères est évaluée en comptant le nombre de sphères et en rapportant ce chiffre au nombre de cellules ensemencées.

- **Tumorigénicité *in vivo*** : Le test fonctionnel considéré comme le « gold standard » est le test de tumorigénicité *in vivo*. Les CSC potentielles sont injectées en nombre limité chez la souris, jusqu'à 100 cellules contre plus d'un million en condition classique de xénogreffe. La présence ou l'absence de développement tumoral définit le caractère souche de la population testée. Un autre paramètre utilisé pour définir le caractère souche est la capacité des cellules à former des tumeurs en série et à reproduire l'hétérogénéité de la tumeur initiale (Visvader & Lindeman, 2008).

### I.3. Classifications des cancers du sein

Le cancer du sein est une pathologie très hétérogène, chaque patiente possède un cancer du sein différent. Il est cependant possible de réaliser plusieurs classifications des cancers du sein en fonction de différents paramètres cliniques, anatomopathologiques ou encore aux profils d'expressions des tumeurs. Ces classifications permettent de prédire la réponse des patientes aux thérapies et de définir le meilleur traitement pour un individu donné.

#### I.3.1. Classification clinique et grade des tumeurs

La **classification TNM (Tumor / Nodes / Metastasis)** a été développé par le Pr. Pierre Denoix entre 1943 et 1952 puis adoptée en 1953 pour son utilisation dans le cancer du sein. Cette classification repose sur l'analyse de plusieurs critères anatomiques de la tumeur afin de déterminer l'avancement de la maladie, le pronostic et les thérapies adéquates à mettre en place. Les paramètres des mesures sont détaillés dans le Tableau 1.

<b>Tumeur primaire (T)</b>		
Tis	Cancer in situ: tumeur intra-épithéliale sans invasion de la lame basale	
TX	La tumeur primaire ne peut pas être analysée	
T1	Invasion par la tumeur de la lame basale ou de la sous-muqueuse	
T2	Invasion par la tumeur de la musculaire propre	
T3	Invasion par la tumeur de la séruse (péritoine viscéral) sans invasion des structures adjacentes	
T4	La tumeur envahit les structures adjacentes	
<b>Ganglions lymphatiques (N)</b>		
NX	Les ganglions lymphatiques ne peuvent être analysés	
N0	Aucune métastase ganglionnaire	
N1	1-6 ganglions métastatiques	
N2	7-15 ganglions métastatiques	
N3	>15 ganglions métastatiques	
<b>Métastases distantes (M)</b>		
MX	Les métastases ne peuvent être déterminées	
M0	Pas de métastase à distance	
M1	Métastases à distance	

Classification en Stade	Classification TNM	Survie à 5 ans (%) Cohorte 1944-1999
Stade 0	Tis, N0, M0	100
Stade I	T1, N0, M0	98
Stade II:		80
.II A	T2, N0, M0 T3, N0, M0	
.II B	T1, N1, M0 T2, N1, M0	
Stade III	T3, N1, M0 T4, Tous N, M0	70
Stade IV	Tous T, Tous N, M1	20

**Tableau 1 : Classification TNM et stade des tumeurs du sein**

Concrètement, le premier élément considéré est la taille de la tumeur qui est mesurée selon un score T (entre 0 et 4). Une valeur élevée correspond à un volume tumoral plus important. Le second paramètre correspond à l'envahissement ganglionnaire qui est mesuré selon un score N (entre 0 et 3). Le score M permet de juger de la présence ou l'absence de métastases (0 ou 1). A partir de ces 3 scores, un stade d'avancement de la maladie est exprimé entre 0 et IV.

Il est également possible de classer les tumeurs du sein selon le grade de **Scarff, Bloom & Richardson (SBR)**. Les tumeurs sont classées selon trois grades en fonction du score de plusieurs paramètres. Le premier critère correspond à l'architecture tumorale qui prend notamment en compte l'état de différenciation des cellules. Les atypies nucléaires telles que la taille et la forme des noyaux sont pris en compte comme deuxième critère. Le dernier critère correspond à l'index mitotique, dont le score augmente avec le nombre de mitoses mesurées. Chaque critère est noté de 1 à 3 et le score total permet de déterminer le grade : un score entre 3 et 5 correspond à une tumeur de grade I ; un score de 6 à 7 définit une tumeur de grade II et les scores de 8 à 9 sont associés à des tumeurs de haut grade peu différenciées.

### 1.3.2. Classification moléculaire

Le développement et la démocratisation des outils de biologie moléculaire permettant l'étude du génome et du transcriptome a permis de déterminer une nouvelle classification des cancers du sein basée sur les profils d'expressions géniques des tumeurs. Des travaux précurseurs de l'équipe de Perou et collaborateurs publiés au début des années 2000 ont permis la classification des cancers en fonction de l'expression génique des tumeurs (Perou et al., 2000, Sørlie et al., 2003). Les auteurs ont identifié que les tumeurs ER+ et ER- peuvent être distinguées selon leur profil d'expression génique. Au moins 5 sous-types moléculaires importants : luminal A et B, HER2+ (*Human epidermal growth factor receptor 2*), basal-like et normal-like avec des caractéristiques moléculaires et cliniques distinctes ont été déterminés (Figure 7). La classification de ces 5 sous-groupes repose principalement sur l'expression de récepteurs : le corécepteur à l'EGF (HER2), le récepteur aux œstrogènes (ER) et le récepteur à la progestérone (PR). Le pronostic des patientes va être différent en fonction du sous-groupe déterminé par leur profil d'expression (Tableau 2).

Sous-type	Status ER/PR/HER	Fréquence	Autres marqueurs	gènes caractéristiques	Grade histologique	Pronostic
luminal-A	ER+ PR+ HER2-	50-60%	CK8/18+	<i>LJVI, FOXA1, XBPI, GATA3, BLC2, erBb3, erbB4</i>	I ou II	Bon
luminal-B	ER+/- PR+/- HER2+/-	15-20%	Ki67 -	<i>v-MYB, GGH, LAPTM4, NSEPI, CCNE1</i>	II ou III	Intermédiaire
HER2 positive	ER- PR- HER2+	15-20%	Ki67+; <i>TP53</i> muté	<i>EGFR, P-cadherine</i>	III	Défavorable
basal-like	ER- PR- HER2-	8-37%	CK5/14/17+; lamine +; <i>TP53</i> muté, Ki67+	<i>P-cadherine, calvéoline, EGFR</i>	III	Très défavorable
normal-like	ER-/+ PR? HER2-	5-10%	CK5-; EGFR-	nd	I	Intermédiaire
moleculair apocrine	ER- PR- HER2+/-	nd	CK5+	<i>AR, FAS, erbB2, XBPI</i>	II ou III	Intermédiaire
claudin low	ER- PR- HER2-	12-14%	CK5, CDH1 bas/-; CLND1 bas/-	<i>CD44, SNAI3</i>	III	Défavorable

**Tableau 2 : Caractéristiques des différents sous-types moléculaire de cancers du sein**

D'autres sous-types ont également été mis en évidence plus récemment : les tumeurs « molecular apocrine » qui possèdent une activation de la signalisation du récepteur aux androgènes (Farmer et al.,2005) et les tumeurs « claudin-low » qui se caractérisent par une faible expression de marqueurs de jonction cellulaire (Prat et al.,2010). Les caractéristiques de ces sous-types sont décrites dans le Tableau 2. Les tumeurs « basal-like », « molecular apocrine » et « claudin-low » sont regroupées sous le terme de tumeurs triple négatives (TNBC, Triple Negative Breast Cancer). Ces tumeurs sont les plus agressives et possèdent un pronostic très défavorable.

Cette classification moléculaire permet de mieux prédire le pronostic et de guider le schéma thérapeutique des patients.

## **I.4. Stratégies thérapeutiques**

Le cancer du sein est considéré comme un cancer à bon pronostic puisque la survie nette des femmes diagnostiquées entre 1989 et 2010 est de 87% à 5 ans et de 76% à 10 ans (INCA). De nombreuses stratégies thérapeutiques sont utilisées pour traiter les patients atteints de cancers du sein. Le schéma thérapeutique d'une patiente est décidé par l'oncologue en fonction des différents critères que nous venons de voir. Lorsqu'un cancer du sein doit être traité, une chirurgie est essentielle pour le traitement de la pathologie. Chez les patientes possédant une tumeur importante un traitement néoadjuvant est généralement prescrit (Chimiothérapies et/ou thérapie ciblée anti-HER2). Dans le cas d'une petite tumeur, une tumorectomie est préférable. Cette opération consiste à enlever le tissu tumoral mais également une partie du tissu sain adjacent. Ce type de chirurgie permet d'éviter l'ablation complète du sein. L'autre possibilité est le recours à la mastectomie qui correspond à l'ablation complète du sein et qui est nécessaire lorsque les tumeurs sont de tailles importantes. A la suite de la chirurgie, plusieurs types de thérapies sont possibles.

### ***I.4.1. Les chimiothérapies***

La chimiothérapie est un traitement adjuvant (postopératoire) ou néoadjuvant (préopératoire) qui a pour but d'éliminer les cellules cancéreuses en ciblant les mécanismes de réplication de l'ADN et de la division cellulaire. Dans le cancer du sein, 4 grandes familles d'agents sont utilisées : les antimétabolites (5-Fluorouracil, méthotrexate), les agents alkylants (cyclophosphamide), les taxanes (paclitaxel, docetaxel) et les anthracyclines

(doxorubicine, épirubicine). Plusieurs agents de chimiothérapies sont souvent prescrits ensemble et en combinaison avec de la radiothérapie, des thérapies ciblées ou des hormonothérapies.

### ***1.4.2. La radiothérapie***

La radiothérapie est un traitement local qui permet d'arrêter la prolifération cellulaire et d'induire la mort cellulaire dans les cellules de la zone ciblée. Les cellules cancéreuses caractérisées par leur prolifération accrue sont donc particulièrement sensibles à la radiothérapie. Les effets de l'exposition aux rayons ionisants d'une zone d'intérêt vont dépendre de la dose reçue et de la localisation de la source d'irradiation. L'effet direct des radiations est l'excitation de l'ADN qui induit d'importants dommages à l'ADN et peut amener à l'induction de l'apoptose. Des effets indirects participent également à l'efficacité de la mort cellulaire, comme l'apparition d'espèces réactives de l'oxygène (ROS). Ces ROS sont à l'origine d'un stress oxydant qui impacte la biologie de la cellule et peut également conduire à l'apoptose.

### ***1.4.3. Les thérapies personnalisées***

La classification moléculaire des cancers du sein a permis l'utilisation de la médecine personnalisée dans cette pathologie. Les patientes possédant des caractéristiques spécifiques à un sous-groupe donné peuvent être traitées avec une stratégie thérapeutique optimale donnée.

Par exemple, une **hormonothérapie** ne sera prescrite que chez des patientes dont la tumeur est hormonosensible, c'est-à-dire lorsque la tumeur exprime le récepteur aux œstrogènes (ER) et à la progestérone (PR). Il existe notamment deux classes d'hormonothérapie ciblant les récepteurs aux œstrogènes. Les SERM (Selective Estrogen Receptor Modular) qui sont des antagonistes du récepteur ER qui se lient de manière compétitive au récepteur ER afin de diminuer la transcription des gènes régulés par les œstrogènes, le plus connu étant le tamoxifène. Les SERD (Selective Estrogen Receptor Degradation) qui sont une autre classe d'antagonistes qui permettent de diminuer l'expression des récepteurs ER.

Des **thérapies ciblées** sont également utilisées dans le cancer du sein et seront également spécifiques au contexte d'expression génique de la patiente. Par exemple, pour des



tumeurs qui présentent une amplification du gène codant le corécepteur HER2, il est possible d'utiliser un anticorps bloquant spécifique. Cet anticorps monoclonal humanisé est le trastuzumab (Herceptin®). En se fixant sur HER2, il empêche la fixation de l'EGF et l'induction de sa voie de signalisation impliquée dans la prolifération cellulaire.

#### ***1.4.4 Résistances aux traitements***

Certaines tumeurs ne vont pas répondre aux traitements et continuer de proliférer malgré la présence d'une thérapie ciblée. Par exemple, plusieurs mécanismes de résistance à l'Herceptin® ont été décrits (Pohlmann et al.,2009). Il peut y avoir une suractivation d'une voie de signalisation en aval de notre cible, comme par exemple la suractivation de la voie PI3K/Akt due à la perte de fonction de *PTEN*. Cette suractivation de la voie PI3K/Akt dans des cellules surexprimant HER2 va rendre inefficace les thérapies comme l'hérceptine qui ciblent HER2. L'existence des cellules souches cancéreuses, qui subissent peu l'effet des chimiothérapies, peut également expliquer, en partie, les phénomènes de résistances aux traitements et de récurrences des cancers du sein.



## II. Les ARN non codants

Trente ans après la caractérisation du premier long ARN non codant (lncRNA), *H19*, chez l'homme, l'importance des lncRNA dans tous les aspects de la vie cellulaire n'est plus à démontrer. De nombreuses classes d'ARN non codant ont été mises à jour ces dernières décennies notamment grâce aux avancées dans les technologies de séquences hauts débits et les outils de biologie moléculaire.

### II.1. De la découverte de l'ADN au rôle primordial des ARNs

Depuis le XIX<sup>ème</sup> (et jusqu'à une bonne partie du XX<sup>ème</sup> siècle), les protéines étaient considérées comme le composant cellulaire le plus important. Dans les années 1860, un jeune étudiant suisse en médecine nommé Friedrich Miescher voulait découvrir les principes fondamentaux de la vie cellulaire en étudiant un modèle cellulaire considéré comme le « plus simple et le plus indépendant » : les lymphocytes. En 1869, lorsqu'il essayait de développer une technique d'isolation des noyaux, il découvrit une substance riche en phosphore qui pour lui n'était pas protéique. Il fut le premier à décrire cette substance acide qui provenait du noyau et qu'il nomma « nucléine ». Il a même réussi à déterminer que cette substance est composée à 22,5% de P<sub>2</sub>O<sub>5</sub> de sa masse totale, ce qui est très proche des 22,9% admis actuellement, et a indiqué que ce phosphore est présent sous la forme d'acide phosphorique. Cette découverte fut réalisée 3 ans après les travaux de Mendel (1866) sur les lois de l'hérédité, mais à l'époque la communauté scientifique pensait que les protéines étaient les messagers de l'information génétique. Au cours des 20 années suivantes, Walther Flemming donna le nom de « chromatine » (traduction du grec « couleur »), au composant coloré dans le noyau à la suite de différents colorants chimiques. Bien que ce soit Wilhelm Hofmeister qui observa en premier les chromosomes au cours d'une division cellulaire en 1848, c'est en 1888 que Wilhelm Waldeyer utilisa le terme de « chromosome » (traduction du grec « corps colorés »), pour désigner le matériel filaire retrouvé dans le noyau qui pouvait se colorer. Les scientifiques ne savaient pas si la chromatine et la nucléine sont en réalité la même chose. En 1889, Richard Altman a réussi à obtenir un échantillon de nucléine sans contamination protéique et proposa le terme d'acide nucléique à la place de nucléine. Au fil du temps, les scientifiques se sont rendu compte qu'il y avait différents types d'acides nucléiques et différents types d'ARN, il n'y a donc pas de personnes ni de date précise pour la découverte de l'ARN. Dans les années 1930, les composants moléculaires de la nucléine ont été déduits et la nucléine est devenue

*desoxyribose nucleic acid puis deoxyribose nucleic acid (DNA)*. Pendant plusieurs décennies, d'importantes découvertes sur la nature de l'ADN ont été réalisées avec notamment le modèle hélicoïdal incorrect de Linus Pauling, la découverte du ratio 2 bases puriques – 2 bases pyrimidiques par Erwin Chargaff, la cristallographie par rayon X de la molécule de l'ADN par Rosalind Franklin qui a permis à Watson et Crick de décrire la structure de l'ADN en forme de double hélice en 1953.

***A cette époque les protéines sont toujours considérées comme étant l'élément clé de la vie cellulaire.*** Le rôle des ARN ainsi que le lien entre ARN et protéine était toujours mal connu. Cependant, en 1939, Torbjörn Caspersson et Jean Brachet ont démontré de manière indépendante que le cytoplasme était riche en ARN et que les cellules qui produisaient le plus de protéines sont les cellules qui ont le plus d'ARN. Ce fut la première fois qu'on suppose un rôle des ARN dans la synthèse protéique. En 1958, Francis Crick proposa le « **dogme central de la biologie moléculaire** » qui théorise les liens entre ADN, ARN et Protéines (Figure 7).

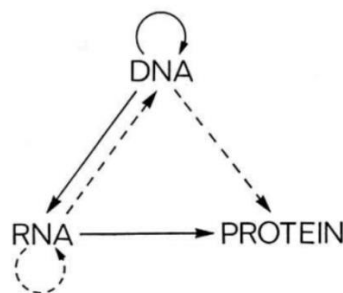


Fig. 3. A tentative classification for the present day. Solid arrows show general transfers; dotted arrows show special transfers. Again, the absent arrows are the undetected transfers specified by the central dogma.

### Figure 7 : Dogme central de la biologie moléculaire selon Crick

Il propose notamment l'existence d'ARNs métaboliques qui pourraient être des transcrits fonctionnels non traduits. Son hypothèse fut validée quelques années plus tard par l'identification des premiers ARNs non codants tels que les ARNs ribosomiaux, les ARNs de transferts et les petits ARNs non codants nucléaires. C'est en **1960** que François Jacob et Jacques Monod ont mis à jour l'existence d'une molécule qui contient l'information génétique nécessaire à la synthèse protéique, cette molécule est l'**ARNmessenger (mRNA)**. A la même époque, Crick publia un article définissant le code génétique : trois nucléotides codent pour un acide aminé. Plus tard, des études ont montré que ce code génétique est universel à travers tous les organismes vivants.

***Bien que les travaux de Monod, Jacob et Crick ont montré de manière indépendante l'importance de l'ARN, de nombreux scientifiques pensaient que les ARNs ne sont que des***

***molécules intermédiaires instables.*** C'est à partir des années 80, avec la découverte d'ARN possédant des activités catalytiques (ribozymes), que les scientifiques ont commencé à s'intéresser plus sérieusement aux ARNs. La découverte d'une activité catalytique des ARN a amené le développement de la « théorie du monde ARN » qui assume que la vie prébiotique est apparu grâce à l'ARN (Higgs & Lehman, 2015).

Le premier ARN non codant régulateur a été découvert en 1983 par l'équipe de Masayuki Inoue à partir de la bactérie *E. coli*. L'ARN non codant *micF* a été montré comme pouvant réprimer la traduction d'un ARNm cible, qui code pour une porine (outer membrane protein F, OmpF), impliquée dans le transport passif à travers la membrane cellulaire. Le premier long ARN non codant eucaryote découvert (lncRNA) est le ***lncRNA H19*** en 1990, cette découverte sera décrite dans la partie III. Cette découverte sera suivie par la découverte du lncRNA Xist (X-inactive specific transcript) en 1991 et son implication dans l'inactivation du chromosome X. Très peu de lncRNA ont été découverts avant le séquençage du génome Humain.

Au début des années 90, de nombreux chercheurs observent de manière indépendante un phénomène d'inhibition de la synthèse protéique. Certaines études ont évoqué ce phénomène chez des végétaux, les nématodes ou les champignons mais aucune n'a suspecté le rôle clé des ARN jusqu'à l'identification du premier microRNA (miR) chez le nématode *Caenorhabditis (C.) elegans* en 1993 par l'équipe de Victor Ambros. Cette équipe a découvert que le gène *lin-4* produit des petits ARN de 22 et 61 nucléotides à partir d'un long précurseur ne codant pas pour une protéine. Ces petits ARN sont produits à partir d'une structure en « tige-boucle », qui est coupée pour générer ces petits ARN qui sont complémentaires antisens à la séquence 3'-UTR du transcrite de *lin-14*. La liaison des petits ARN de *lin-4* à l'ARNm de *lin-14* a été proposée comme étant le mécanisme responsable de la mise sous silence de la transcription du gène *lin-14*. En 2000, un autre miR important a été découvert chez *c. elegans* : le miR let-7. De nombreux homologues de let-7 sont identifiés dans différents organismes dont l'Homme. ***La découverte du phénomène d'interférence par ARN a révolutionné la biologie moléculaire et à fournit un argument supplémentaire en faveur de l'importance des ARN non codants dans la vie cellulaire.***

## **II.2. De la poubelle à la lumière : la révolution des ARN non codants**

### **II.2.1. Concept d'ADN poubelle**

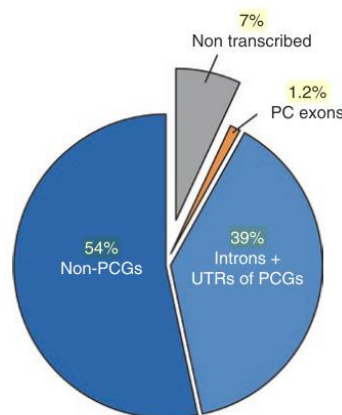
Dans les années 1950, la C-value, ou la quantité d'ADN dans un génome haploïde (taille du génome), n'a pas montré de corrélation avec la taille de l'organisme ou la complexité développementale (Kung et al.,2013). Des petits animaux comme certaines salamandres ont un génome d'une taille supérieure d'environ 15 fois à des animaux plus grands tels que l'homme. Ce paradoxe de la C-value été troublante à l'époque puisque d'un point de vue anthropocentrique les scientifiques pensaient que l'homme était l'espèce la plus complexe sur terre et nécessitait donc d'avoir le maximum de gènes. Ce paradoxe a été en partie résolu par la découverte que la plupart du génome ne code pas pour des protéines. Grâce notamment à des expériences d'hybridations ADN-ARN, il a été déterminé dans les années 1970 que l'homme possède entre 20.000 et 30.000 gènes. Ce qui est très proche du nombre de gènes estimé actuellement aux alentours de 20.000. **Le reste du génome non codant fût appelé ADN poubelle (junk DNA)** du fait de la présence de nombreux transposons, pseudogènes et séquences répétées. Ces séquences non codantes étaient considérées comme un bruit de fond transcriptionnel. En dépit du statut de « poubelle », les séquences non codantes ont intéressé quelques pionniers qui considéraient que « être poubelle ne signifie pas être entièrement inutile » (Comings 1972) et qui ont proposé que ces séquences pouvaient avoir un rôle dans l'intégrité du génome, la régulation génique, la maturation des ARNm et d'autres processus (Kung et al.,2013). A la lumière des dernières avancées dans le domaine, ces chercheurs n'étaient pas loin de la réalité.

### **II.2.2. Révolution du séquençage du génome Humain**

Le premier séquençage d'un organisme a été réalisé sur le bactériophage  $\phi$ X174 par Frederick Sanger en 1978 grâce à la technique qui porte son nom aujourd'hui. Durant les années suivantes, de nombreux génomes viraux vont être séquencés et à la fin des années 90 le projet du séquençage humain (Human Genome Project, HGP) a été développé par la NIH (National Institute of Health). En parallèle du HGP financé par le NIH et d'autres institutions mondiales, le biochimiste Craig Venter fonda une entreprise dans le même but, ce qui créa la course au séquençage. Le premier génome bactérien fut publié en 1995, suivi en 1999 du séquençage d'une portion du chromosome 22 humain. Cette portion couvrait 65% du chromosome 22 connu aujourd'hui et contenait 545 gènes codants pour des protéines, ou prédits codants pour une protéine. Les séquences codantes pour une protéine ne

représentaient que 3% de la séquence totale. La première version du génome humain a été publié par le HGP dans Nature en 2001 et couvrait 96% de l'euchromatine, les travaux de Venter seront publiés le jour suivant dans Science. Le séquençage total fut terminé en 2003. En 2004 les résultats du séquençage humain prédirent 22,287 gènes codants pour des protéines (PCG). Il est établi que seulement 1,2% du génome humain représente des exons codants pour des protéines alors que 24% et 75% sont attribués à de l'ADN non codant intronique et intergénique.

Les données du séquençage humain ont révélé que la plupart du génome est transcrit, que ces séquences codent pour des protéines ou non. Une étude d'hybridation par tiling array avec des sondes ciblant les chromosomes 21 et 22 ont révélé que 90% des transcrits cytosoliques polyadénylés correspondent à des séquences génomiques non codantes, et non à des exons. D'autres études ont donné le même type de résultat. Le consortium **ENCODE** (Encyclopedia of DNA Elements) a compilé jusqu'à 1640 datasets de 147 lignées cellulaires différentes afin d'étudier la composition du génome humain. En 2012 les résultats ont estimé que 93% du génome est activement transcrit et associé à au moins un transcrit primaire. Approximativement 39% du génome représente des gènes codant pour des protéines (défini à partir du promoteur jusqu'au signal poly(A)) et 1,2% correspond aux exons des gènes codant pour des protéines. 54% de ces transcrits ne se trouvent pas dans des PCG (Figure 8).



**Figure 8 : Proportion des gènes codants pour de protéines et des séquences non codantes dans le génome humain**

Figure adaptée de Jarroux et al. 2017.

**Ces résultats montrent donc que la majorité du génome est transcrit en ARN mais qu'une très faible partie de ses ARN sert à produire des protéines. Le génome humain code donc majoritairement pour des ARN non codants. Des résultats similaires ont été obtenus chez la souris et ont permis d'étendre la définition du gène au-delà de sa fonction codant**

**pour une protéine.** Depuis le séquençage humain le nombre d'ARN non codants découverts a augmenté de manière exponentielle. Nous allons décrire ces différentes familles d'ARN non codants dans le paragraphe suivant.

## **II.3. Classification des ARN non codants**

Les transcrits non codants représentent la grande majorité du transcriptome humain et sont exprimés de manière ubiquitaire ou de façon spécifique dans un tissu donné. Les ARN non codants (ncRNA) possèdent de nombreuses fonctions régulatrices, catalytiques et architecturales très importantes dans de nombreux processus biologiques. Une classification basique permet de distinguer deux classes d'ARN non codants en fonction de leur taille. Les petits ARN non codants, dont la taille maximale est fixée arbitrairement à 200 nucléotides (nts) et les longs ARN non codants dont la taille est supérieure à 200 nucléotides. Avec le développement des outils de séquençage et les dernières avancées dans le domaine, de nombreuses classes de ncRNA ont été identifiées. Le rôle de certaines familles de ncRNA dans les processus physiologiques et physiopathologiques reste à démontrer.

### **II.3.1. Les microARN**

#### *II.3.1.1. Généralités sur les petits ARN non codants*

Comme nous l'avons vu dans la partie II.1, au début des années 90 les chercheurs se sont intéressés au phénomène d'inhibition de la synthèse protéique. L'équipe de Victor Ambros a mis à jour les microARN chez *c. elegans*, produit à partir du gène *lin-4*. La liaison des petits ARN de *lin-4* à l'ARNm de *lin-14* a été proposée comme étant le mécanisme responsable de la mise sous silence de la transcription du gène *lin-14*. En **2000**, un autre miR important a été découvert chez *c. elegans* : le miR *let-7*. De nombreux homologues de **let-7** sont identifiés dans différents organismes dont l'Homme.

Les petits ARN non codants ont une taille variant de 18 à 200 nucléotides. Ils sont divisés en deux sous-familles principales : les petits ARNs interférents et les autres petits ARN non codants. Les ARNs interférents comportent les microARN, les siRNAs endogènes (short interfering RNAs), les piRNAs (Piwi-interacting RNAs). Leur taille varie de 18 à 31 nucléotides. Les autres petits ARNs non codants ont une taille de 32 à 200 nucléotides et comprennent notamment les snoRNA (small nucleolar RNA) qui sont impliqués dans les modifications post-transcriptionnelles des autres ARN et les PROMPTs (PROMoter uPstream Transcripts) qui



permettent de réguler la transcription de certains gènes. Les snRNA (small nuclear RNA) sont également des petits ARNs non codants qui sont impliqués dans l'épissage des ARNm (Mohr et al., 2015). Les snRNA vont s'associer avec des protéines impliquées dans l'épissage et former des complexes appelés snRNP (small nuclear ribonucleoproteins ; snRNP). Les snRNA les plus connus sont les snRNA U1 à U6 spliceosomal RNA. Nous allons uniquement nous intéresser aux microARN.

#### *II.3.1.2. Généralités sur les microARN*

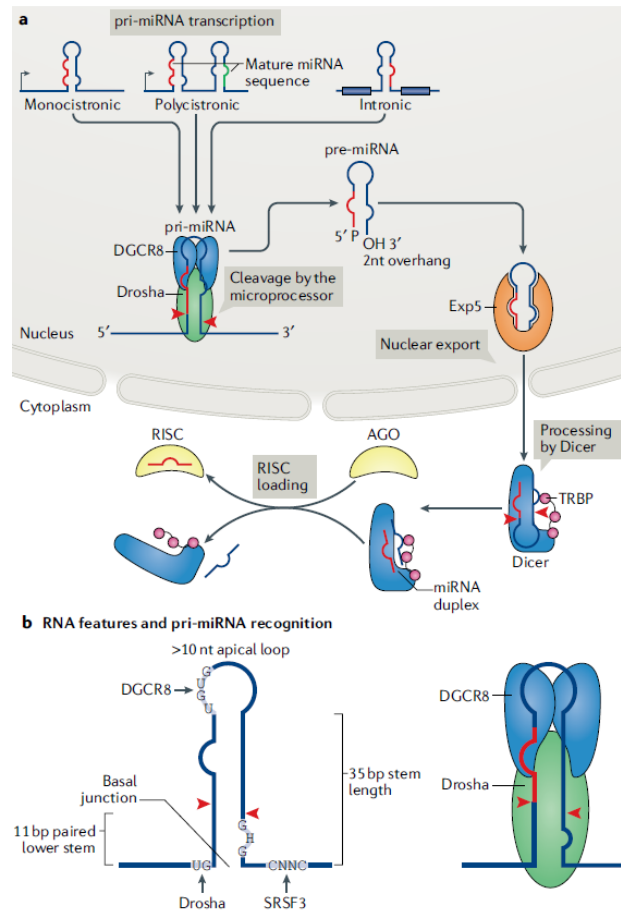
L'avancée des travaux dans le domaine a permis de déterminer plusieurs familles de microARN. Les microARN d'une même famille possèdent la même séquence minimale d'interaction avec son ARNm cible, appelée **seed sequence**, qui permet l'interaction spécifique entre un microARN et son ARNm cible. Il y a une redondance dans l'action des microARN d'une même famille, ils possèdent les mêmes ARNm cibles. Au niveau de la nomenclature, les premiers microARN découverts étaient nommés en fonction des phénotypes associés à leur découverte : lin-7 et let-7 par exemple. Par la suite, les microARN identifiés par séquençage ou clonage ont reçu un numéro : miR-491, miR-675 par exemple. Les gènes permettant la formation des microARN sont notés : mir-xx alors que les microARN matures sont notés : miR-xx. En 2015, 2588 microARN étaient répertoriés chez l'Homme (Hammond, 2015). Grâce à des algorithmes de prédiction de fixation des microARN sur leurs cibles, il est possible de déterminer des cibles putatives des microARN. Ainsi, *in silico* chaque microARN possède des centaines de cibles putatives. De très nombreux microARN sont impliqués dans des mécanismes physiologiques et physiopathologiques.

##### *II.3.1.2.1. Les voies de synthèse des microARN*

Deux voies de synthèse sont à l'origine de la formation des microARN : une voie canonique majoritaire et une voie minoritaire dite non canonique. La description des voies de biogénèse des microARN dans les deux paragraphes suivants est principalement basée sur la revue publiée par Treiber et collaborateurs en Janvier 2019, qui fait le point sur les connaissances actuelles dans le domaine (Treiber et al., 2019).

##### *II.3.1.2.2. La voie canonique*

Les séquences des microARN sont localisées dans des précurseurs d'ARN appelés « pri-miRNA » qui sont des molécules d'ARN contenant des régions en tige-boucle (Figure 9).



**Figure 9 : Voie de biosynthèse canonique des microARN**

A. Les microARNs (miRNAs) des animaux sont encodés comme des gènes individuels (monocistronique), des clusters de gènes (polycistroniques) ou sont situés dans des introns de gènes (introniques). Dans le noyau, l'ARN polymérase II génère le miRNA primaire (pri-miRNA) qui contient une structure tige-boucle (hairpin) et une extrémité 5' et 3'. Le complexe du microprocesseur, constitué du dimère DGCR8 (DiGeorge critical region 8) et Drosha (Drosha ribonucléase III) clive (flèches rouges) le pri-miRNA au niveau de l'origine de la tige et libère un précurseur du miRNA (pre-miRNA) qui possède un résidu hydroxyle (OH) avec une extrémité débordante de deux nucléotides en 3' et une extrémité 5' phosphate (P). L'exportine 5 (Exp5) prend en charge le pre-miRNA afin de faciliter son export vers le cytoplasme, où Dicer effectue un clivage au niveau de l'origine de la boucle (flèches rouges) afin de générer le duplex de miRNA intermédiaire. Dicer, les protéines TRBP (trans-activation-responsive RNA-binding protein) et la protéine argonaute (AGO) s'assemblent pour former le complexe de chargement du RISC (RNA-induced silencing complex). Un des brins du miRNA est transféré sur la protéine argonaute, ce qui amène la formation du complexe RISC. B. Représentation des éléments requis pour le clivage du pri-miRNA. DGCR8 se lie en dimère sur la boucle apicale, alors que Drosha interagit au niveau de la base de la tige. H représente les nucléotides A, C ou U ; N représentent tous les nucléotides. SRSF3 : serine/arginine-rich factor 3. Adapté de Treiber et al., 2019.

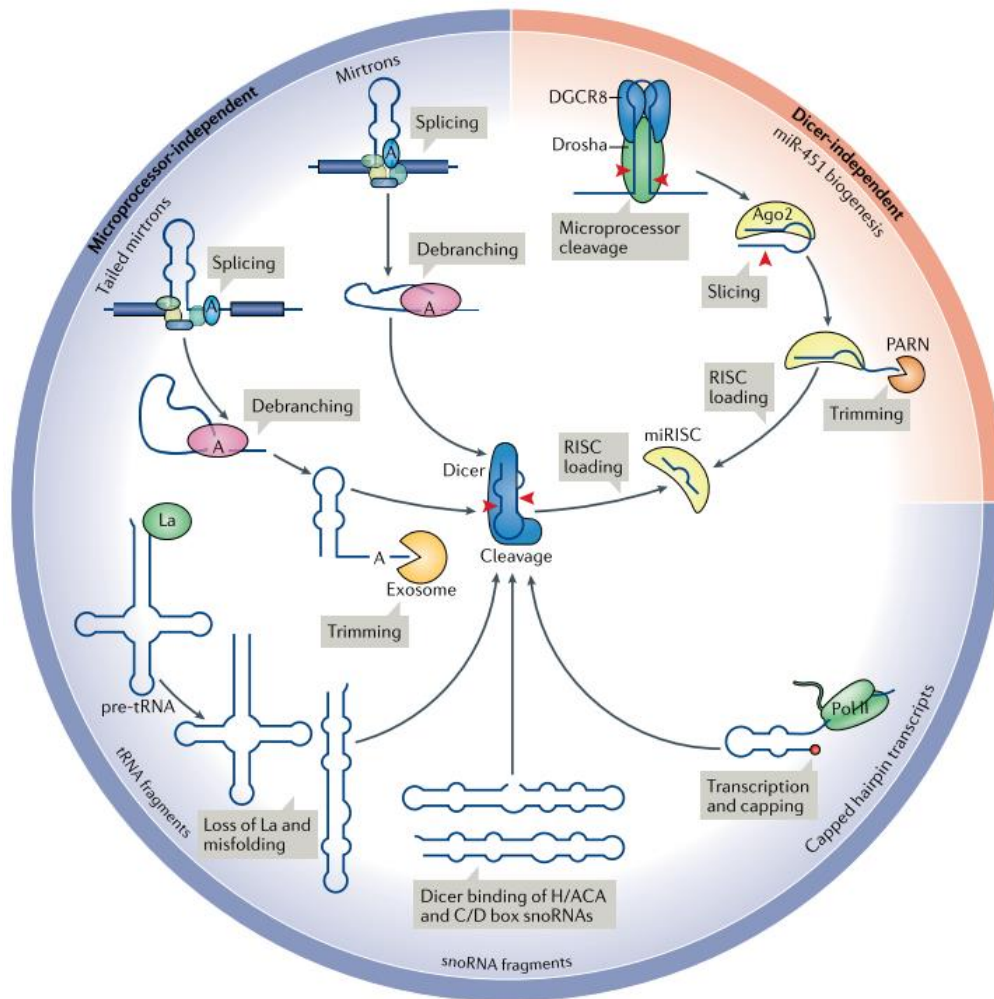
Les pri-miRNA sont généralement portés par les introns des ARNm, les ARN non codants, ou plus rarement, les exons. Dans le noyau, le pri-miRNA est transcrit par l'ARN Pol II puis maturé en pré-miRNA après reconnaissance de la structure tige-boucle du pri-miRNA par le complexe microprocesseur. Le microprocesseur est composé de l'endonucléase Drosha, de la protéine stabilisatrice DGCR8 et d'autres cofacteurs moins connus comme DDX5 et DDX17.

Le clivage du pri-miRNA par Drosha crée une extrémité 3'OH débordante de deux nucléotides et libère un fragment d'environ 60 nucléotides appelé pré-miRNA, qui sera exporté vers le cytoplasme grâce à l'exportin-5.

Dans le cytoplasme, le pré-miRNA est clivé par une deuxième nucléase appelée Dicer (RNase de type III) pour générer un duplex de miRNA de 22 nucléotides. Comme Drosha, la nucléase Dicer est associée à différents cofacteurs comme les protéines TRBP. Dans un processus appelé chargement du complexe d'extinction des gènes induits par l'ARN (RISC), le miRNA produit par Dicer est pris en charge par une nucléase de la famille Piwi/Argonaute (AGO). Le brin guide du miRNA est incorporé de manière sélective dans une protéine AGO, alors que le brin dit passager est dégradé. Aucune marque spécifique n'a été identifiée pour discriminer le brin guide du brin passager, cependant, le brin possédant une extrémité 5' moins stable thermodynamiquement est presque invariablement le brin guide du duplex. La spécificité de la protéine AGO dépend du miRNA associé. Chez l'homme, il existe quatre protéines AGO qui peuvent former le complexe RISC : la protéine AGO2 permettra le clivage de l'ARNm cible tandis que les trois autres (AGO1, AGO3 et AGO4) vont empêcher la traduction de l'ARNm. La formation du RISC va donc s'effectuer par le transfert du miRNA mature vers AGO2 avec l'aide de la protéine TRBP pendant que la protéine AGO2 est maintenue dans une conformation particulière grâce à la protéine chaperonne HSP90 et d'autres co-chaperonnes.

#### *II.3.1.2.3. Les voies non canoniques*

De manière moins fréquente, les miRNA peuvent être synthétisés de façon indépendante du microprocesseur ou de Dicer (Figure 10). Chez les vertébrés, seulement 1% des miRNA sont produits à partir des voies non canoniques. Les « mirtrons » sont une classe importante de miRNA synthétisée par la **voie indépendante du microprocesseur**. Ces miRNA sont originaires d'introns, qui une fois épissés, peuvent être reconnus comme des pré-miRNA dans le noyau et ainsi être exportés sans avoir besoin du clivage par le microprocesseur. Une enzyme de « débranchement » permet une structuration sous forme de pré-miRNA afin de permettre la reconnaissance de l'ARN par Dicer puis la prise en charge dans un complexe RISC. Dans certains cas on parle de « tailed mirtrons », ces mirtrons possèdent des extensions en 5' et 3' qui devront être raccourcies avant l'export du mirtron.



**Figure 10 : Voies de biosynthèse non canoniques des microARN**

Les microARN produits par les voies indépendantes du microprocesseur incluent notamment les mirtrons et les « tailed mirtrons », qui sont générés par épissage puis débranchement. En fonction de leurs séquences, certains ARN de transfert peuvent adopter des conformations alternatives en structure tige-boucle. Ces structures peuvent mimées les pre-miRNAs et être prises en charge directement par Dicer. En condition normale, la chaperone a ARN « La » prévient la formation de ces structures et permet le bon repliement des ARNt. Certains petits ARN nucléolaires (snoRNAs) peuvent servir de substrats à Dicer. Finalement, certains microARN peuvent être synthétisés à partir de structures tige-boucle directement générées par l'ARN Pol II au niveau d'un site start de transcription spécifique. Ces pre-miRNAs ont une coiffe et sont exportées via l'exportine 1. La seule voie de biosynthèse des microARN indépendante de Dicer implique la formation d'une structure tige-boucle inhabituelle sur le miR-451, qui sera directement cliver par Ago2 et prise en charge par la protéine PARN (poly(A)-specific ribonuclease). DGCR8, DiGeorge critical region 8 ; RISC, RNA-induced silencing complex. Adapté de (Treiber et al., 2019).

La comparaison des profils d'expression de miRNA entre des cellules souches embryonnaires wild-type et DGCR-knockout et Dicer-knockout a permis d'identifier des miRNAs qui ne dépendent pas du microprocesseur, mais qui ont besoin de Dicer (Babiarz et al., 2008).

Beaucoup de ces microARNs sont produits à partir d'autres ARN non codants comme les petits ARN nucléolaires (snoRNAs) ou les ARN de transferts (ARNt). De nombreux snoRNAs sont bifonctionnels et sont capables de donner naissance à un microARN. Cependant, très peu de données sont disponibles sur la relevance physiologique de ces microARNs. Certains fragments d'ARNt semblent être chargés dans un complexe RISC. Par exemple, l'ARN de transfert de l'isoleucine peut être digéré par les RNases P et RNase Z pour générer un ARN semblable à un pré-miRNA. Ce pré-miRNA est reconnu par Dicer et permet la formation du miR-1983.

Une autre classe de microARN produite de manière indépendante du microprocesseur sont les miRNA originaires des régions 5' de gènes transcrit par l'ARN Pol II, où le transcrit naissant peut prendre des structures de tige-boucle. Une terminaison précoce de la transcription peut libérer ces ARN possédant des structures tige-boucle, qui pourront servir de substrat à Dicer. Ces ARN étant des transcrits issus de l'ARN Pol II, ils possèdent une coiffe 5' 7-methylguanylate (m<sup>7</sup>G), qui ne sera pas retirée par le microprocesseur (contrairement à la coiffe des pri-miRNA canoniques). Cette coiffe va faciliter l'export de ces ARN via l'exportin-1. À la suite du clivage par Dicer, la coiffe empêche le chargement du bras 5' dans le RISC, l'expression des miRNA de ces précurseurs est donc réduite au bras 3' (Xie et al., 2013). Un seul exemple de **mécanisme de synthèse d'un miRNA Dicer indépendant** est décrit dans la littérature pour le miR-451. Le pri-miR-451 est clivé par le microprocesseur dans le noyau pour former le pré-miR-451 d'une taille de 18 nucléotides et se retrouve exporté dans le cytosol via l'exportin-5. Ce pré-miR-451 est trop petit pour être reconnu par Dicer, il est donc directement pris en charge par AGO2 afin de former un complexe RISC. Le pré-miR451 est clivé dans le complexe RISC par AGO2 afin de former le miR-451 mature.

#### *II.3.1.2.4. Modes d'actions des microARN*

Aujourd'hui il est admis que l'action des miRNA passe par un effet négatif sur leurs ARNm cibles. Cependant, quelques publications et des résultats du laboratoire montrent un effet positif de certains miRNA sur leurs ARNm cibles dans des contextes cellulaires particuliers.

##### *II.3.1.2.4.1. Régulation négative induite par les miRNA*

Brièvement, lorsque le miRNA mature est chargé dans le RISC, le miRNA va permettre au complexe RISC de trouver la séquence complémentaire à la seed sequence du miRNA, qui est souvent présente dans la région 3'UTR des ARNm cibles. La seed sequence du miRNA

localisé entre les nucléotides 2 et 7 est requise pour la reconnaissance spécifique de la cible du miRNA. Au stade précoce de l'inhibition génique induite par les miRNA, l'ARNm cible est toujours stable mais sa traduction est bloquée. Au stade tardif, les ARNm avec une petite queue poly(A) sont dégradés. Cependant, en fonction de la spécificité de l'interaction entre la seed sequence d'un miRNA et l'ARNm cible, un des deux processus peut être favorisé. L'inhibition génique induite par les miRNA va donc empêcher la traduction d'un ARNm et/ou sa dégradation. Les détails moléculaires ne sont pas encore clairement mis à jour.

#### *II.3.1.2.4.2. Régulation positive induite par les miRNA*

En 2007, Vasudevan et collaborateurs ont mis en évidence un effet positif du miR-369-3p sur sa cible, l'ARNm TNF $\alpha$  (Vasudevan et al., 2007). Le miR-369-3p reconnaît des séquences riches en AU (ARE) au sein de l'ARNm de TNF $\alpha$  et s'y fixe. En condition normale de prolifération, le miR-369-3p est associé au complexe RISC et induit la répression de la traduction de l'ARNm. Cependant, les auteurs ont démontré l'association de la protéine FXR1 avec la protéine AGO2 lors d'un arrêt du cycle cellulaire. La présence de FXR1 au niveau des ARE de l'ARNm de TNF $\alpha$ , proche du site de fixation du miR-369-3p, induit un effet positif du miRNA sur sa cible. Lorsque les cellules sont en sevrage, le miR-369-3p aura donc un rôle activateur sur son ARNm cibles. Les auteurs ont également démontré ce processus pour deux autres microARN. Depuis, quelques études ont démontré des résultats semblables (Vasudevan, 2012). Les résultats du laboratoire semblent montrer que le miR-675 est également capable d'induire une activation d'une de ses cibles, la protéine FADD. Nous verrons ces résultats dans la partie résultat.

#### *II.3.1.2.5. MicroARNs et cancer*

Les microARNs sont essentiels au bon fonctionnement cellulaire, une dérégulation des réseaux de microARNs engendre des dysfonctionnements pouvant amener à des pathologies. De nombreuses études ont démontré l'implication des microARNs dans les cancers (Hayes et al. 2014). Tous comme les protéines, les microARNs peuvent agir en tant que suppresseurs de tumeur ou oncogène. Les microARNs ayant une fonction oncogène sont appelés « oncomiRs » et vont avoir pour cibles les ARNm de gènes suppresseurs de tumeurs. À l'inverse, les microARNs ayant une fonction suppresseur de tumeurs vont avoir pour cibles les ARNm des oncogènes. Les membres de la famille des microARNs let-7 sont les microARNs suppresseurs de tumeurs les plus connus. Dans différents cancers, une diminution d'expression de ces microARNs est observée (Balzeau et al., 2017). Les microARNs let-7 ont

pour cibles de nombreux gènes impliqués dans le développement tumoral. *A contrario*, le miR-675 qui est synthétisé à partir du lncRNA *H19* a un rôle d'oncogène. Il est capable de réguler négativement de nombreuses cibles impliquées dans la régulation de la prolifération et du pouvoir migratoire des cellules.

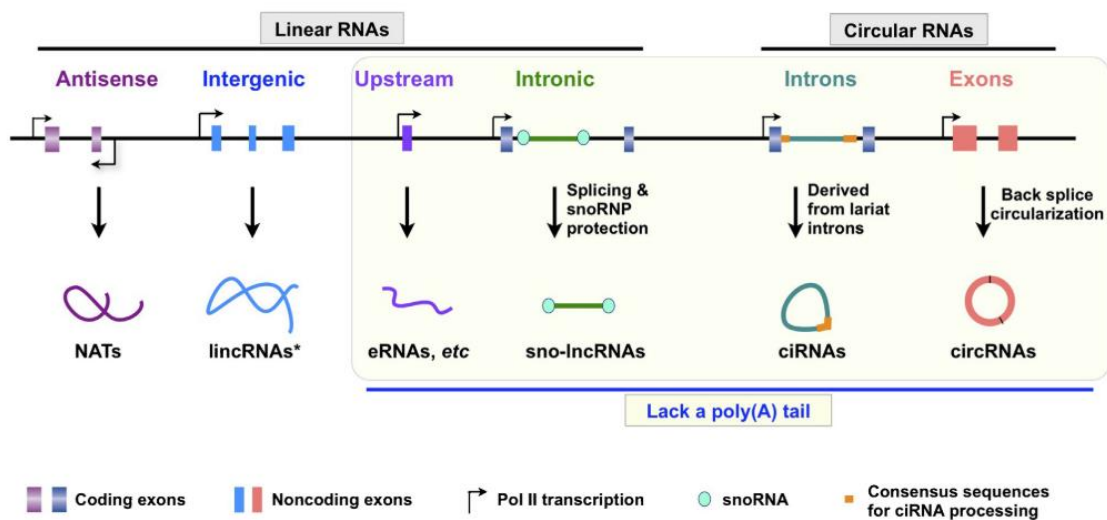
Un des éléments clef dans la compréhension de l'action des microARNs et leurs implications dans les cancers est la notion de « réseaux de microARNs ». Un même microARN peut cibler plusieurs protéines impliquées dans la même voie de signalisation afin de réguler cette voie (Mott & Mohr, 2015). Cependant, plusieurs microARNs peuvent également avoir les mêmes ARNm cibles et donc réguler les mêmes voies de signalisations, on parle alors de réseaux de microARNs. Une modification dans l'expression de certains microARNs peut être compensée par l'expression d'autres microARNs ayant les mêmes cibles. La compréhension des réseaux de microARNs pourrait nous permettre une meilleure compréhension des mécanismes impliqués dans la tumorigenèse. En utilisant les progrès technologiques récents (séquençage haut-débit des microARNs, techniques liées à l'études des microARNs) et des outils de simulation bio-informatique de plus en plus performants, il est maintenant possible d'étudier ces réseaux de microARNs (Dragomir et al., 2018).

Des techniques ont également été développées pour identifier toutes les cibles fonctionnelles des microARNs *in vitro*. Imig et collaborateurs ont décrit une technique permettant l'identification du « miRNA targetome » d'un microARN, c'est-à-dire l'ensemble des cibles d'un microARN donné dans un contexte cellulaire particulier (Imig et al., 2015). Dans cette étude, ils ont trouvé plus de 600 cibles spécifiques du miR-17-5p dont le lncRNA *H19*, ce qui a démontré que le lncRNA *H19* est capable de séquestrer le miR-17-5p. La technique utilisée par les auteurs est le : microRNA crosslinking and immunoprecipitation (miR-CLIP). Brièvement, une sonde spécifique à notre microARN d'intérêt est synthétisée et couplée avec un groupement psoralen et un groupement biotine. Cette sonde est transfectée dans la cellule d'intérêts et les cellules sont exposés à des rayons afin d'induire un crosslinking entre la sonde et les ARNm cibles *via* le groupement psoralen. Une immunoprécipitation est réalisée en utilisant un anticorps anti-Ago2 et les ARNm associés à ce complexe sont récupérés par purification d'affinité en utilisant le groupement biotine de la sonde. Les ARNm sont ensuite rétro-transcrit et soumis à séquençage haut débit afin d'identifier les ARNm cibles de notre microARN d'intérêt. Une sonde contrôle sans biotine est utilisé afin d'obtenir une valeur d'enrichissement.

## II.3.2. Les longs ARN non codants

### II.3.2.1. Généralités et classification

Les longs ARN non codants (lncRNA) sont une classe des ARN non codants, ces lncRNA ont une taille moyenne entre 1 et 10kb. Au niveau moléculaire, les lncRNAs sont proches des ARN messagers (ARNm). Ils sont synthétisés grâce à l'ARN polymérase II et subissent des modifications identiques aux ARNm : ils sont épissés, coiffés en 5' et possèdent une queue polyA en 3'. Cependant, certains lncRNAs comme les eRNAs (enhancers RNA) sont maturés d'une manière différente et ne possèdent pas cette queue polyA (Figure 11 ; Zhang et al., 2014).



**Figure 11 : Différentes classes de lncRNAs chez les mammifères**

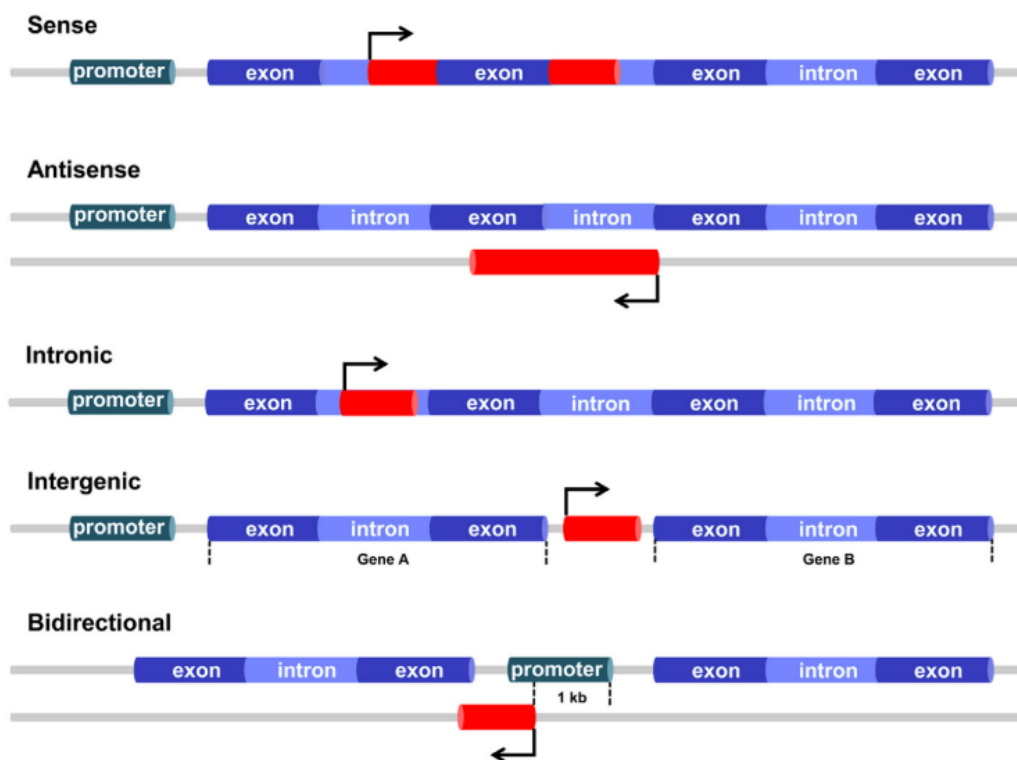
NATs : natural antisens transcripts ; lincRNAs : large intergenic noncoding RNAs ; eRNAs : enhancer RNAs ; sno-lncRNAs : snoRNA-related lncRNAs ; ciRNAs : circular intronic RNAs ; circRNAs : circular RNAs. La légende indique les lncRNAs linéaires et circulaires ainsi que l'absence de la queue polyA. Voir le texte pour plus de détails. Adapté de Zhang et al. 2014

Découverts en 2010, les eRNAs sont produits à partir de séquences d'ADN des régions enhancers du génome (Kim et al. 2010). Peu de données sont disponibles sur le rôle des eRNAs, ils sont impliqués dans des mécanismes épigénétiques et dans certains cancers (Huang et al., 2018). Les sno-lncRNAs (snoRNA-related lncRNAs) sont une classe de lncRNAs qui seront épissés puis protégés à leurs extrémités par la présence de snoRNA (petits ARN non codants). Cependant, certains sno-lncRNA ne possèdent une coiffe snoRNA qu'en 5' et posséderont tout de même une queue poly-A en 3', on parlera de SPA (5'-end snoRNP-capped, 3' polyadenylated lncRNAs). La revue de Xing et collaborateurs recensent les informations connues sur les sno-lncRNAs (Xing et al., 2018). En plus de ces lncRNAs linéaires ne possédant pas de queue polyA, il existe également des lncRNAs qui sont eux circulaires. Ces lncRNAs circulaires (circRNAs) ont



longtemps été considérés comme des sous-produits d'épissages sans fonction biologique. Cependant, de nombreuses publications récentes révèlent un rôle important des circRNAs (Santer et al., 2019). On peut également distinguer les circRNAs dont les séquences proviennent des exons contrairement aux ciRNAs (circular intronic RNAs) dont les séquences proviennent des introns. Sur cette figure nous pouvons également distinguer les lincRNA (long intergenic RNA ; comme *H19*) qui sont des lincRNAs produits à partir de séquences non codantes pour des protéines et les NATs (Natural Antisens) qui sont des lincRNAs produits au niveau d'exons codant pour des protéines mais de manière antisens.

Il existe une dizaine de classification pour classer les ARN non codants et les lincRNAs : en fonction de leurs tailles, leurs structures, leurs fonctions biologiques (St. Laurent et al., 2015). Une des classifications la plus utilisée pour classer les lincRNAs dépend de leur origine de transcription dans le génome (Figure 12).



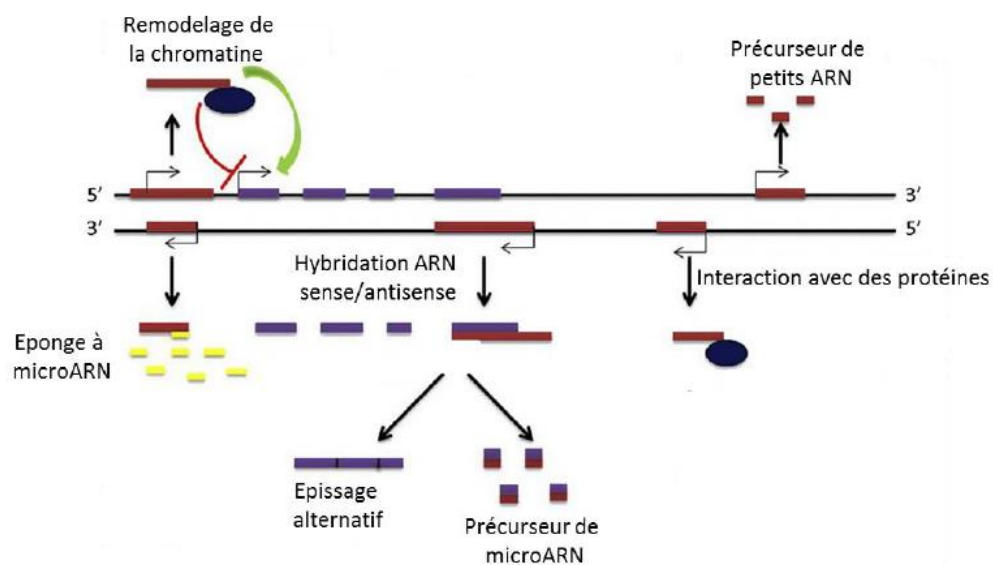
**Figure 12. Classification des long ARN non codants basée sur leurs localisations génomiques**

Les long ARN non codants (lincRNAs) sens sont transcrits à partir du même brin et dans la même direction que les gènes codants pour des protéines (PCGs) proches ; ils peuvent être à la fois introniques et (multi-)exoniques. Les lincRNAs antisens sont transcrits à partir d'un brin opposé aux PCGs et peuvent également être introniques et (multi-)exoniques. Les lincRNAs introniques sont localisés entièrement dans une région intronique d'un PCG. Les lincRNAs intergéniques sont localisés entre deux PCGs et sont transcrits dans la même direction. Les lincRNAs bidirectionnels sont localisés à l'intérieur d'une zone de 1kb d'une région promotrice d'un PCG proche, mais sont transcrits à partir du brin opposé.

En plus des cas représentés sur le schéma, il existe les lncRNAs issus des pseudogènes. Les pseudogènes sont des gènes qui ont perdu la capacité de coder pour une protéine suite à des mutations non-sens ou des décalages de phases ouvertes de lecture (ORF). De nombreux pseudogènes sont transcrits en lncRNAs qui peuvent acquérir de nouvelles fonctions dans la cellule.

### II.3.2.2. Modes d'actions des longs ARNs non codants

De nombreux modes d'actions des lncRNAs ont été décrits dans la littérature (Figure 13).



**Figure 13 : Représentation des principaux modes d'action des longs ARN non codants.**

Dans le noyau, les lncRNAs peuvent réguler l'expression des gènes en recrutant des protéines modificatrices de la chromatine et induire un remodelage de la chromatine. Ils peuvent réguler l'épissage des ARNm en formant des sous-domaines nucléaires, c'est le cas pour le lncRNA MALAT1. Dans le cytoplasme, les lncRNAs peuvent être précurseurs de microARN. Ils sont capables d'empêcher la fonction des microARN en servant d'éponge à microARN. Enfin, ils peuvent fixer des protéines pour modifier leur localisation et/ou leur fonction mais également pour former des complexes particuliers. Illustration réalisée par C. Vennin (2016).

La diversité des modes d'actions des lncRNAs sera illustrée en prenant l'exemple du lncRNA *H19* dans la prochaine partie. Les lncRNAs peuvent interagir avec des protéines afin de participer à des complexes RNP (ribonucléoprotéique) en servant de molécule d'échafaudage. Ils peuvent également interagir avec des complexes de modification de la chromatine afin de participer au remodelage de la chromatine, dans la plupart des cas les lncRNAs vont permettre l'adressage du complexe sur des régions cibles. Un des mécanismes d'action des lncRNAs le mieux décrit est le rôle d'éponge à microARNs. Les lncRNAs peuvent séquestrer des microARNs

afin de diminuer leurs biodisponibilités et lever la répression de leurs cibles. Certains lncRNAs comme MALAT1 peuvent réguler l'épissage des ARNm. Les lncRNAs peuvent également servir de précurseur à des microARNs ou d'autres petits ARN non codants, comme par exemple la synthèse du miR-675 à partir de l'exon 1 du gène *H19*.

### *II.3.2.3. Implications des lncRNAs dans les cancers*

Les lncRNAs peuvent avoir un rôle suppresseur de tumeur ou oncogène. De nombreux lncRNA suppresseurs de tumeurs ont été découverts. Par exemple, le lncRNA *MT1JP* possède un rôle suppresseur de tumeur dans les cancers de la vessie (Yu et al., 2019). En effet, ce lncRNA est capable de séquestrer le miR-214-3p, qui lui est un oncomiR qui favorise la prolifération et l'invasion des cellules cancéreuses. Les auteurs ont montré que le lncRNA *MT1JP* est sous-exprimés dans les patients atteints d'un cancer de la vessie ainsi que dans les lignées cellulaires cancéreuses alors que l'expression du miR-214-3p est augmentée.

De nombreux lncRNAs possédant une activité oncogénique ont également été découvert, dont le lncRNA *H19*. Un autre lncRNA oncogène bien connu est le lncRNA *HOTAIR* (*HOX antisens intergenic RNA*). Le lncRNA *HOTAIR* est un régulateur épigénétique très important, en effet, il se fixe au complexe de modification de la chromatine PRC2 (polycomb repressive complexe 2) et permet la fixation du complexe sur des régions d'ADN d'intérêts afin d'induire une répression de gènes cibles. Il est également capable de se lier à d'autres complexes comme LSD1/CoREST/REST qui est un autre complexe de modification de la chromatine. Dans de nombreux cancers (poumon, sein, gastrique...) une surexpression d'*HOTAIR* est observée et est corrélée à un mauvais pronostic (Tang et al., 2018). Les mécanismes moléculaires impliqués sont divers et peuvent s'expliquer par une dérégulation des marques épigénétiques mais également par un effet d'éponge à microARN. En effet, dans les cancers du poumon le miR-326, qui est un suppresseur de tumeur, est séquestré par *HOTAIR* (Wang et al., 2016). Une liste non exhaustive des lncRNAs impliqués dans les cancers du sein a été réalisé par Youness et collaborateurs (Tableau 3 ; Youness et al., 2019):

lncRNA	Chromosomal Location	Expression Profile	Tumor suppressor or Oncogene	Molecular Function
<i>H19</i>	11p15.5	Upregulated	Oncogene	Sponge Let-7 family miRNAs miRNA precursor for miR-675-3p
<i>HOTAIR</i>	12q13.13	Upregulated	Oncogene	Molecular scaffold
<i>CCAT1</i>	8q24.21	Upregulated	Oncogene	Epigenetic gene silencing
<i>MALAT-1</i>	11q13.1	Upregulated	Oncogene	Sponge Let-7 family miRNAs
				Activates ERK/MAPK pathway
				Induces expression of <i>B-MYB</i>
				Promotes EMT by activating Wnt signaling
				Sponge miR-1; miR-101; miR-217
<i>PVT1</i>	8q24.21	Upregulated	Oncogene	Promotes KLF5/ $\beta$ -catenin signaling pathway
<i>MVIH</i>	10q22 at RPS24	Upregulated	Oncogene	Promotes cellular proliferation, cell cycle progression and inhibiting apoptosis
<i>BCAR4</i>	16p13.13	Upregulated	Oncogene	Required for noncanonical Hedgehog/GLI 2 signal transduction pathways
<i>LINC00152</i>	2p11.2	Upregulated	Oncogene	Promotes cellular proliferation, migration and invasion
<i>PCAT-1</i>	8q24.21	Upregulated	Oncogene	Post-transcriptional repression of the <i>BRCA2</i> 3'UTR
<i>CCAT2</i>	8q24.21	Upregulated	Oncogene	Regulation of Wnt/ $\beta$ -catenin signaling pathway
<i>UCA1</i>	19p13.12	Upregulated	Oncogene	Sponge miR-26a; miR-184; miR-203; miR-129
				Regulation of KLF4-KRT6/13 signaling pathway
<i>Z38</i>	unknown	Upregulated	Oncogene	Unknown
<i>SPRY4-IT1</i>	5q31.3	Upregulated	Oncogene	Unknown
<i>HULC</i>	6p24.3	Upregulated	Oncogene	Sponge miR-186 and miR-372
				Repress the expression of P18
<i>ANRIL</i>	9p21.3	Upregulated	Oncogene	Sponge miR-199a
<i>AFAP1-AS1</i>	4p16.1	Upregulated	Oncogene	Unknown
<i>UCA1</i>	19p13.12	Upregulated	Oncogene	Affects Wnt/ $\beta$ -Catenin Pathway
				Affects mTOR signaling Pathway
				Sponge miR-181a; miR-182; miR-122-5p
<i>SPRY4-IT1</i>	5q31.3	Upregulated	Oncogene	Targets ZNF703
<i>LINP1</i>	10p14	Upregulated	Oncogene	Molecular scaffold
<i>lincRNA-RoR</i>	18q21.31	Upregulated	Oncogene	miRNA sponge
<i>AK058003</i>	Unknown	Upregulated	Oncogene	Regulating $\gamma$ -synuclein gene ( <i>SNCG</i> ) expression
<i>LINK-A</i>	Unknown	Upregulated	Oncogene	Regulation of HIF1 $\alpha$ signaling pathway
<i>DSCAM-AS1</i>	21q22.2	Upregulated	Oncogene	Unknown
<i>HOXA-AS2</i>	7p15.2	Upregulated	Oncogene	Sponge miR-520c-3p
<i>ATB</i>	Unknown	Upregulated	Oncogene	Promotes epithelial-mesenchymal transition by upregulating the miR-200c/ <i>Twist1</i> axis
<i>XIST</i>	Xq13.2	Upregulated	Controversial	X-chromosome silencing
<i>PANDAR</i>	6p21.2	Upregulated	Controversial	Regulation of G1/S transition
<i>LOC554202</i>	9p21.3	Controversial	Controversial	Activation of specific caspase cleavage cascades
<i>NBAT-1</i>	6p22.3	Downregulated	Tumor suppressor	Mediating transcriptional silencing
<i>EPB41L4A-AS2</i>	5q22.2	Downregulated	Tumor suppressor	Unknown
<i>FGF14-AS2</i>	13q33.1	Downregulated	Tumor suppressor	Unknown
<i>BC040587</i>	3q13.31	Downregulated	Tumor suppressor	Unknown
<i>GASS</i>	1q25.1	Downregulated	Tumor suppressor	Interaction with mTOR signaling pathway
<i>LINC00472</i>	6q13	Downregulated	Tumor suppressor	Unknown
<i>MA-LINC1</i>	5q31.3	Downregulated	Tumor suppressor	Vital regulator of cell cycle
<i>LINC RNA-P21</i>	Unknown	Downregulated	Tumor suppressor	Regulates P21 mRNA and protein levels
<i>PTENP1</i>	9p13.3	Downregulated	Tumor suppressor	Upregulates PTEN via its ceRNA interaction on miR-19b
<i>PLNCRNA-1</i>	21q22.12	Downregulated	Tumor suppressor	Induces Apoptosis
				Upregulates TGF- $\beta$ 1
				Downregulates PHGDH
<i>MEG3</i>	14q32.2	Downregulated	Tumor suppressor	Sponge miR-21; miR-421
				Activation of P53

### Tableau 3 : Caractérisation fonctionnelle des lncRNAs dans le cancer du sein

Ce tableau répertorie les lncRNAs impliqués dans les cancers du sein. Leurs profils d'expression, rôle dans la tumorigénèse (oncogène ou suppresseur de tumeurs) ainsi que la voie moléculaire impliqués sont également indiqués. Ce tableau est adapté de Youness et al. 2019.

Un dernier exemple du rôle essentiel des lncRNA est l'action du lncRNA issue du pseudogène *PTENP1*. La gène *PTEN* produit la protéine PTEN qui régule les processus de prolifération et de migration cellulaire. Cette protéine a une fonction suppresseur de tumeur. Le lncRNA *PTENP1* est produit à partir d'un pseudogène ayant une forte homologie avec le gène *PTEN*. En 2010, l'équipe de Poliseno et collaborateurs a démontré que le pseudogène *PTENP1* possède une action suppresseur de tumeur (Poliseno et al., 2010). Certains microARNs ayant une action oncogène vont cibler l'ARNm de la protéine PTEN afin de déréguler la cellule

et favoriser le développement tumoral. Ils ont démontré qu'un lncRNA dérivant de *PTENP1* est ciblé par les mêmes microARN que *PTEN*. Il va donc y avoir une compétition entre la fixation de ces microARN entre les ARN *PTEN* et *PTENP1*. Lorsque *PTENP1* est exprimé, il y aura moins de microARN disponible pour cibler les ARNm de *PTEN* et donc plus de protéines *PTEN* exprimées. Une corrélation positive entre l'expression de *PTEN* et *PTENP1* est retrouvée dans des tumeurs. Dans certains cancers, le locus de *PTENP1* est sélectivement perdu. En protégeant l'ARNm de *PTEN*, le lncRNA *PTENP1* est capable d'avoir un rôle suppresseur de tumeur.

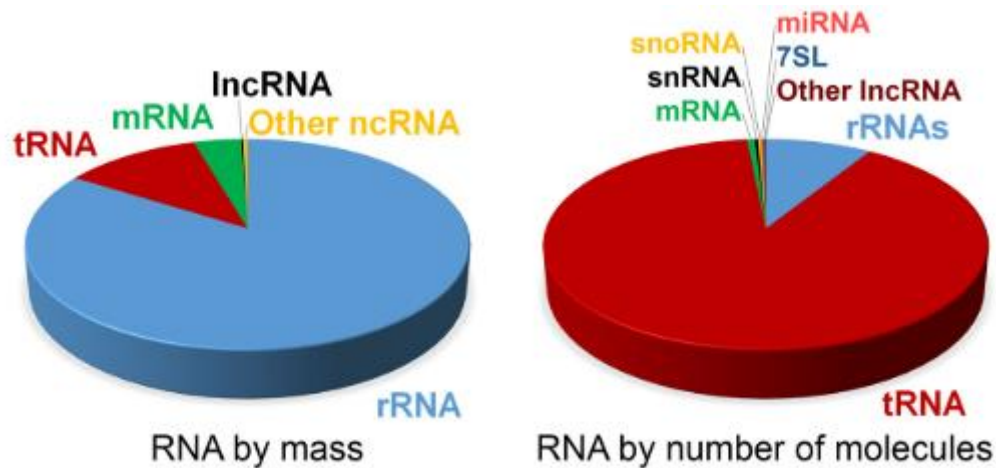
#### II.3.2.4. Implications des lncRNAs dans des phénotypes « non-cancer »

Les lncRNAs sont essentiels au bon fonctionnement des individus et ne sont pas uniquement impliqués dans les cancers. Ils sont également impliqués dans d'autres pathologies comme la maladie d'Alzheimer (Luo et al., 2016). La maladie d'Alzheimer se caractérise notamment par une accumulation de peptides B-amyloïdes. Les chercheurs ont mis en évidence l'implication du lncRNA *BACE1-AS* (B-site amyloid precursor protein cleaving enzyme-1 antisens transcript) dans ce processus. Ce lncRNA est un transcrit antisens au gène codant la protéine *BACE1*, qui est essentiel à la production des dépôts toxiques de peptides B-amyloïdes qui sont impliqués dans le développement de la maladie d'Alzheimer. Il a été démontré que le stress cellulaire induit l'expression de *BACE1-AS* qui va en retour induire l'expression de *BACE1*, ce qui va favoriser la production des peptides B-amyloïdes toxiques.

Certains lncRNAs sont également essentiels au bon fonctionnement cellulaire physiologique. Le lncRNA *Xist* est un lncRNA essentiel à la réalisation de l'inactivation du chromosome X. Contrairement au mâle, les femelles mammifères possèdent deux chromosomes X et il est nécessaire d'inactiver un de ces deux chromosomes. En effet, de nombreux gènes impliqués dans le fonctionnement cellulaire se situent sur le chromosome X et un mécanisme de compensation de dose est nécessaire au bon fonctionnement cellulaire. Le lncRNA *Xist* est indispensable à ce mécanisme d'inactivation du chromosome X. Très brièvement, *Xist* est exprimé à partir du futur chromosome X inactif, il va se propager au niveau du futur X inactif et permettre le processus d'inactivation en entraînant une répression transcriptionnelle du chromosome et l'acquisition de marques épigénétiques répressives sur la chromatine. D'autres lncRNAs impliqués dans l'inactivation du chromosome X ont également été découverts chez la souris : *Tsix* et *Jpx* (Tian, Sun, & Lee, 2008) .

#### II.3.2.4. Places relatives des lncRNAs et des microARNs dans la famille des ARNs non codants

On estime qu'environ 80 à 90% de la masse des ARNs totaux d'une cellule correspond aux ARNs ribosomiques (ARNr), 10 à 15% aux ARNs de transferts (ARNt), 3 à 7% les ARNm, 0,003 à 0,2% les microARNs et les lncRNAs représentent 0.03 à 0,02%. Les lncRNAs et les microARNs sont donc minoritaires par rapport à la masse total des ARNs d'une cellule (Figure 14; Palazzo & Lee, 2015).



**Figure 14 : Estimation des niveaux d'ARNs dans une cellule mammifère**

Figure adaptée de Palazzo & Lee, 2015.

Les deux classes principales d'ARN non codant sont donc les ARNr et les ARN de transferts qui vont avoir pour rôle la synthèse des protéines. Les ARNr vont s'associer avec de nombreuses protéines pour former un complexe ribonucléoprotéique appelé Ribosome. Les ARNt vont interagir avec les ribosomes pour permettre la synthèse protéique. Ils vont être responsable de la reconnaissance des codons présents sur l'ARNm et permettre l'assemblage du peptide naissant au sein du ribosome.

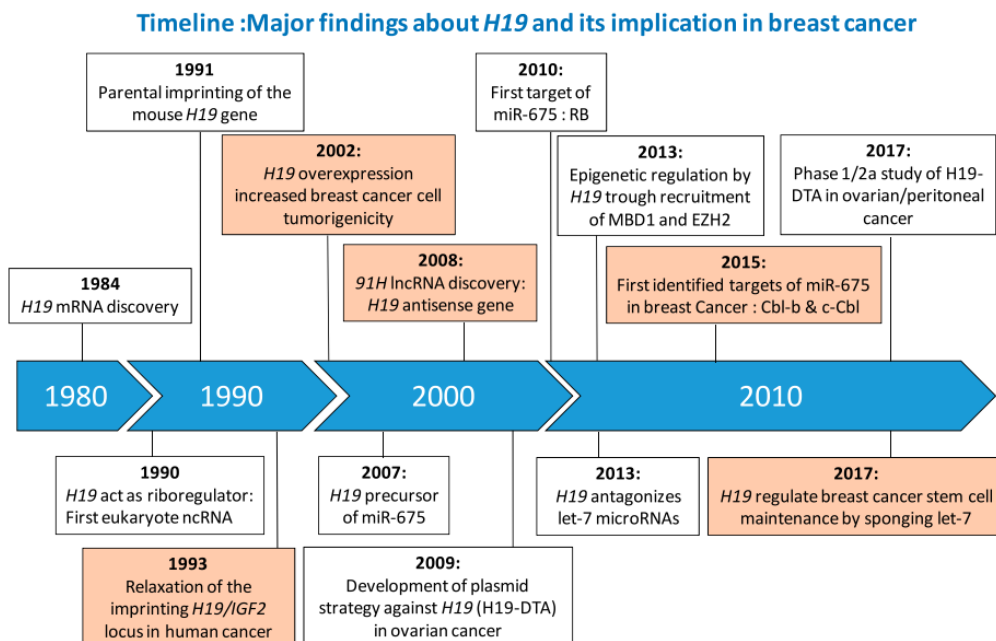
**Bien que les lncRNAs et les microARNs ne représentent qu'une faible partie des ARN non codants présents dans la cellule, leurs implications dans la physiologie humaine et de nombreuses pathologies sont extrêmement importantes. Dans la partie suivante, nous allons nous intéresser au gène *H19* qui est le précurseur du lncRNA *H19* et du miR-675.**

### III. Le long ARN non codant *H19*

#### III.1. Historique du gène *H19*

##### III.1.1. Découverte du gène *H19* et de son action en tant qu'ARN non codant

Le long ARN non codant *H19* est le premier long ARN non codant (lncRNA) eucaryote à avoir été caractérisé comme tel en 1990. Le gène *H19* a été découvert chez la souris par l'équipe de S. Tilghman en 1984 lors d'une étude ayant pour objectif d'identifier les gènes régulés par les gènes *Rif* et *raf*, qui régule en trans l'expression de l'AFP ( $\alpha$ -foetoprotéine) après la naissance (Pachnis et al., 1984). Dans cette étude les auteurs ont criblé une banque de cDNA de foie fœtal chez la souris afin d'identifier ces gènes d'intérêts. Le 19<sup>ème</sup> clone de la ligné H se trouvait être un gène étant sous le contrôle de *raf* et fut nommé *H19*. Les découvertes majeures sur *H19* et son implication dans les cancers du sein sont illustrées dans la Figure 15.



**Figure 15 : Découvertes majeures sur *H19* et son implication dans le cancer du sein.**

Les informations concernant l'implication d'*H19* dans le cancer du sein sont colorées en orange. IGF2, Insulin-like Growth Factor 2 ; lncRNA, Long Non-Coding RNA; RB, Retinoblastoma; MBD1, Methyl-CpG Binding Domain, EZH2, Enhancer of Zeste Homolog 2; Cbl, Casitas B-lineage Lymphoma; *H19*-DTA, Plasmide codant la chaîne A de la toxine diphtérique (DT-A) qui s'exprime sous le contrôle de région régulatrice du promoteur de *H19*. (Collette et al., 2017)

L'équipe de S. Tilghman a été la première à étudier spécifiquement l'ARN produit par le gène *H19* (Brannan et al., 1990). Ils ont démontré qu'un ARN de 2,3kb est produit à partir des 5 exons du gène *H19* grâce à l'ARN Pol II et que cette ARN est épissé, coiffé et polyadénylé. Plusieurs arguments en faveur d'une action du gène *H19* en tant qu'ARN et non de protéine

ont été apportés. Malgré le fait que le gène *H19* soit conservé chez les mammifères, les 35 petites ORFs retrouvés chez l'homme ne sont pas conservés. Ils ont également montré qu'*H19* était présent dans le cytoplasme mais n'était pas associé à la machinerie de traduction et ont conclu que le gène *H19* ne codait pas pour une protéine, mais un ARN qui agira dans la cellule en tant qu'ARN (non codant).

D'autres travaux ont montré que le gène *H19* été surexprimé dans les carcinomes embryonnaires (Wiles, 1988) et qu'il est activé lors de la différenciation in vitro des cellules souches embryonnaires (Poirier et al., 1991). Il a été démontré que le gène *H19* est fortement exprimé durant l'embryogénèse, notamment dans les tissus dérivés du mésoderme et de l'endoderme (Leighton et al., 1995). Son expression est réprimée après la naissance mis à part dans le muscle squelettique et le cœur, cependant, son expression peut être réactivée dans l'utérus et les glandes mammaires notamment lors d'une grossesse (Adriaenssens et al., 1999). Depuis le début des années 2000, le gène *H19* a été retrouvé surexprimé et impliqués dans de nombreux cancers (seins, ovaires, prostate,). Nous verrons dans une partie suivante comment cet ARN non codant peut induire des réponses cellulaires en prenant l'exemple du cancer du sein.

### **III.1.2. Informations sur le locus *H19/IGF2***

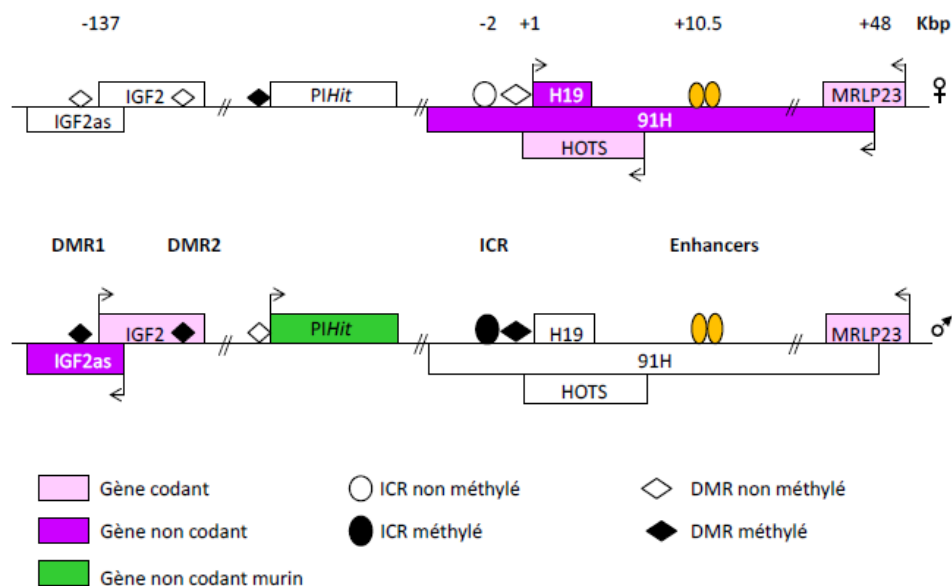
Le gène *H19* fait partie du locus *H19/IGF2* qui est situé chez l'homme sur le chromosome 11 en position 15. Dans ce locus, les deux gènes principaux sont *H19* et *IGF2*. Le gène *H19* est exprimé uniquement à partir de l'allèle maternel alors que le gène *IGF2* est exprimé uniquement à partir de l'allèle paternel. L'expression monoallélique de ces gènes est assurée par le mécanisme de l'empreinte parentale. Le mécanisme de l'empreinte ne sera pas détaillé dans ce manuscrit mais brièvement :

- L'empreinte parentale permet l'expression monoallélique de gènes contenus dans des locus spécifiques
- La mise en place de l'empreinte parentale a lieu durant la gamétogénèse, seul moment où il est possible de différencier les deux génomes parentaux
- Les gènes soumis à l'empreinte jouent un rôle important dans le développement et la croissance fœtale et placentaire
- L'absence d'expression monoallélique des gènes soumis à l'empreinte provoque des pathologies liées au développement



Les dysfonctionnements dans les mécanismes d’empreinte au niveau du locus *H19/IGF2* sont à l’origine de deux syndromes développementaux : les syndromes Beckwith-Wiedmman et Silver-Russel. Le syndrome Beckwith-Wiedmman se caractérise par une expression biallélique du gène *IGF2* ainsi qu’une perte d’expression du gène *H19*. Au niveau pathologique, il est notamment observé un gigantisme du nouveau-né, une macroglossie et une viscéromégalie. Il affecte une naissance sur 12.000 alors que le syndrome de Silver-Russel est plus rare et touche une naissance sur 100.000. Ce syndrome se caractérise au niveau génétique par l’expression biallélique du gène *H19* et la répression du gène *IGF2*. Ce syndrome est associé à un retard de croissance du fœtus et du nouveau-né, une dysmorphie faciale et une asymétrie corporelle sont également retrouvés dans 50% des cas.

De nouveaux transcrits ont été identifiés dans le locus *H19/IGF2* au cours des 20 dernières années (Figure 16).



**Figure 16 : Transcrits identifiés sur le locus *H19/IGF2***

Les transcrits exprimés à partir de l’allèle maternel (♀) ou paternel (♂) sont colorés. A partir de l’allèle maternel sont transcrits : le lncRNA *H19* et *91H* ainsi que le gène codant *HOTS*. Les lncRNAs *IGF2as* et *PIHit* sont transcrits à partir de l’allèle paternel, le gène codant *IGF2* également. Cependant, l’ARN non codant *PIHit* a été identifié uniquement chez la souris. Illustration réalisée par C. Vennin (thèse) et adaptée de (Angrand et al., 2015).

Un ARN non codant transcrit en orientation antisens par rapport au gène *IGF* a été découvert : *IGF2-as* (Okutsu et al., 2000). Notre équipe a découvert un ARN non codant transcrit en orientation antisens par rapport au gène *H19* : *91H*. Le lncRNA *91H* est un transcrit de 120kbp chez l’Homme qui est conservé entre les mammifères. Notre équipe a démontré

que *91H* est surexprimé dans les cancers du sein et possède un rôle oncogène. Une étude a également mis en évidence un transcrit maternel antisens au gène *H19* et compris dans l'ARN non codant *91H* : le transcrit HOTS (*H19* opposite tumor suppressor). Ce transcrit code pour la protéine HOTS qui semble avoir un rôle de suppresseur de tumeurs.

### **III.1.3. Le gène *H19* produit un microARN : le miR-675**

En 2007, Cai & Cullen ont démontré que le lncRNA *H19* est un précurseur de microARN. Les miR-675-3p et miR-675-5p sont produits à partir d'une structure tige-boucle contenue dans l'exon 1 du gène *H19* et qui est conservé entre les espèces (X. Cai & Cullen, 2007). En 2010, Tsang et collaborateurs ont trouvé la première cible du miR-675 : le gène suppresseur de tumeur *Rb* (Tsang et al., 2010). Dans les cellules cancéreuses du colon, le miR-675-5p est capable de se fixer sur l'UTR-3' de l'ARNm de *Rb* afin d'induire une diminution d'expression de cette protéine. L'action du miR-675-5p dans le cancer du côlon est donc une action oncogénique, au même titre que *H19*. Depuis cette étude, de nombreuses cibles du miR-675 ont été identifiées dans différents processus physiologiques et pathologiques. L'implication du miR-675 dans le cancer du sein sera détaillée dans la partie suivante.

## **III.2. Modes d'action du long ARN non codant *H19* : exemple du cancer du sein**

L'implication du lncRNA *H19* dans la tumorigenèse et sa surexpression ont été rapportées dans de nombreux cancers solides comme celui de la prostate, de la vessie ou du sein (Collette et al., 2017). Nous avons démontré qu'*H19* est surexprimé dans 73% des tissus cancéreux de patientes atteints d'un cancer du sein, comparé aux tissus sains (Adriaenssens et al., 1998). De nombreuses études ont montré qu'*H19* est contrôlé par les hormones stéroïdes dans les glandes mammaires normales et cancéreuses ainsi que dans l'utérus et la prostate (Berteaux et al., 2004; Basak et al., 2015). Dans le cancer du sein, l'expression d'*H19* est plus importante dans les cellules ER $\alpha$ + comme les MCF-7, cependant, dans les cellules ER $\alpha$ - comme les MDA-MB-231, la surexpression ectopique d'*H19* est associée avec une prolifération accrue (Adriaenssens et al., 1998; Lottin et al., 2005). Collectivement, ces données montrent qu'*H19* favorise le développement du cancer du sein. Nous allons voir les différents mécanismes qui permettent l'action du lncRNA *H19* dans le cancer du sein. Comme nous l'avons vu dans la partie II.3.2., les lncRNA vont agir *via* différents mécanismes afin d'induire une réponse

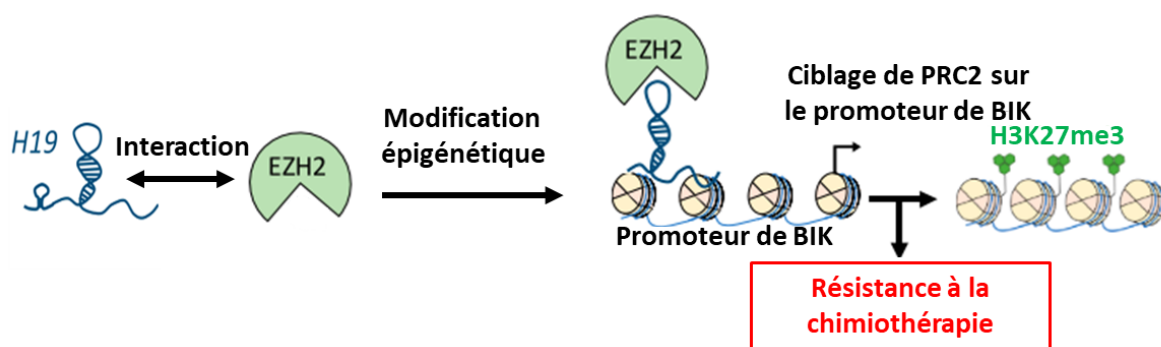
cellulaire. Nous allons prendre l'exemple du lncRNA *H19* et de son implication dans le cancer du sein pour détailler ces modes d'actions. Nous pouvons lister les quelques modes d'actions importants des long ARN non codants :

- Interaction avec des protéines de modification de la chromatine afin d'induire des modifications épigénétiques
- Interaction avec des protéines pour former des complexes ribonucléoprotéiques
- Précurseurs de microARN
- Eponges à microARN afin d'empêcher leurs actions

### III.2.1. Interaction avec des protéines

#### III.2.1.1. Induction de modifications épigénétiques

Dans un contexte de chimiorésistance au paclitaxel, *H19* est capable d'interagir avec la protéine EZH2 afin de cibler le complexe de répression de la chromatine PRC2 (qui contient EZH2) sur des gènes cibles favorisant la survie des cellules (Figure 17; Si et al., 2016)



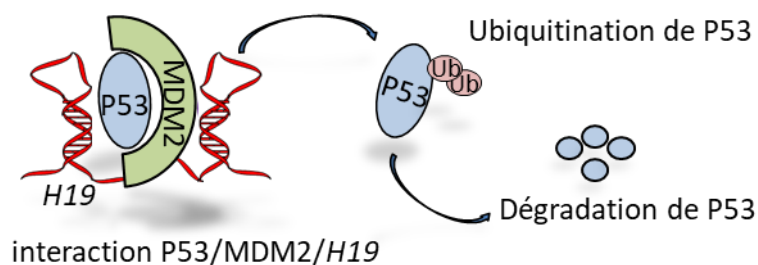
**Figure 17 : *H19* interagit avec EZH2 pour permettre le ciblage du complexe PRC2 sur des gènes cibles dans le cancer du sein**

H3K27me3 : Triméthylation de la lysine 27 de l'histone H3. Figure adaptée de Collette et al., 2017.

Les auteurs ont notamment montré que la répression des gènes *Bik* et *Noxa* observé dans des cellules résistantes au paclitaxel est induite par la surexpression d'*H19*. Dans ces cellules, *H19* interagit avec EZH2 et permet le ciblage du complexe polycomb PRC2 sur ces gènes et induit la triméthylation de la lysine 27 de l'histone H3, ce qui empêche l'activation transcriptionnelle de ces gènes.

### III.2.1.2. Formation de complexes ribonucléoprotéique

En 2012, Yang et collaborateurs ont montré que dans les cancers gastriques, le lncRNA *H19* régule négativement la protéine p53 et est capable d'interagir physiquement avec cette protéine (Yang et al., 2012). Cependant, l'impact de l'interaction physique entre p53 et *H19* n'était pas clair. Mes résultats obtenus au cours de ma thèse nous ont permis de démontrer qu'*H19* est capable de se lier avec p53 mais également MDM2 afin de former un complexe ternaire (Figure 18). Nous pensons que c'est la formation de ce complexe qui est responsable de la régulation négative de p53 en augmentant la dégradation de p53 et en empêchant sa translocation dans le noyau.

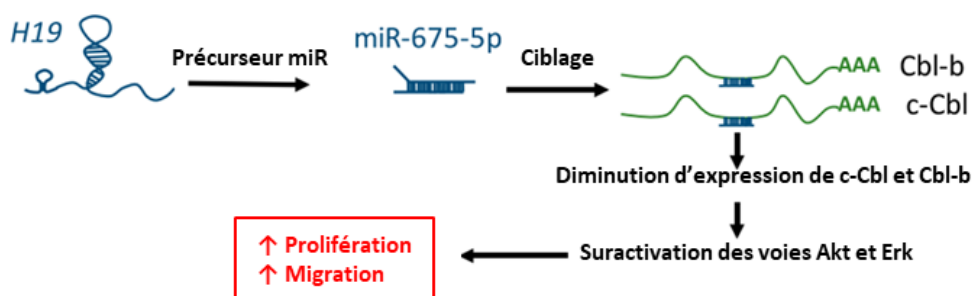


**Figure 18 : Représentation du complexe ternaire formé par l'interaction entre *H19*, P53 et MDM2**

Ub : Ubiquitine. Figure adaptée de Collette et al., 2019 (en révision).

### III.2.2. *H19* est un précurseur de microARN

Comme nous l'avons vu dans la partie précédente, *H19* est le précurseur du miR-675-5p et miR-675-3p. Dans le cancer du sein, nous avons démontré que le miR-675-5p est capable de cibler les ARNm des protéines c-Cbl et Cbl-b (Figure 19; Vennin et al., 2015).



**Figure 19 : Le miR-675-5p réprime les protéines Cbl-b et C-cbl dans le cancer du sein**

Figure adaptée de Collette et al., 2017.

Ces deux protéines sont des E3 Ubiquitin ligase qui régule la stabilité des récepteurs EGFR et c-Met par exemple. Le ciblage de c-Cbl et Cbl-b par le miR-675-5p augmente la stabilité et

l'activation de l'EGFR et de C-Met, ce qui induit la suractivation de la voie Akt et Erk et ainsi augmente la prolifération et la migration cellulaire. A l'heure actuelle, aucune autre cible du miR-675 dans les cancers du sein n'a été identifiée dans la littérature. Cependant, nous verrons dans la partie résultat que nous avons identifié une nouvelle cible de ce microARN : FADD. Une liste non exhaustive des cibles du miR-675 est disponible dans le tableau suivant :

Targets of miR-675-5p	Cellular context	Proteins function
c-Cbl & Cbl-b	Breast cancer	Ubiquitin ligase E3
HDAC 4/5/6	Bone Marrow Mesenchymal Stem Cells	Histone deacetylase
CaMKIId	Cardiomyocyte hypertrophy	Serine threonine protein kinase
USP10	c-kit+ cardiac progenitor cells	Ubiquitin-specific protease
RB	Colorectal cancer	Cell cycle regulator
DDB2	Colon cancer cells	Transcriptional repressor
VDR	Colon cancer cells	Vitamin D receptor
VDAC1	Diabetic cardiomyopathy	Required for mitochondria-mediated apoptosis
REPS2	Esophageal squamous cell carcinoma	Repressor of cell proliferation and migration
CALN1	Gastric cancer	Calcium-binding protein
RUNX1	Gastric cancer	Transcription factor
FADD	Gastric cancer	Apoptotic adaptor that recruits caspase 8 or 10
Cadherin 13	Glioma cell	Atypical cadherin lacking the cytoplasmic domain
RB & TWIST1	Hepatocellular carcinomas	Twist1: Transcription factor
GPR55	Non-small cell lung cancer	G protein-coupled receptor
TGF-β1	Osteoblast differentiation	Growth factor
TGF-β1	Prostate cancer	Growth factor
NOMO1	Placental trophoblast cell	Nodal signaling pathway
ATP8A2	Skeletal cell	Catalytic component of a P4-ATPase flippase complex
CDC6	Skeletal muscle	Essential for the initiation of DNA replication
VDR	Ulcerative Colitis	Vitamin D receptor

Targets of miR-675-3p	Cellular context	Proteins function
Cadherin 11	Melanogenesis	Type II classical cadherin
MITF	Melanogenesis	Transcription factor
IGF1R	Placenta	Insulin-like growth factor 1 receptor
TGF-β1	Osteoblast differentiation	Growth factor
SMAD1 & SMAD5	Skeletal muscle	Intracellular signal transducer and transcriptional modulator

**Tableau 4 : Cibles des miR-675-5p et miR-675-3p**

Ce tableau avec les références associées est disponible dans ma revue (Collette et al.,2017), disponible en annexe.

Certaines de ces cibles pourraient expliquer le rôle oncogénique du gène *H19* dans le cancer du sein. Par exemple, le miR-675 stimule la migration et l'invasion en ciblant TGF- β1 dans les cellules cancéreuses de la prostate, Cadherin13 dans les cellules de gliomes ou RUNX1 dans les cellules cancéreuses gastriques. Nos résultats ont démontré que la surexpression ectopique du lncRNA *H19* et son microARN augmente la migration cellulaire (Vennin et al., 2015). Ce phénomène pourrait être en partie médié par l'action du miR-675 sur ces cibles, cependant, les cibles des microARN semblent être tissu spécifique. Par exemple, le miR-675 est capable de cibler l'expression de Rb dans les tumeurs colorectales afin de promouvoir le développement tumoral (Tsang et al., 2010). Malgré le fait que ce résultat se soit confirmé dans les carcinomes hépatocellulaires (Hernandez et al., 2013), nous n'avons pas réussi à montrer une interaction entre le miR-675-5p et l'ARNm de Rb dans les cellules cancéreuses

mammaires (Vennin et al., 2015). Depuis la publication de ma revue, de nombreuses nouvelles cibles du miR-675 ont été identifiées et sont répertoriées dans la revue de Clément Lecerf (Lecerf et al., 2019).

### III.2.3. *H19* est une éponge à microARN

Un même microARN peut avoir de très nombreux ARNm cibles, ce qui a amené les scientifiques à émettre la théorie des « competing endogenous RNA », c'est-à-dire l'existence de cibles de microARN qui rentre en compétition pour la fixation du microARN. On dit qu'un ARN est une éponge à microARN lorsqu'il est capable de fixer un microARN cible afin d'empêcher la fixation de cet ARN sur ses cibles. Il est important de comprendre que dans ce concept, le microARN régule négativement l'expression de sa cible, mais la cible du microARN est également capable de réguler l'expression du microARN. Tay et collaborateurs ont démontré que les lncRNAs peuvent rentrer en compétition avec les ARNm pour la fixation de microARN communs et ont nommé ces lncRNAs comme des « competing endogenous RNAs (ceRNAs) » ou éponge à microARN. *H19*, comme de nombreux lncRNAs, peut agir via ce mécanisme (Figure 20).



**Figure 20** Représentation du rôle d'éponge à microARN de *H19* dans le cancer du sein

Figure adaptée de Collette et al., 2017.

Récemment, Peng et collaborateurs ont démontré l'implication d'*H19* dans la maintenance du phénotype souche des cellules cancéreuses mammaires en séquestrant les microARN de la famille let-7 (Peng et al., 2017). La surexpression d'*H19* diminue la biodisponibilité de let-7 et ainsi augmente l'expression de sa cible : le facteur de pluripotence LIN28. La surexpression de LIN28 diminue en retour l'expression de let-7 et augmente l'expression d'*H19*. Une autre étude a montré que la surexpression d'*H19* permet l'augmentation d'expression de l'ADN méthyltransférase DNMT1 grâce à la séquestration du miR-152 par *H19* (Z. Li et al., 2017). Cette augmentation d'expression de DNMT1 favorise la prolifération et l'invasion cellulaire. Les auteurs ont également montré une corrélation entre

la surexpression d'*H19*, de DNMT1 et la sous-expression du miR-152 dans des tissus de patientes atteintes d'un cancer du sein. Une étude menée par Zhou et collaborateurs en 2017 a permis de démontrer que le lncRNA *H19* régule la transition épithélio-mésenchymateuse (EMT) et la transition mésenchymateuse-épithéliale (MET ; Zhou et al., 2017). Cette régulation se fait par la séquestration différentielle des microARN mir-200a/b et let-7b par *H19* en utilisant un modèle de souris développant de manière spontanée des métastases du cancer du sein. En 2018, Li et collaborateurs ont démontré que le miR-93-5p est séquestré par le lncRNA *H19* dans les cellules cancéreuses mammaires (Li et al., 2019). En effet, la surexpression d'*H19* diminue la biodisponibilité du miR-93-5p et permet d'accroître les capacités de prolifération, d'invasion et de migration des cellules grâce à la dérégulation de sa cible : la protéine STAT3. En 2019, une étude menée par Haiyan et collaborateurs a démontré que le miR-138 régule l'expression du lncRNA *H19* mais également de SOX4 (Haiyan et al., 2019). Cette étude a démontré que le lncRNA *H19* régule le miR-138 ainsi qu'une de ses cibles, la protéine SOX4, afin de participer à la tumorigenèse mammaire. Une liste non exhaustive des microARNs séquestrés par *H19* est disponible dans le tableau suivant :

miRNAs sponged by <i>H19</i>	Cellular context
hsa-miR-19a/b	Acute myelocytic leukemia
miR-29b-3p	Bladder cancer
miR-152	Breast cancer
let-7	Breast cancer stem cells
miR-455	Cardiac fibrosis
let7	Endometriosis
miR-181-d	Glioblastoma
let-7	HEK293
miR-106-a & miR-17-5p	Hela Cells, myoblast
let-7	Muscle cells
let-7b & miR-200b/c	Mouse breast cancer
miR 141 miR 22	Osteoblast
mir-200s	Osteosarcoma
miR-17-5p	Thyroid cancer

**Tableau 5 : Liste des microARN cibler par *H19***

Ce tableau avec les références associées est disponible dans ma revue (Collette et al.,2017), disponible en annexe.

L'impact d'*H19* dans l'accroissement des capacités métastatiques des cellules cancéreuses mammaires pourrait être dû à la séquestration de ces microARN. Dans le cancer de la thyroïde, *H19* promeut la prolifération, la migration et l'invasion cellulaire en séquestrant

le miR-17-5p (L. Liu et al., 2016). *H19* régule les propriétés tumorigènes des cellules de glioblastomes en séquestrant le miR-181-d (Wu et al., 2017). Cependant, nous pouvons penser que comme la régulation des cibles de microARN, la régulation des microARN par les lncRNA pourrait être tissu spécifique. Il serait donc intéressant d'étudier ces cibles dans le cancer du sein. D'autres modes d'action ont été décrits pour des lncRNA mais ne sont pas, à ce jour, décrit pour *H19*. Par exemple, certains lncRNAs peuvent réguler l'épissage des protéines, former des interactions en duplex ou triplex avec de l'ADN ou d'autres ARN.

#### ***III.2.4. Intérêts cliniques des niveaux d'expression du lncRNA H19 et du miR-675***

En plus de leurs actions sur les phénotypes associées à la tumorigénèse mammaire, les niveaux d'expressions du lncRNA *H19* et du miR-675 dans les tissus cancéreux ainsi que dans le plasma permettent la détection des cancers du sein et renseignent sur leurs gravités. En effet pour le miR-675, des études ont montré que celui-ci est surexprimé dans les tumeurs du sein et qu'il pourrait servir de biomarqueur (Zhai et al., 2015). Zhai et collaborateurs ont étudié l'expression du miR-675-5p dans des tissus fixés au formol et inclus dans la paraffine (FFIP) provenant de 100 patientes atteintes d'un cancer du sein (Zhai et al., 2015). Les auteurs ont montré que ce microARN est surexprimé dans le tissu des patientes atteintes d'un cancer du sein, comparé aux individus contrôles. Cependant, cette surexpression n'est pas corrélée avec le statut clinique et pathologique incluant notamment l'expression des récepteurs ER et PR, l'âge ou le stade d'envahissement ganglionnaire. Une autre étude a évalué les niveaux de méthylation de l'ADN de 517 gènes codant pour des microARN présents dans les leucocytes du sang périphérique, contenu dans des échantillons pré-diagnostiques, de patientes ayant développé un cancer du sein ou un cancer colorectal ainsi que ceux de patients qui sont restés cliniquement sains (Cordero et al., 2015). Ils ont trouvé 8 microARNs, incluant le miR-675-5p, méthylés de manière différentielle dans les sujets ayant développé un cancer du sein. Dans ces sujets, le gène du miR-675-5p est significativement hypométhylés par rapport aux autres patientes, ce qui suggère que ce microARN pourrait servir de biomarqueur dans le cancer du sein. En dehors du contexte du cancer du sein, le miR-675 est retrouvé surexprimé dans le plasma de patients atteints d'un cancer gastrique, comparé aux patients sains (Ghaedi et al., 2019). Dans cette même étude les auteurs ont montré que le lncRNA *H19* est également surexprimé dans les patients atteints d'un cancer gastrique. De manière très intéressante, les auteurs ont réussi à identifier les échantillons issus d'un patient malade et les échantillons



issus d'un patient sain à l'aide uniquement de l'expression du lncRNA *H19*, du miR-675 et d'un autre lncRNA surexprimé dans les patients malades : le lncRNA *MEG3*.

D'autres publications montrent que le lncRNA *H19* est surexprimé dans le plasma de patients atteints de cancer dans différents cancers : sein (Jiao et al., 2018), poumon (Luo et al. 2018) et gastrique (Hashad et al., 2016; Zhou et al., 2015). Comme nous l'avons vu précédemment, la surexpression du lncRNA *H19* dans les cellules cancéreuses a été décrite dans de nombreux cancers.

Un nouvel aspect de l'implication du gène *H19* dans les cancers est en train d'émerger dans la littérature. En effet, des publications récentes montrent l'implication de variant génomique du gène *H19* et des SNP dans le gène *H19* pourraient être associés aux cancers.

Les SNPs (Single Nucleotide Polymorphisms) sont des polymorphismes génétiques qui se caractérisent par la variation d'un seul nucléotide. On recense environ 4 à 5 millions de SNP dans le génome de chaque individu. Lorsqu'un SNP apparaît dans un gène ou une séquence d'ADN importante cela peut conduire à des pathologies. En 2018, Cui et collaborateurs ont démontré qu'un SNP dans le gène *H19* est associé au risque de développement d'un cancer du sein dans une cohorte de patientes chinoises (Cui et al., 2018). Ce SNP a été associé au risque de développer un cancer du sein ER négatif. Le mécanisme permettant l'action de ce SNP reste à déterminer. Une autre étude a montré que d'autres SNPs dans le gène *H19* confèrent également un risque de développement de cancer du sein, dans une cohorte iranienne (Safari et al., 2019).

### **III.2.5. Rôle suppresseur de tumeur du gène *H19***

Aujourd'hui il est clairement établi que le lncRNA *H19* et le miR-675 sont des oncogènes dans le cancer du sein et dans de nombreux cancers. Cependant, des études ont montré un rôle suppresseur de tumeurs pour *H19* et son microARN dans certains contextes. En effet, l'absence d'expression d'*H19* est responsable d'une prédisposition aux tumeurs de Wilms (tumeurs rénales de l'enfant ; Reik et al., 1995). Fukuzawa et collaborateurs ont démontré qu'une inactivation du gène *H19* est fréquemment retrouvée dans les hépatoblastomes sporadiques (Fukuzawa et al., 1999). Plus récemment, Ge et collaborateurs ont également démontré qu'une des voies moléculaires de la cancérisation du foie induite par le virus de

l'hépatite B passe par le lncRNA *H19*. En effet, le virus HBV est capable d'activer des cibles qui vont permettre la surexpression d'*H19* qui pourra réguler négativement l'apoptose et participer à la tumorigenèse (Ge et al., 2019). En 2008, l'équipe de Dandolo a utilisé des modèles de souris pour démontrer un rôle suppresseur de tumeurs de *H19 in vivo* (Yoshimizu et al., 2008). Ils ont utilisé différentes souris transgéniques liés au locus *H19/IGF2* :

- Des souris *H19 $\Delta$ 3* : délétion du gène *H19* sur l'allèle maternel
- Des souris *H19 $\Delta$ Enh* : délétion des enhanceurs en aval du gène *H19* sur l'allèle maternel

Ces souris ont ensuite été croisées avec des souris transgéniques formant des cancers spontanés (colon et foie). Ils ont mis en évidence que l'absence d'expression du gène *H19* favorise l'apparition des tumeurs dans ces modèles, ce qui montre un rôle suppresseur de *H19 in vivo*. Plus récemment, l'équipe du Dr. Ségal-Bendirdjian a démontré que le lncRNA *H19* régule l'activité de la télomérase dans un modèle de leucémie aiguë promyélocytaire (El Hajj et al., 2018). Les auteurs ont démontré que l'utilisation d'ATRA (all-trans retinoic acid) permet de cibler la croissance tumorale en agissant sur la télomérase via deux mécanismes. Premièrement, en diminuant l'expression du gène *hTERT* (human telomerase reverse transcriptase ; sous-unité catalytique de la télomérase) et deuxièmement en augmentant le niveau d'expression d'*H19*. L'augmentation d'expression du lncRNA *H19* va permettre d'inhiber l'action de la télomérase en empêchant l'interaction entre *hTERT* et *hTR* (human telomerase RNA ; ARN nécessaire à l'activité de la télomérase). Dans ce contexte, le lncRNA *H19* possède un rôle de suppresseur de tumeur et permet l'action anticancéreuse de l'ATRA. Cependant, dans un autre contexte (cellules souches cancéreuses du foie) Pu et collaborateurs ont montré que le lncRNA *H19* est capable d'augmenter l'activité de la télomérase *via* un autre mécanisme (Pu et al., 2015). Le rôle du lncRNA *H19* dans la régulation de la télomérase reste donc controversé et on peut imaginer que son action puisse dépendre du contexte cellulaire. En 2018, Wu et collaborateurs ont démontré le rôle suppresseur de tumeur du lncRNA *H19* dans les cancers de la glande pituitaire (Wu et al., 2018). Les auteurs ont découvert qu'*H19* est fréquemment sous-exprimés dans les cancers pituitaires et que son expression est négativement corrélée à la progression tumorale. La surexpression ectopique du gène *H19* dans les cellules tumorales pituitaires inhibe la prolifération cellulaire *in vitro* et la croissance des tumeurs *in vivo*. Le lncRNA *H19* agit en tant qu'inhibiteur compétitif de la liaison entre 4E-BP1 (protéine en aval de la voie mTOR) et Raptor (protéine qui interagit avec mTOR), grâce à un site de fixation identifié sur la protéine 4E-BP1. Dans ce contexte, *H19* possède un rôle de

suppresseur de tumeur en empêchant l'interaction entre 4E-BP1 et Raptor, ce qui a pour effet d'empêcher l'action protumorale de la voie mTOR.

Le miR-675 qui est produit à partir du lncRNA *H19* est également décrit comme un oncomiR dans la grande majorité des publications. Cependant, He et collaborateurs ont montré un rôle suppresseur de tumeur pour le miR-675-5p. En effet, le miR-675-5p est sous-exprimé dans les cellules de cancers du poumon non à petites cellules (NSCLC; He et al., 2015). Ils ont également démontré une corrélation négative entre l'expression du miR-675-5p dans le tissu de patients atteints de NSCLC et le pronostic associé. La cible du miR-675-5p dans ce contexte est la protéine pro-tumorigène GPR55.

**Pour conclure, l'action du long ARN non codant *H19* et du miR-675 dans les cancers est encore sujet à débat et semble être dépendant du contexte cellulaire. Cependant, pour la majorité des modèles et notamment le cancer du sein, l'action du lncRNA *H19* est une action oncogénique.**



# Objectifs



## Objectifs

Le cancer du sein reste, à ce jour, un problème de santé publique. En dépit d'un dépistage plus précoce et d'une meilleure prise en charge thérapeutique, la mortalité associée au cancer du sein reste stable. Le processus de cancérisation qui conduit à la mort d'un patient comporte de nombreuses étapes et il est possible d'étudier chacune de ces étapes par des modèles *in vitro* et *in vivo*. La compréhension des mécanismes impliqués dans la régulation de la tumorigenèse mammaire rendra possible le développement de nouveaux espoirs thérapeutiques. Les lncRNA sont impliqués dans de nombreuses fonctions physiologiques et dans de nombreux cas, les altérations de ces ARN sont associées à des troubles pathologiques. *H19*, le premier lncRNA découvert, est impliqué dans la tumorigenèse mammaire et a récemment été suggéré comme biomarqueur du cancer dans le plasma. Notre équipe a montré que la surexpression d'*H19* augmente les capacités tumorigènes des cellules cancéreuses mammaires mais les mécanismes d'actions associés restent peu clairs. **L'objectif de ma thèse a donc été de mettre en évidence de nouveaux mécanismes d'actions du lncRNA *H19* et de son microARN dans la tumorigenèse mammaire.**

J'ai étudié les mécanismes d'actions associés à la régulation de la protéine p53 par *H19*. Des résultats préliminaires du laboratoire ont identifié que le lncRNA *H19* régule négativement la protéine p53 dans les cellules cancéreuses mammaires. Nous avons mis à jour que le lncRNA *H19* interagit non seulement avec la protéine p53 mais également avec la protéine MDM2 afin de former un complexe tripartite qui régule négativement l'activité de la protéine p53. De plus, j'ai poursuivi mes travaux de Master 2 qui montraient que le lncRNA *H19* régule le phénotype souche des cellules cancéreuses mammaires. Au cours de ma thèse, j'ai pu mettre en évidence que le miR-675-5p est également impliqué dans cette régulation. J'ai également travaillé à la détermination des gènes cibles du lncRNA *H19* et de son microARN par une approche de séquençage de nouvelle génération. Un autre projet du laboratoire a permis d'identifier un nouveau mode d'action inédit du miR-675. Le miR-675 est capable de réguler positivement l'expression d'une de ses cibles, l'ARNm de FADD. J'ai eu l'opportunité de réaliser quelques expériences en lien avec ce projet pour conclure cette étude.





# Résultats



# I. Le long ARN non codant *H19* inhibe p53 et favorise l'instabilité génomique

## Contexte :

Le long ARN non codant *H19* possède un rôle clef dans la progression et la métastase de nombreux cancers. Nous avons déjà démontré que le gène *H19* est impliqué dans la prolifération du cycle cellulaire en étant activé par E2F1 et réprimé par les suppresseurs de tumeurs p53 et RB. Quelques études suggèrent qu'*H19* est capable de réguler l'expression de p53, mais rien n'a été réalisé dans le cancer du sein et le mécanisme moléculaire associé reste à découvrir.

## Matériels et méthodes :

Nous avons mesuré le niveau d'expression de p53 après surexpression d'*H19* dans des lignées de cancers du sein. L'effet d'*H19* sur l'activité transcriptionnelle de p53 en mesurant l'expression des gènes cibles de p53 par qRT-PCR et en utilisant un système rapporteur. La localisation de la protéine p53 a été réalisée par immunocytochimie et fractionnement cellulaire. Nous avons réalisé des immunoprécipitations de l'ARN pour déterminer l'interaction entre *H19* et les protéines p53 et MDM2. Pour étudier l'impact de la surexpression d'*H19* sur l'interaction p53-MDM2 nous avons réalisé des co-immunoprécipitations entre p53 et MDM2. Nous avons étudié l'impact de la surexpression d'*H19* sur les phénotypes cellulaires régulés par p53 en réalisant différentes expériences (cycle cellulaire, apoptose et senescence assay). Finalement, nous avons réalisé un HPRT assay pour déterminer l'implication d'*H19* dans l'instabilité génomique.

## Résultats :

Nous avons démontré qu'*H19* est capable d'interagir avec p53 dans les cancers du sein. Cette interaction induit la dégradation de p53 mais empêche également la fonction de p53 en empêchant sa translocation dans le compartiment nucléaire. Nous montrons qu'*H19* interagit avec p53 mais également MDM2, afin de former un complexe ternaire. De plus, la surexpression d'*H19* diminue l'activité transcriptionnel de p53 et altère son action sur le contrôle du cycle cellulaire, l'induction de l'apoptose et la sénescence des cellules après des dommages à l'ADN. Nous avons également démontré que l'expression d'*H19* favorise l'instabilité génomique, ce qui permet l'accumulation de mutations comme le révèle le HPRT assay.

## **Conclusion :**

L'ensemble de nos résultats révèle un nouveau mécanisme d'action protumoral du long ARN non codant *H19* via la répression de p53 et l'induction d'une instabilité génomique.

## **The long non coding RNA *H19* inhibits p53 and promotes genetic instability**

Jordan Collette<sup>1§</sup>, Constance Vennin<sup>1§</sup>, Clément Lecerf<sup>1</sup>, Mélissa Gilles<sup>2</sup>, Nathalie Spruyt<sup>3</sup>, Nathalie Martin<sup>3</sup>, Corinne Abbadie<sup>3</sup>, Bruno Lefebvre<sup>2</sup>, Xuefen Le Bourhis<sup>1</sup>, Eric Adriaenssens<sup>1</sup>.

<sup>1</sup> INSERM U908, University of Lille, Villeneuve d'Ascq, France

<sup>2</sup> INSERM UMRS1172, University of Lille, CHU-Lille, Lille, France

<sup>3</sup> CNRS UMR 8161, University of Lille, Lille, France

<sup>§</sup> The first two authors contribute equally to this work.

*Correspondence to:* Eric Adriaenssens. INSERM U908, Cité scientifique, Bât SN3, University of Lille, 59655 Villeneuve d'Ascq Cedex, France. Email: [eric.adriaenssens@univ-lille.fr](mailto:eric.adriaenssens@univ-lille.fr)

## **Abstract**

*H19* is a long non-coding RNA 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 E2F, repressed by p53 and RB tumor suppressors and implicated in cell cycle progression. Few studies suggest that *H19* can regulate the expression of p53, but nothing has been done in breast cancer and the functional mechanism remains undeciphered.

## **Methods**

We examined the protein expression level of p53 after *H19* overexpression in breast cancer cells. The effects of *H19* on p53 transcriptional activity were determined using qRT-PCR on p53

target genes and with a reporter system. The localization of p53 protein was determined by immunocytochemistry and cell fractionation. RNA immunoprecipitation was used to determine the interaction between *H19* and the proteins p53 and MDM2. Co-immunoprecipitation of p53 and MDM2 was performed to determine the impact of *H19* overexpression on p53-MDM2 interaction. We examine the impact of *H19* overexpression on p53-regulated cell behavior of breast cancer cells using diverse experiments (cell cycle, apoptosis and senescence assay). Finally, we performed an HPRT assay to determine the implication of *H19* in genomic instability.

## **Results**

We demonstrated that *H19* can interact with p53 in breast cancer cells. This interaction induces p53 degradation but also impairs p53 function under DNA damage by preventing its translocation into the nuclear compartment. We showed that *H19* interacts not only with p53 but also with MDM2 to form a ternary complex. Moreover, *H19* reduces p53 transcriptional activities and impairs cell cycle blockage, apoptosis induction and senescence of cells after DNA damage. Furthermore, we found that *H19* expression favors also genetic instability, allowing for the accumulation of gene mutations as revealed by HPRT assay.

## **Conclusions**

Taken together, our data highlight a novel mechanism of protumoral action of *H19* through the repression of p53 and favoring genetic instability.

## **Keywords**

H19; lncRNA; breast cancer; p53; genetic instability

## **Background**

The *H19* gene, located in human in 11p15.5 locus, is submitted to genomic imprinting. It is expressed only from the maternal allele and transcribed by the RNA polymerase II (Marise S.

Bartolomei, Sharon Zemel, 1991). Although the transcript is spliced, polyadenylated, capped and exported into the cytosol, no protein has been associated. Consequently, Brannan et al. proposed that *H19* RNA functions as a riboregulator (Brannan et al., 1990). *H19* is highly expressed in the embryo, both in the extraembryonic (placenta), and fetal tissues. After birth, its expression is repressed even if a basal expression subsists in several tissues including mammary gland, adrenal gland and uterus (E Adriaenssens et al., 1999b; Eric Adriaenssens et al., 1998; Ariel et al., 1997; J. Liu, Kahri, Heikkila, Ilvesmaki, & Voutilainen, 1995). Although the role of *H19* in cancer is still a matter of debate, increasing studies demonstrate that *H19* is an oncogene in numerous cancers such as bladder, gastric and breast cancer (Ariel et al., 1995; Lottin et al., 2002; Tsang et al., 2010). To promote tumorigenesis, the *H19* long non-coding RNA acts as a regulator at different levels of gene expression including chromatin organization, transcriptional regulation, and post-transcriptional control (Collette et al., 2017).

The p53 protein has been discovered in 1979 (Deleo et al., 1979). It is a transcription factor of 53kDa composed of several domains: a transactivator domain, a prolin-rich domain, a DNA-binding domain, a tetramerisation domain and a Ct domain (Dippold, Jayt, Deleo, Khouryt, & Old, 1981; Rotter, 1983). The p53 protein is constitutively expressed in cells and continuously degraded by the proteasome. In unstressed cells, p53, through its transactivator domain, is associated with the E3 ubiquitin-ligase, MDM2. MDM2 induces the ubiquitination of p53 on its C-terminal domain, which is recognized by the proteasome and promotes p53 proteasomal degradation (Chao, 2014). Under cellular stress including DNA break, irradiation or genotoxic stress, the activated ATM and ATR proteins phosphorylate p53 and MDM2, inducing the dissociation of p53/MDM2 complex and promoting p53 translocation into the nucleus (Toledo, Bluteau, & Simeonova, 2007). According to the cellular stress, p53 promotes cell cycle arrest gene expression such as *p21<sup>waf1</sup>*, or apoptotic gene expression such as *bax* (Freeman & Espinosa, 2012). It is well known that a low level of DNA damage induces a low level of p53-

activated *p21<sup>waf1</sup>* expression; whereas a high level of DNA damage induces a high p53 activation which promotes pro-apoptotic gene expression (Aubrey, Kelly, Janic, Herold, & Strasser, 2017). Thanks to its functions, the p53 protein is a genome keeper and prevents the transformation of normal cells to cancer cells.

We have previously shown that *H19* expression is negatively regulated by p53 in breast cancer cells (Dugimont et al., 1998). In gastric cancer cells, it is established that *H19* physically interacts with p53 protein and this association results in partial p53 inactivation (Yang et al., 2012). However, the biological significance of this binding is not clear. In this study, we identified new mechanisms of *H19*-p53 interaction in breast cancer cells and determined the role of this association on p53 function. Particularly, we found that *H19* interacts with both p53 and MDM2 in a tripartite manner. This interaction favors p53 degradation, and impairs cell cycle blockage, apoptosis induction and cell senescence after DNA damage. This new mechanism associated with *H19* RNA in breast cancer cells promotes the proliferation of aggressive breast cancer cells and could be involved in genetic instability and cell transformation.

## **Materials and Methods**

### ***Cells culture***

The estrogen-sensitive MCF-7 and T-47D breast cancer cell lines were maintained routinely in Dulbecco's Modified Eagle Medium (DMEM, Gibco) or *Roswell Park Memorial Institute medium* (RPMI, Gibco) containing 10% foetal bovine serum (FBS) and 0.01% Zell Shield. The Calu-6 lung cancer cell line was maintained routinely in RPMI containing 10% FBS and 0.01% Zell Shield. The CHO-K1 cell line was maintained routinely in RPMI containing 10% FBS and 0.01% Zell Shield. Cell lines were obtained from the American Type Culture Collection and cultured at 37°C with 5% CO<sub>2</sub>, 95% of air in humidified atmosphere.

### **Construction of plasmids**

The p53 and *H19* constructions were performed as previously described (Dugimont et al., 1998; Vennin et al., 2015). p53 response elements primers were annealed in 95°C gradient. The dimers were cloned in the pGL4.10[luc2] (Invitrogen) plasmid between *Kpn* and *Xho*I enzyme site. The normalising vector pRL-null has no promoter sequence to drive expression of the *Renilla* luciferase gene and was purchased from Promega.

Plasmid productions were performed in *E. Coli* TOP 10 (Invitrogen). Plasmid extraction was performed by using Nucleobond PC100 (Macherey-Nagel).

### **H19 and p53 cDNA transfection**

*H19* and *p53* encoding-plasmids were transfected in Calu-6 and MCF-7 cells previously seeded at  $1.25 \times 10^5$  cells/well in 6-well plates. The desired quantity of plasmid was transfected with Exgen-500 (Euromedex) according to furnisher recommendations. To maintain the same amount of DNA in each sample, the DNA quantity was maintained at 1.5µg with pcDNA3.1(-) empty vector (Invitrogen). Transfection medium was removed after 6 h at 37°C. Proteins were extracted 48 h after transfection.

### **Establishment of H19 overexpressing cell lines**

To establish cell lines overexpressing *H19*, Calu-6 cell line and T-47D breast cancer cell lines were transfected with 1 µg of pcDNA3.1-H19 or pcDNA3.1 empty vector as a control using Exgen-500 and cells were allowed to recover for 48 h. Cells were then selected in the presence of G418 at 0.5 mg/mL (for Calu-6 cell line) or 1 mg/mL (for T-47D cell line) for at least one month before their use for *in vitro* experiments. *H19* overexpressing MDA-MB-231 and MCF-7 cell lines were previously described (Vennin et al., 2015).

### **siRNA transfection**

For siRNA transfection,  $3 \times 10^6$  MCF-7 cells were plated in 100 mm dishes. After 24 h, cells were transfected with negative control (SR-CL000-005, Eurogentec), p53 or MDM2 siRNA (Table S1)



using jetPRIME® Transfection Reagent, according to manufacturer's guidelines (Polyplus Transfection®). Cells were lysed 48 h after transfection.

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

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

### ***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: anti-MDM2 (sc-965, Santa Cruz), anti-p53 (Do-1, sc-126, Santa Cruz), anti RAS-GAP (sc-63, Santa Cruz), anti-PARP-1 (sc-7150, Santa Cruz) and anti-actin (A2066-2ML, Sigma). Actin was used as a loading control. 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).

### ***Nuclear and cytoplasmic protein extraction***

Cells were washed twice with ice-cold PBS and lysed in hypotonic buffer (10mM HEPES, 2mM MgCl<sub>2</sub>, 10 mM KCl, 0.2mM PMSF, 3mM DTT, 10mM EDTA, 0.1% Igepal). After centrifugation (5 min, 4°C, 12.000g), the supernatants were collected as cytosol. Nuclear extracts were prepared by resuspension of the crude nuclei in high salt buffer (50 mM HEPES, 50mM MgCl<sub>2</sub>, 300mM NaCl, 0.2mM PMSF, 3mM DTT, 10mM EDTA, 10% glycerol) during 10 min at 4°C. The

nucleoprotein-containing supernatants were collected after centrifugation (5 min, 4°C, 12 000 g) and conserved at -80°C until use.

#### ***Cells transfection and luciferase activity***

A total of  $20 \times 10^4$  cells/well was seeded in 6-well plates. Twenty four hours after, cells were transfected with 1µg p53-pGL4 and 12ng pRL-null with Exgen-500 (Euromedex) according to manufacturer's guidelines. The luciferase activity was measured 24 h after transfection by using the dual-luciferase reporter assay system according to manufacturer's instructions (Promega). *Firefly* luciferase was normalized with *Renilla* luciferase.

#### ***RNA immunoprecipitation (RIP)***

RNA associated with p53 or MDM2 proteins were immunoprecipitated by using "RNA ChIP-IT® Magnetic Chromatin Immunoprecipitation Kit" (#53024, Active Motif) and an anti-p53 (p53 (DO-1), sc-126, Santa Cruz) or an anti-MDM2 (MDM2 (SMP14), sc-965, Santa Cruz) antibody, according to manufacturer's guidelines.

#### ***Co-immunoprecipitation***

Cells were treated with 10µM of proteasome inhibitor MG132 (#474790, Calbiochem) for 6 h and lysed in HEPES buffer (40mM HEPES pH 7.5, 1mM EDTA pH 8.0, 120mM NaCl, 10mM NaPPi, 1% Triton X-100, 0.1% SDS, 10% glycerol) supplemented with protease and phosphatase inhibitors (50mM NaF, 1.5mM sodium orthovanadate, 1mM PMSF, cOmplete™ protease inhibitor cocktail from Roche 1/100). After centrifugation at 14,000 g, 4°C, 10 min, proteins were extracted and quantified with BCA protein assay (Pierce). The protein G beads were blocked with 1mg of protein 1 h at 4°C on rolling shaker. On the cleared supernatant, 5µg antibody were added and incubated 2 h at 4°C on rolling shaker. Antibodies used were anti-p53, anti-MDM2. Pre-blocked beads were added to antibody-protein complex and incubated for 2 h at 4°C on rolling shaker. The beads were rinsed three times with complete lysis buffer. Immunoprecipitated proteins were eluted and reduced in Laemmli buffer (12.5mM Tris, 2%

glycerol, 0.4% SDS, 1%  $\beta$ -mercaptoethanol, bromophenol blue) at 95°C for 7 min. Proteins were separated on 10% acrylamide gel and analyzed by western blotting as described above.

### ***GST pull down***

Expression of wild-type GST-tagged MDM2 protein and His-tagged p53 protein was induced in *Escherichia coli* strain One Shot™ BL21 (Invitrogen™). His-tagged p53 protein was purified with Ni-NTA fast start kit (Qiagen). MDM2 protein was purified using Glutathione sepharose 4B beads (Sigma Aldrich). The human p53-(1-393) plasmid was a gift from Cheryl Arrowsmith (Addgene plasmid # 24859) and pGEX-4T MDM2 WT plasmid was a gift from Mien-Chie Hung (Addgene plasmid # 16237). Modulation of p53 binding to MDM2 by *H19* presence was analyzed using a GST pull down assay. Briefly, GST-tagged MDM2 and His-tagged p53 proteins were generated as described above and *H19* RNA was produced using Riboprobe *in vitro* transcription systems (Promega) and pcDNA3.1-H19 vector (Bauderlique-Le Roy et al., 2015). A control RNA, SRA was also generated using an expression vector kind gift of Dr B O'Malley (Lanz et al., 1999). GST-tagged MDM2 proteins were incubated with glutathione sepharose beads on rolling shaker at 4°C for 1 h, beads were rinsed twice with PBS-T (PBS, 0.1% Triton) and twice with wash buffer (100mM NaCl). After washing, p53 was added to the mix to allow for its potential binding to MDM2 proteins for 1 h at 4°C. The *in vitro* synthesized RNA was then added to the mix and incubated overnight on rolling shaker at 4°C. The beads were rinsed thrice with wash buffer and proteins were reduced in Laemmli buffer. Proteins were separated by SDS-PAGE and immunoblotted with p53 and MDM2 antibodies as described above.

### ***Immunocytochemistry***

Cells were seeded on collagen-coated coverslips and cultured for 24 h. After treatment with doxorubicine (D1515, Sigma-Aldrich) at 0.2 $\mu$ g/ml for 6 or 24 h, cells were fixed with 4% paraformaldehyde for 10 min, permeabilised with methanol for 10 min at -20°C and then blocked with Ultravision Protein Block (TA-125-PBQ, Thermo Scientific) for 10 min at room

temperature. Cells were incubated with the antibody against p53 (DO-1, sc-126, Santa Cruz, 1/500) overnight at 4°C, rinsed with PBS-tween and then incubated with secondary antibody (A211020, Invitrogen, 1/500) for 1 h at room temperature. The nuclei were counterstained with DAPI (D1306, Invitrogen) for 5 min. Cell staining was analyzed under microscope (Axioplan2, Zeiss).

#### ***Induction of senescence and SA-b-Gal assay***

A total of  $25 \times 10^3$  cells was seeded on 60mm dish and cultured for 24h. To induce senescence, cells were treated with 0.067 $\mu$ g/mL doxorubicin for 72 h. To detect senescence, we employed a senescence-associated-b-Galactosidase (SA-b-Gal) assay using the C<sub>12</sub>FDG b-Galactosidase substrate. The C<sub>12</sub>FDG substrate, cleaved by the  $\beta$ -galactosidase, an enzyme highly active in senescent cells, produces a fluorescent product retained by the cells. Senescent cells were stained with C<sub>12</sub>FDG probe (F2756, Sigma-Aldrich) as previously described (Goehe et al., 2012) and cell fluorescence was quantified by flow cytometry using CyAn™ ADP Analyser (Beckman Coulter).

#### ***Cell cycle analysis***

A total of  $3 \times 10^5$  cells was seeded on 60mm dish and cultured for 24 h. Cells were synchronised in G1/S with 6 $\mu$ M Aphidicolin (Enzo Life Sciences, BML-CC101) for 24 h. Cells were rinsed twice in PBS and treated with 0.2 $\mu$ g/ml doxorubicin for another 24 h. Cells were then permeabilised in 70% cold-ethanol overnight, rinsed in PBS and stained in PBS supplemented with propidium iodide (1 $\mu$ g/ml) and RNase A (1mg/ml) (R6513, Sigma-Aldrich) for 30 min at 37°C. Cell cycle was analyzed using CyAn™ ADP Analyser (Beckman Coulter) and MultiCycle AV software.

#### ***Hprt mutant Frequency determination***

The determination of *Hprt* mutant frequency (MF) was done by measuring clonogenicity of cells treated with cisplatin, as previously described (Silva et al., 2005). Briefly, CHO-K1 cells were grown in HAT medium (Sigma #H0262-10VL) for 5 days to eliminate preexisting *Hprt*

mutants, then  $1.5 \times 10^6$  cells were plated in 75cm<sup>2</sup> flasks. After 24 h of culture, cells were treated or not with 15µM cisplatin (Sigma #C2210000) for 1 h. To determine the potential effect of H19 on colony formation independently of Hprt mutation, cells were plated at a density of 200 viable cells per 35mm dishes (12 dishes/condition) and cultured for ten days before colony estimation. The remaining cells were maintained in exponential growth for eight days to allow the expression of the *Hprt* mutant phenotype. After this period,  $2 \times 10^5$  cells were plated in 100mm petri dishes (20 dishes per condition), in the complete medium containing 6µg/mL of 6-thioguanine (6-TG; Sigma #A4882) for the selection of *Hprt* mutants. In parallel, 200 cells were seeded in 35mm dishes (12 dishes/condition), with non-selective medium. Ten days later, colonies formed in both selective and non-selective media were fixed, stained with 0.5% crystal violet, and scored. Colony efficiency (CE) was expressed as a ratio between the number of formed colonies and the number of seeded cells. MF was expressed as a ratio between the CE of *Hprt* mutants and that of cells cultured in non-selective medium.

### ***Statistical analysis***

Data are expressed as mean values  $\pm$  standard error 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***

### ***H19 reduces the level of p53 protein***

We previously demonstrated that p53 regulates *H19* gene expression (Dugimont et al., 1998) and that *H19* is implicated in cell cycle control (Bertaux, Nathalie; Séverine et al., 2005). To determine the role of *H19* in the regulation of p53 activities, we first examined the expression level of p53 in control (mock) and *H19*-overexpressing (*H19*) cells. As shown in Figure 1A, the levels of p53 protein were decreased in *H19*-overexpressing breast cancer cells,

especially in MCF-7 cells, which possess a wild-type form of p53, but also in T-47D (with mutated form of p53). In MDA-MB-231 (another breast cancer cell lines with mutated form of p53) which expressed high p53 quantities, the decrease is not detectable. Interestingly, the level of p53 mRNA was not significantly modified while the expression of several p53 target genes such as *bax*, *cdkn1a* and *puma* was decreased in MCF-7 cells stably overexpressing *H19* (Fig 1B). Similar results were also observed after transient transfection of cells with a plasmid encoding *H19* (Fig 1C). These results indicate that *H19* post-transcriptionally down-regulates p53 protein, with in consequence a decreased expression of several p53 target genes. To determine if *H19*-induced down-regulation of p53 could be a general mechanism, we co-transfected epithelial Calu-6 cells (p53-null) with different quantities of *p53* and *H19* expression plasmids (Fig 1D). Both genes were expressed under the control of strong promoters. As revealed by western blot, in cells transfected with *H19*, the level of p53 protein is strongly decreased. However, by qRT-PCR, we did not detect any effect of *H19* expression on p53 mRNA level. Overall, these results indicate that *H19* is able to decrease the level of p53 protein and the expression of p53 target genes.

### **H19 mRNA simultaneously associates with both p53 and MDM2 proteins to induce p53 degradation**

The decreased level of p53 protein in *H19* overexpressing cells prompted us to determine if *H19* could favor p53 degradation. It is well known that p53, through its transactivator domain, can associate with the E3 ubiquitin-ligase MDM2, leading to its ubiquitination and proteasomal degradation. We first evaluated the levels of MDM2 in control (mock) and *H19*-overexpressing (*H19*) MCF-7 cells. As revealed by western blot analysis (Fig 2A), an increased level of MDM2 protein was observed in *H19*-overexpressing MCF-7 cells. However, in T-47D cells no difference of MDM2 expression was observed, indicating that other mechanism could exist. Moreover, by performing co-immunoprecipitation of p53 and MDM2 in T-47D cells

treated with the proteasome inhibitor MG132, we found that p53 and MDM2 association was increased in *H19*-overexpressing cells (Fig 2B). Similar results were also observed in MCF-7 cells. Thus, *H19* was able to increase p53 association with its main regulator MDM2. We then performed RNA-immunoprecipitation (RIP) assay to determine the potential interaction of *H19* with p53 or MDM2. Significant high levels of *H19* were found to be co-immunoprecipitated with p53 (Fig 2C) or with MDM2 (Fig 2D) in *H19*-overexpressing (H19) cells, indicating that *H19* did interact with both p53 and MDM2. To determine if this interaction was direct or indirect, we performed GST pulldown assays by using an immobilized glutathione S-transferase (GST)-MDM2 fusion protein with *in vitro* translated His-p53 in presence or not of *in vitro* synthesized *H19* mRNA (Fig 2E). His-p53 interacted with the GST-MDM2 fusion protein but not with the GST alone. Adding *H19* mRNA increased p53/MDM2 association, compared to the conditions in absence of any RNA or in presence of the SRA mRNA, thus indicating the direct interaction of *H19* with p53 and MDM2. To further determine if *H19* could interact with one of the proteins, independently of the other, we performed RIP analysis after knock-down of p53 or MDM2 using siRNA approaches (Fig 2F, G, H). Efficiency of knockdown was checked by western-blotting (Fig 2H). Interestingly, a siRNA against p53 abolished MDM2/*H19* interaction (Fig 2F), and conversely a siRNA against MDM2 abolished p53/*H19* interaction (Fig 2G). In addition, RIP analysis in Calu-6 cells, which are devoid of p53, revealed no enrichment in *H19* mRNA after MDM2 precipitation (Fig 2I). Taken together, these data show that *H19* mRNA directly interacts with p53 and MDM2 in a tripartite manner. This interaction favored p53 and MDM2 binding, leading to p53 degradation (Fig 2J).

### ***H19 impairs p53 translocation from the cytoplasm to the nucleus***

As a transcriptional factor, p53 protein needs to be translocated into the nucleus to induce the expression of its target genes. We determined whether *H19* could modify this process in cells

stressed by DNA damage. For this, we treated MCF-7 cells overexpressing or not *H19* with 0.2 $\mu$ g/mL doxorubicin for 6 h (Fig 3). We observed, by immunofluorescence, that after doxorubicin treatment, the p53 protein was increased and translocated into the nuclei in control cells (Fig. 3A, left panel). In *H19*-overexpressing cells, p53 seemed to be less increased in the nuclei (Fig. 3A, right panel). To confirm these observations, we next performed western blot on cytosolic and nuclear fractions. *H19* overexpression strongly decreased the level of p53 protein in the nuclear fraction, whatever in the non-stressed situation or after doxorubicin treatment (Fig 3B).

### ***H19 impairs cell behaviors regulated by p53***

In order to determine the functional consequence of p53 down regulation by *H19* in breast cancer cells, we first assessed p53 transcriptional activity using a reporter system, in which the luciferase reporter gene was put under the control of p53 response element in pGL4.10[luc2] plasmid. As shown in Fig 4A, the relative luciferase activity in *H19*-overexpressing cells decreased to 50% of control (mock), indicating that p53 transcriptional activities are impaired by *H19*. We then determined several well-known p53-regulated cell behaviors including cell cycle arrest, apoptosis induction and senescence by treating breast cancer cells with the genotoxic drug doxorubicin. For cell cycle analysis, we synchronized cells in G1-phase with Aphidicolin (6  $\mu$ M, 24 h) and then treated them with doxorubicin for 24 h, before analysis by flow cytometry. Doxorubicin induced an accumulation of cells in G2-phase in control cells, in *H19*-overexpressing cells, a slight but significant increase of cells in S-phase and a decrease of cells in G2-phase were observed, compared to control cells (Fig. 4B). This indicates that *H19* overexpression reduces the expected cell cycle arrest. Cell death induction by doxorubicin (0.2  $\mu$ g/mL, 24 h) was checked by flow cytometry after Annexin V/propidium iodide labeling (Fig 4C). While doxorubicin increased necrosis (up to 70%), *H19* overexpression impaired cell death



(less than 45%), indicating that *H19* overexpression reduces necrosis induced by the genotoxic drug. For senescence induction, cells were treated with a lower dose of doxorubicin (0.067 $\mu$ g/mL) for 3 days and senescent cells (C<sub>12</sub>FDG<sup>+</sup>) were detected by flow cytometry (Fig 4D). Mock cells exhibited 15% of C<sub>12</sub>FDG<sup>+</sup> cells, while 10% of C<sub>12</sub>FDG<sup>+</sup> cells were detected in *H19* overexpressing cells (Fig 4D), indicating that *H19* overexpression reduces doxorubicin-induced senescence. Collectively, all these data indicate that *H19* expression reduces transcriptional activities of p53 and impairs cell cycle arrest, apoptosis and senescence, well-known to be mediated by p53 after DNA damage.

### ***H19 favors genomic instability***

To measure the consequence of cell death and cell cycle arrest inhibition mediated by *H19* expression on genomic stability, we performed an *Hprt* assay in CHO K1 cells which possess only one *Hprt* allele. CHO K1 cells were first transfected with *H19* encoding vector to allow stable *H19* expression (Fig 5A). As shown in Fig 5B, *H19* did not modify the colony formation with or without cisplatin, a drug known to induce DNA damages and genomic mutations (Silva et al., 2005). Cells were then cultured in the selection medium containing 6-TG, which is metabolized by HPRT enzyme to give a toxic derivative inducing cell death, so only cells mutated in *Hprt* gene can grow in this condition. As shown in Fig 5C, the *H19* overexpression for at least 10 weeks induced an increase of mutation frequency even in absence of genotoxic drug. As expected, cisplatin increased mutation frequency, this increase was further enhanced in *H19* expressing cells.

## **Discussion**

Given the importance of p53 in diverse cellular pathways, the level of this protein is under strict control by several sophisticated regulatory systems. Among them, MDM2 the main regulator of p53 expression, is known to induce the ubiquitination and proteosomal

degradation of p53. However, other additional mechanisms could participate to p53 expression control. For example, in gastric cancer cells, it is reported that *H19* RNA interacts with the p53 protein (Yang et al., 2012) though the functional consequences of this interaction remain to be elucidated.

In this study, we identified new mechanisms associated with tumor promoting effect of *H19* in breast cancer cells. We demonstrated that *H19* associates with both p53 and MDM2. This association promotes p53 degradation and prevents p53 activation under DNA damage leading to the impairment of cell cycle arrest and cell death induction. In addition, we demonstrated that *H19* is able to induce cell genomic instability (Fig 6).

MDM2 protein is composed of at least four independent functional domains (Shimizu et al., 2002); a N-terminal domain that recognizes the BOX-1 domain of p53, a core domain involved in the binding of p14<sup>ARF</sup> tumor suppressor gene, a Zinc binding domain and a RINF-finger RNA binding domain (Elenbaas, Dobbstein, Roth, Shenk, & Levine, 1996). The p53 protein interacts with MDM2 through its BOX-1 domain located in N-terminal region of p53. Importantly, MDM2 can also bind RNAs. Once bound to RNA, MDM2 changes its conformation. This conformational change reduces MDM2 affinity for the primary binding site in the N-terminal domain of p53 and switches its specificity for a distinct peptide motif in the core domain of p53 (Shimizu et al., 2002). Indeed, MDM2-RNA complexes can bind to p53 protein lacking BOX-1 domain (Burch, Midgley, Currie, Lane, & Hupp, 2000). These data indicate that RNAs can favor MDM2/p53 interaction using binding domains different from those known to be involved in the MDM2/p53 direct interaction. Other long non coding RNAs are known to regulate p53 biology, and in peculiar p53 stability. For example, in atherosclerosis, the association of lincRNA-p21 and MDM2 liberates p53, enabling it to interact with p300 to transactivate its target genes (G. Wu et al., 2014). Another LncRNA, *LincRNA-RoR* has been

described as a strong repressor of p53 in MCF-7 (A. Zhang et al., 2012). In this case, *RoR* suppresses p53 translation through direct interaction with hnRNP I. Our data demonstrate for the first time a tripartite interaction between MDM2/*H19*/p53, leading to p53 degradation. To our knowledge, it is also the first time that a long non coding RNA is found to impair p53 activities by favoring its degradation. Our results together with the previous work of Zhang et al (A. Zhang et al., 2012) indicate that long non coding RNAs such as *H19* and *LincRNA-RoR* constitute an additional regulatory network in the regulation of p53 expression. In addition, our data demonstrate that *H19* is able to impede the p53 translocation in nucleus. Even after treatment with doxorubicin which strongly induces the p53 expression, the phenotypes associated with a transcriptional activity of p53 remain weaker in *H19*-overexpressing cells probably due to p53 retention in the cytoplasm.

The tumor suppressor p53 plays a critical role in preserving DNA fidelity from diverse insults through the regulation of cell-cycle checkpoints, DNA repair, senescence and apoptosis. Defects in the p53 tumor suppressor gene have been shown to affect cell cycle control and lead to genetic instability (as gene amplification, allelic loss, karyotype abnormalities) in murine and human cell lines as well as in primary human breast carcinomas (Eyfjã, Steinarsdottir, Anamthawat-jonsson, & Helga, 1995). In addition, p53 is altered in more than half of human cancers (Chao, 2014). This alteration results from the appearance of point mutation or loss of heterozygosity. Nevertheless, several clinical studies have determined that the use of specific inhibitors of MDM2 (peptides, antibodies, antisense oligonucleotides) makes it possible to reduce tumor growth when p53 is functional (Toledo et al., 2007). However, these inhibitors have a high toxicity rate for the surrounding tissues. Although further studies are clearly needed, our data suggest that alternative therapeutic strategies by targeting *H19* can be envisaged to increase the tumor suppression function of p53 in cancers expressing functional p53.

Although p53 mutation status has been recently shown to modify survival of breast cancer patients (Ungerleider et al., 2018), its status is not used for breast cancer management. However, genetic alterations of p53 are common in breast cancer and are associated with molecular subtypes of tumors; 26% of luminal tumors, 50% of tumors amplified by HER2 and 88% of basal carcinomas show p53 mutations (Bertheau et al., 2013). On the other hand, *H19* is overexpressed in breast carcinomas, independently of molecular subtypes (Eric Adriaenssens et al., 1998; Vennin et al., 2017). As *H19* inhibits p53 activities in breast cancer cells, it is possible that *H19* overexpression mimics p53 mutation in terms of p53 repression, and that loss of p53 activity rather than the mutation alone should be a prognostic factor.

In cells expressing *H19*, transcriptional activities of p53 are impaired. Target genes of p53 such as p21 or bax are down regulated, leading to the decrease of cell cycle arrest and apoptosis and probably the accumulation of cells harboring genetic mutations. In this line, *H19* could contribute to tumorigenesis by inducing genetic alterations. Interestingly, the mutation rate increases in cells expressing *H19* without exposure to genotoxic drugs. So, *H19* could be a factor allowing genomic instability in persons not exposed to genotoxins, through a mechanism involving p53/*H19* interplay. Further experiments should be performed to show if this regulation exists in normal cells.

We previously showed that *H19* is down regulated by p53 (Dugimont et al., 1998). Therefore, our present data indicate that *H19* and p53 are reciprocally regulated and form feedback loops. Interestingly, the same feedback loop exists in the case of *RoR*/p53 since p53 transcriptionally induces *RoR* expression which represses, in turn, p53 activities (A. Zhang et al., 2012). *H19* seems to exert a pivotal role in cell cycle progression. It is activated by E2F factors (Nathalie Berteaux et al., 2005), negatively regulated by RB (Tsang et al., 2010) and p53 (Dugimont et

al., 1998) and we showed here that *H19* RNA represses p53 activities to allow cell cycle progression.

Altogether, we highlight a new mechanism of *H19* RNA in cancer cells. We demonstrated that *H19*/p53 association induces p53 degradation but also impairs p53 function under DNA damage by preventing its translocation into the nuclear compartment. This mechanism, which could allow the proliferation of cells with DNA breaks, highlights that *H19* can be involved in cell transformation.

## References

1. Bartolomei MS, Zemel S, Tilghman SM. Parental imprinting of the mouse H19 gene. *Nature*. 1991;351(6322):153–5.
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(1):28–36.
3. Adriaenssens E, Dumont L, Lottin S, Bolle D, Leprêtre A, Delobelle A, et al. 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(5):1597–607.
4. Adriaenssens E, Lottin S, Dugimont T, Fauquette W, Coll J, Dupouy JP, et al. Steroid hormones modulate H19 gene expression in both mammary gland and uterus. *Oncogene*. 1999;18(31):4460–73.
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(2):492–6.
6. Ariel I, Weinstein D, Voutilainen R, Tamar S, Lustig-Yariv O, Nathan D, et al. Genomic Imprinting and the Endometrial Cycle. The Expression of the Imprinted Gene H19 in the Human Female Reproductive Organs. *Diagnostic molecular pathology: the American journal of surgical pathology*. 1997;6(1):17–25.
7. Ariel I, Lustig O, Tamar S, D GPM, Mally S, Nathan D, et al. The imprinted H19 gene as a tumor marker in bladder carcinoma. *Urology*. 1995;45(2):335–8.
8. Tsang WP, Ng EKO, Ng SSM, Jin H, Yu J, Sung JJY, et al. Oncofetal H19-derived miR-675 regulates tumor suppressor RB in human colorectal cancer. *Carcinogenesis*. 2010;31(3):350–8.
9. Lottin S, Adriaenssens E, Dupressoir T, Berteaux N, Montpellier C, Coll J, et al. Overexpression of an ectopic H19 gene enhances the tumorigenic properties of breast cancer cells the phenotype of human breast epithelial cells. *Carcinogenesis*. 2002;23(11):1885–95.
10. Collette J, Le Bourhis X, Adriaenssens E. Regulation of human breast cancer by the long non-coding RNA H19. *Int J Mol Sci*. 2017;18(11).
11. Deleo AB, Jay G, Appella E, Dubois GC, Law LW, Old LJ. Detection of a transformation-related antigen in chemically induced sarcomas and other transformed cells of the mouse *Immunology*. *Proc Natl Acad Sci*. 1979;76(5):2420–4.
12. Dippold WG, Jay G, Deleo AB, Khoury G, Old LJ. p53 transformation-related protein: Detection by monoclonal antibody in mouse and human cells. *Proc Natl Acad Sci*. 1981;78(3):1695–9.
13. Rotter V. p53, a transformation-related cellular-encoded protein, can be used as a biochemical marker for the detection of primary mouse tumor cells. *Proc Natl Acad Sci*. 1983;80(9):2613–7.

14. Chao CC. Mechanisms of p53 degradation. *Clin Chim Acta*. 2014;438,139-147.
15. Toledo F, Bluteau O, Simeonova I. Réactivation de p53 dans les tumeurs. Une stratégie antitumorale prometteuse. *Médecines/Sciences*. 2007;23(6-7):565–7.
16. Freeman JA, Espinosa JM. The impact of post-transcriptional regulation in the p53 network. *Briefings in functional genomics*. 2012;12(1):46–57.
17. Aubrey BJ, Kelly GL, Janic A, Herold MJ, Strasser A. How does p53 induce apoptosis and how does this relate to p53-mediated tumour suppression ? *Cell Death and Differentiation*.2017;25(1):104–13.
18. Dugimont T, Montpellier C, Adriaenssens E, Lottin S, Dumont L, lotsova V, et al. The H19 TATA-less promoter is efficiently repressed by wild-type tumor suppressor gene product p53. *Oncogene*. 1998;16(18):2395–401.
19. Yang F, Bi J, Xue X, Zheng L, Zhi K, Hua J, et al. Up-regulated long non-coding RNA H19 contributes to proliferation of gastric cancer cells.*The FEBS Journal*. 2012;279(17):3159-3165.
20. Vennin C, Spruyt N, Dahmani F, Julien S, Bertucci F, Finetti P, et al. The long non coding RNA-derived miR-675 enhances tumorigenesis and metastasis of breast cancer cells by downregulating c-Cbl and Cbl-b. *Oncotarget*. 2015;6(30):29209–23.
21. Berteaux N, Aptel N, Cathala G, Genton C, Coll J, Daccache A, et al. A Novel H19 Antisense RNA Overexpressed in Breast Cancer Contributes to Paternal IGF2 Expression. *Mol Cell Biol*. 2008;28(22):6731–45.
22. Bauderlique-Le Roy H, Vennin C, Brocqueville G, Spruyt N, Adriaenssens E, Bourette RP. 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*. 2015;24(10):1252–62.
23. Lanz RB, Mckenna NJ, Onate SA, Albrecht U, Wong J, Tsai SY, et al. A Steroid Receptor Coactivator, SRA, Functions as an RNA and Is Present in an SRC-1 Complex. *Cell*. 1999;97(1):17–27.
24. Goehe RW, Di X, Sharma K, Bristol ML, Henderson SC, Valerie K, et al. The Autophagy-Senescence Connection in Chemotherapy: Must Tumor Cells (Self) Eat Before They Sleep? *The Journal of Pharmacology and Experimental Therapeutics*. 2012;343(3):763–78.
25. Silva MJ, Costa P, Dias A, Valente M, Louro H, Boavida MG, et al. Comparative Analysis of the Mutagenic Activity of Oxaliplatin and Cisplatin in the Hprt Gene of CHO Cells. *Environ Mol Mutagen*. 2005;46(2):104–15.
26. Bertaux S, Lottin S, Monté S, Quatannens B, Coll J, Hondermarck H, Curgy J, et al. H19 mRNA-like Noncoding RNA Promotes Breast Cancer Cell Proliferation through Positive Control by E2F1.*The Journal of Biological Chemistry*. 2005;280(33):29625–36.
27. Shimizu H, Burch LR, Smith AJ, Dornan D, Wallace M, Ball KL, et al. The Conformationally Flexible S9 – S10 Linker Region in the Core Domain of p53 Contains a Novel MDM2 Binding Site Whose Mutation Increases Ubiquitination of p53 in Vivo. *The Journal of Biological Chemistry*. 2002;277(32):28446–58.
28. Elenbaas B, Dobbstein M, Roth J, Shenk T, Levine AJ. The MDM2 Oncoprotein Binds Specifically to RNA through Its RING Finger Domain. *Molecular Medecine*.1996;2(4):439-451.
29. Burch LRY, Midgley CA, Currie RA, Lane DP, Hupp TR. Mdm2 binding to a conformationally sensitive domain on p53 can be modulated by RNA. *The FEBS Letters*. 2000;472(1):93–8.
30. Wu G, Cai J, Han Y, Chen J, Huang Z, Chen C, et al. LincRNA-p21 Regulates Neointima Formation, Vascular Smooth Muscle Cell Proliferation , Apoptosis and Atherosclerosis by Enhancing p53 Activity.*Circulation*. 2014;130:1452-1465.
31. Zhang A, Zhou N, Huang J, Liu Q, Fukuda K, Ma D, et al. The human long non-coding RNA-RoR is a p53 repressor in response to DNA damage. *Cell Research*. 2012;23(3):340–50.
32. Eyfjã JE, Steinarsdottir M, Anamthawat-jonsson K, Helga MÃ. p53 Abnormalities and Genomic Instability in Primary Human Breast Carcinomas.*Cancer Research*. 1995;55(3):646–51.

33. Ungerleider NA, Rao SG, Shahbandi A, Yee D, Niu T, Frey WD, et al. Breast cancer survival predicted by TP53 mutation status differs markedly depending on treatment. *Breast Cancer Research*.2018;20(1):115.
34. Bertheau P, Lehmann-che J, Varna M, Dumay A, Poirot B, Porcher R, et al. p53 in breast cancer subtypes and new insights into response to chemotherapy. *The Breast*. 2013;22:S27-S29.
35. Vennin C, Spruyt N, Robin YM, Chassat T, Le Bourhis X, Adriaenssens E. The long non-coding RNA 91H increases aggressive phenotype of breast cancer cells and up-regulates H19/IGF2 expression through epigenetic modifications. *Cancer Letters*. 2017;385:198–206
36. Berteaux N, Lottin S, Monté D, Pinte S, Quatannens B, Coll J, et al. H19 mRNA-like noncoding RNA promotes breast cancer cell proliferation through positive control by E2F1. *J Biol Chem*. 2005;280(33):29625–36.

**Funding.** This work was supported by INSERM and grant from INCA (PLBio 2010-180) and “Ligue contre le cancer”.

C.V. was supported by doctoral fellowships from the University of Lille 2 and by ARC Fondation fellowships. J.C.

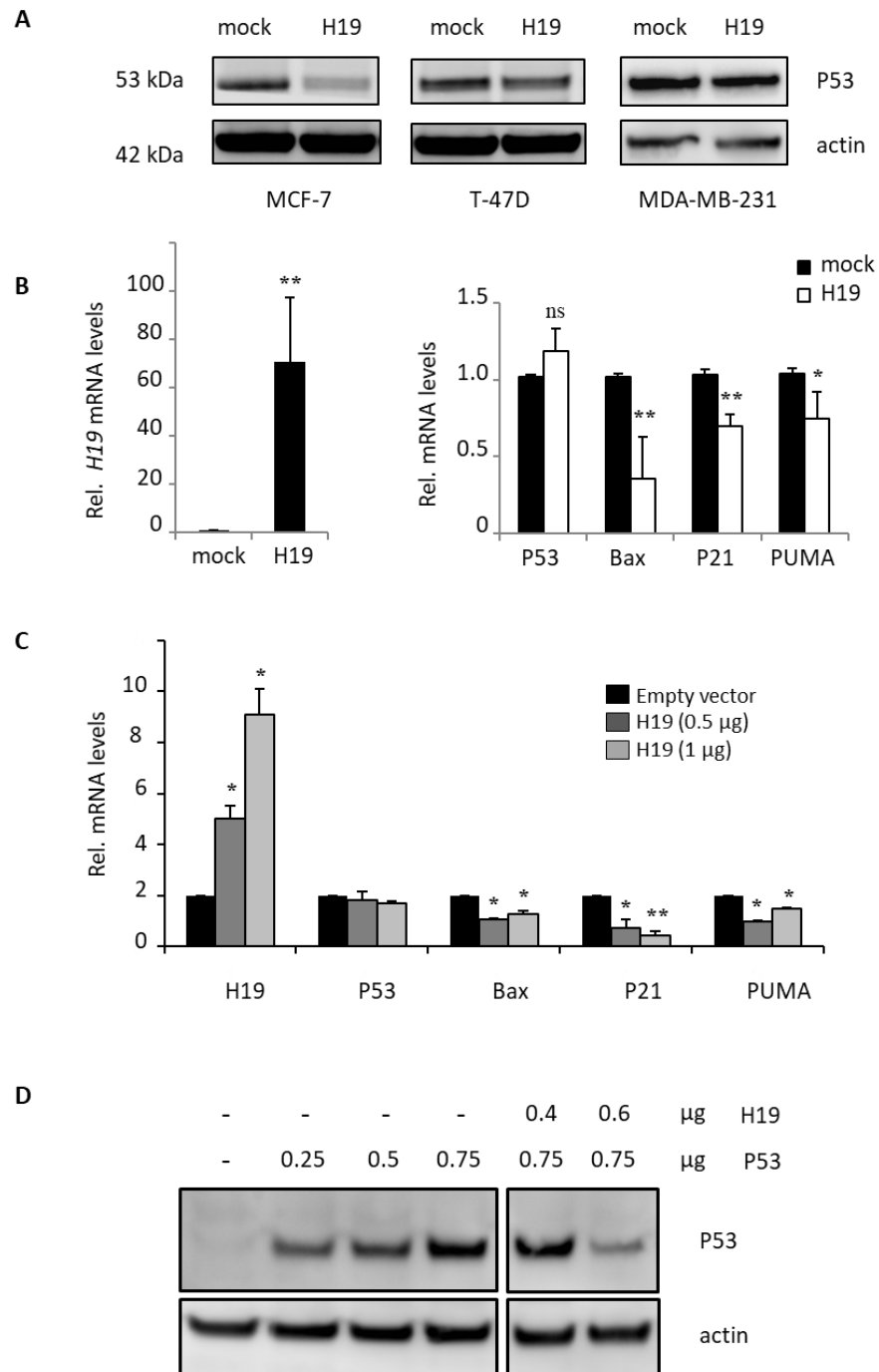
and C.L. were supported by doctoral fellowships from the University of Lille.

**Acknowledgements.** We thank all members of laboratory for helpful discussion.

**Conflicts of interest.** Authors declare that they have no conflicts of interest.

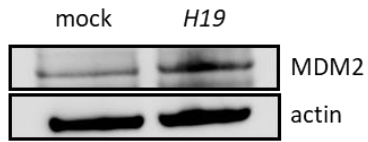
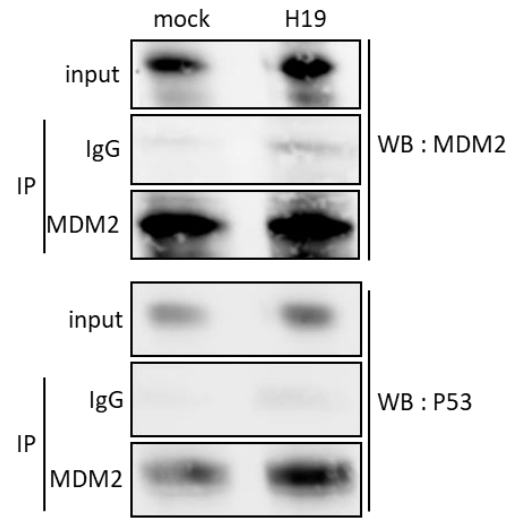
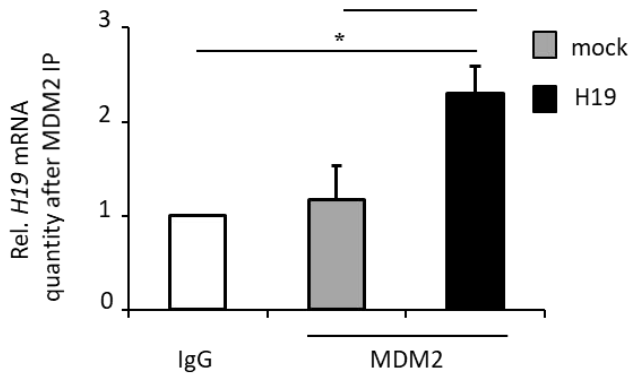
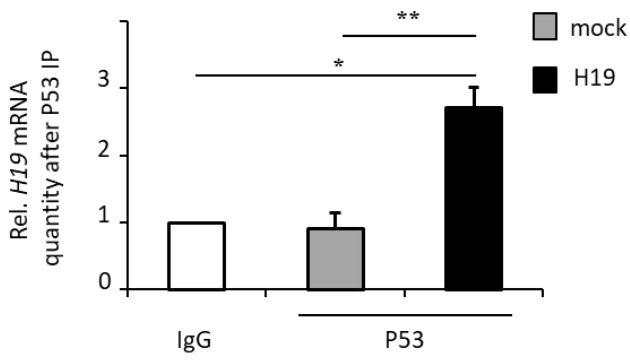
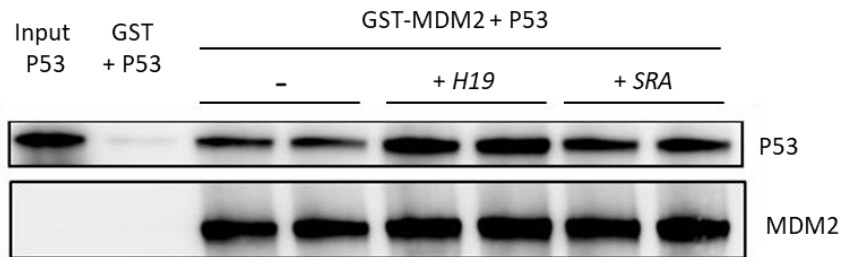
**Availability of data and material.** All materials are fully available on request.

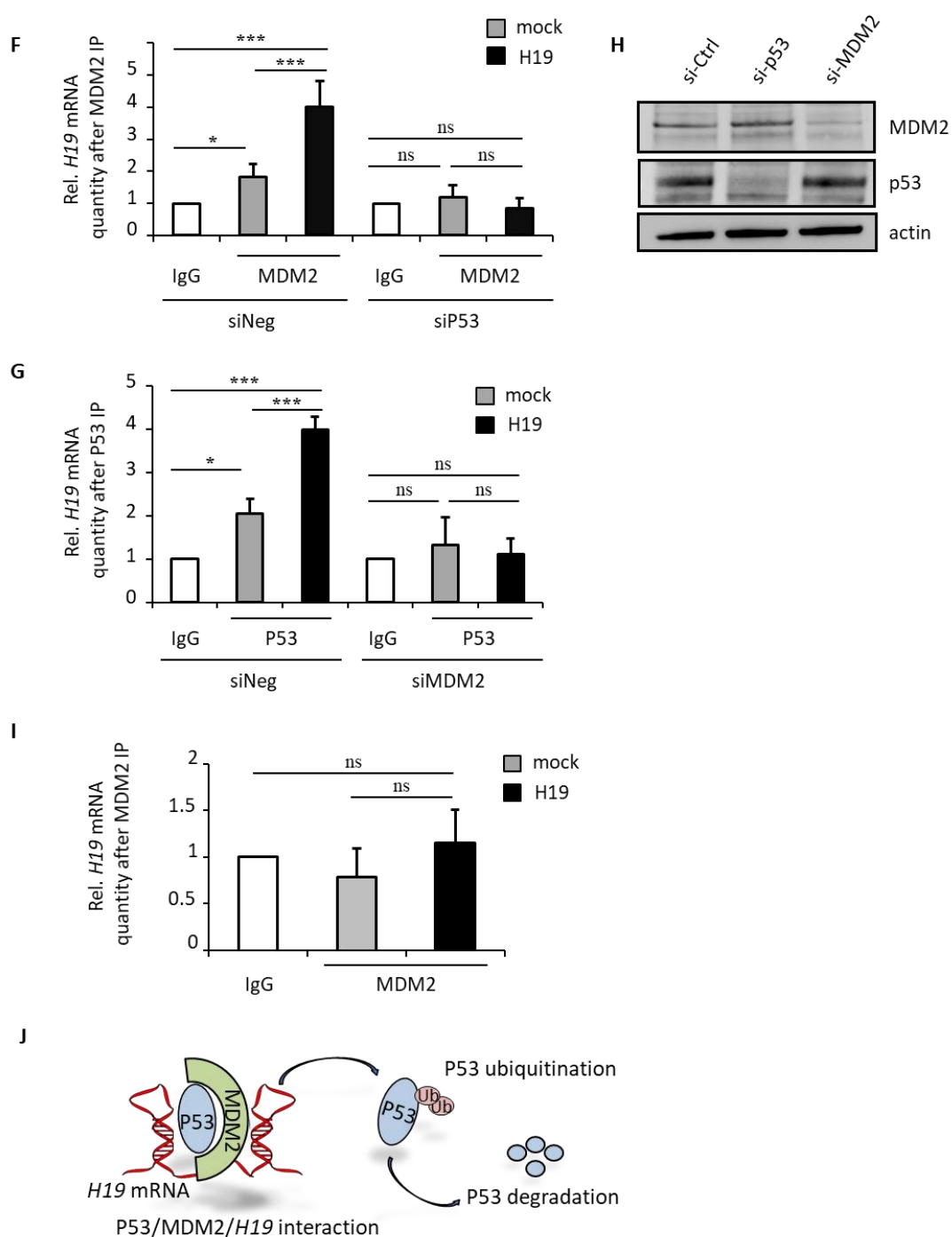
**Authors' contributions.** CV, JC, CL, MG, NS and NM performed experiments. CV, JC, EA, XLB, CA and BL wrote the manuscript. CV, EA, XLB and BL designed the experiments and analyzed the data.



**Figure 1. H19 down-regulated P53 protein expression in breast cancer cells.** **A.** P53 expression in MCF-7, T-47D or MDA-MB-231 control (mock) or stably overexpressing H19 (H19), determined by western blot analysis. Actin was used as equiloading control. **B.** Relative expression of H19, P53, P21, Bax and PUMA RNA in MCF-7 control (mock) or stably overexpressing H19 (H19), determined by qRT-PCR. **C.** Relative expression of H19, P53, P21, Bax and PUMA RNA in MCF-7 transiently transfected with increasing amounts of H19 encoding plasmid. **D.** Relative expression of P53 in P53-null cells (Calu-6) transfected with P53 and/or H19 encoding plasmids. P53 expression was determined by western blot and actin was used as equiloading control. All experiments were performed at least three times. For graphs, data represent mean of three independent experiments and error bar sem. Statistical analyses were performed as described. \* $p < 0.05$ ; \*\* $p < 0.01$ ; ns not significant.

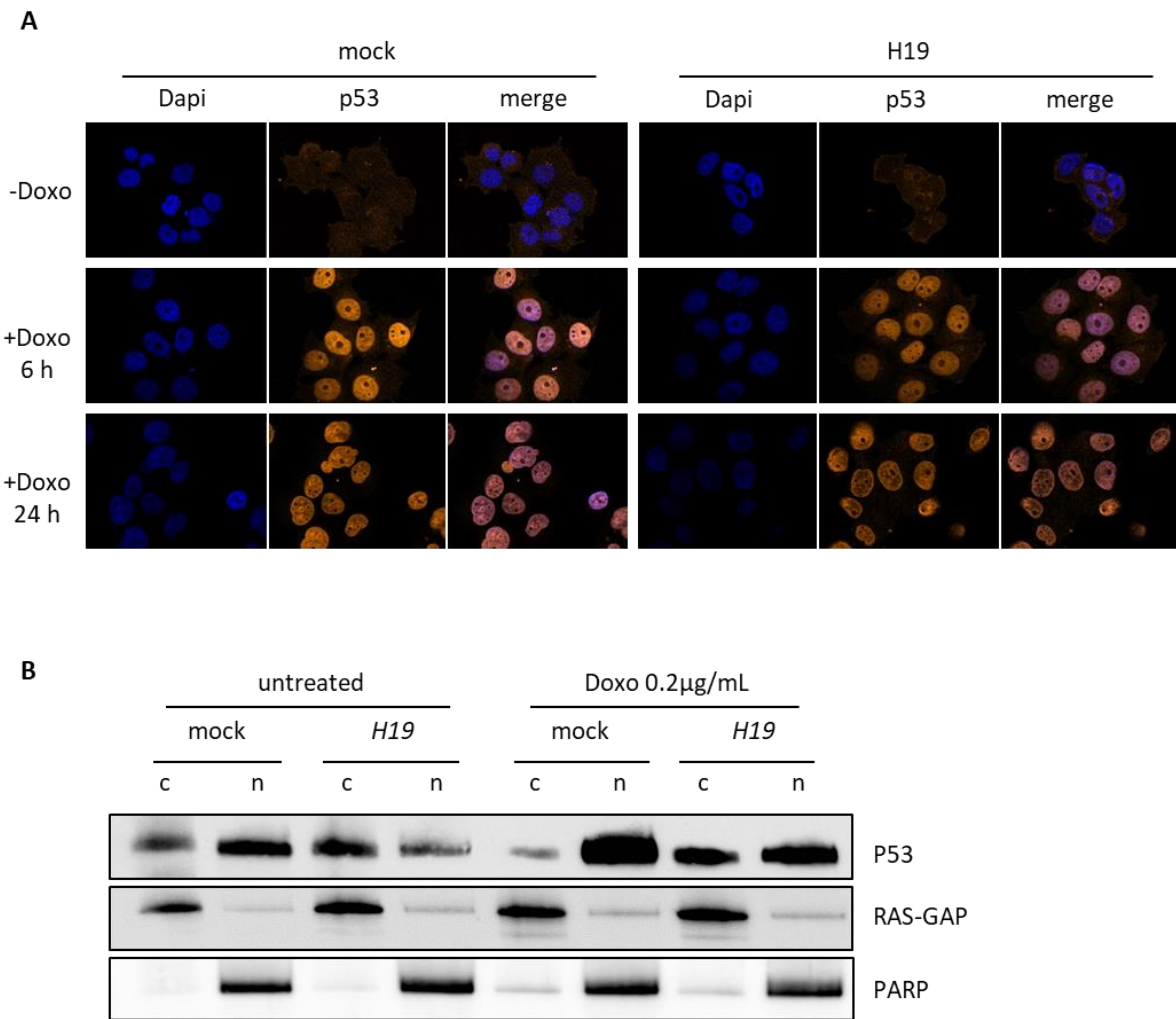


**A****B****C****D****E**

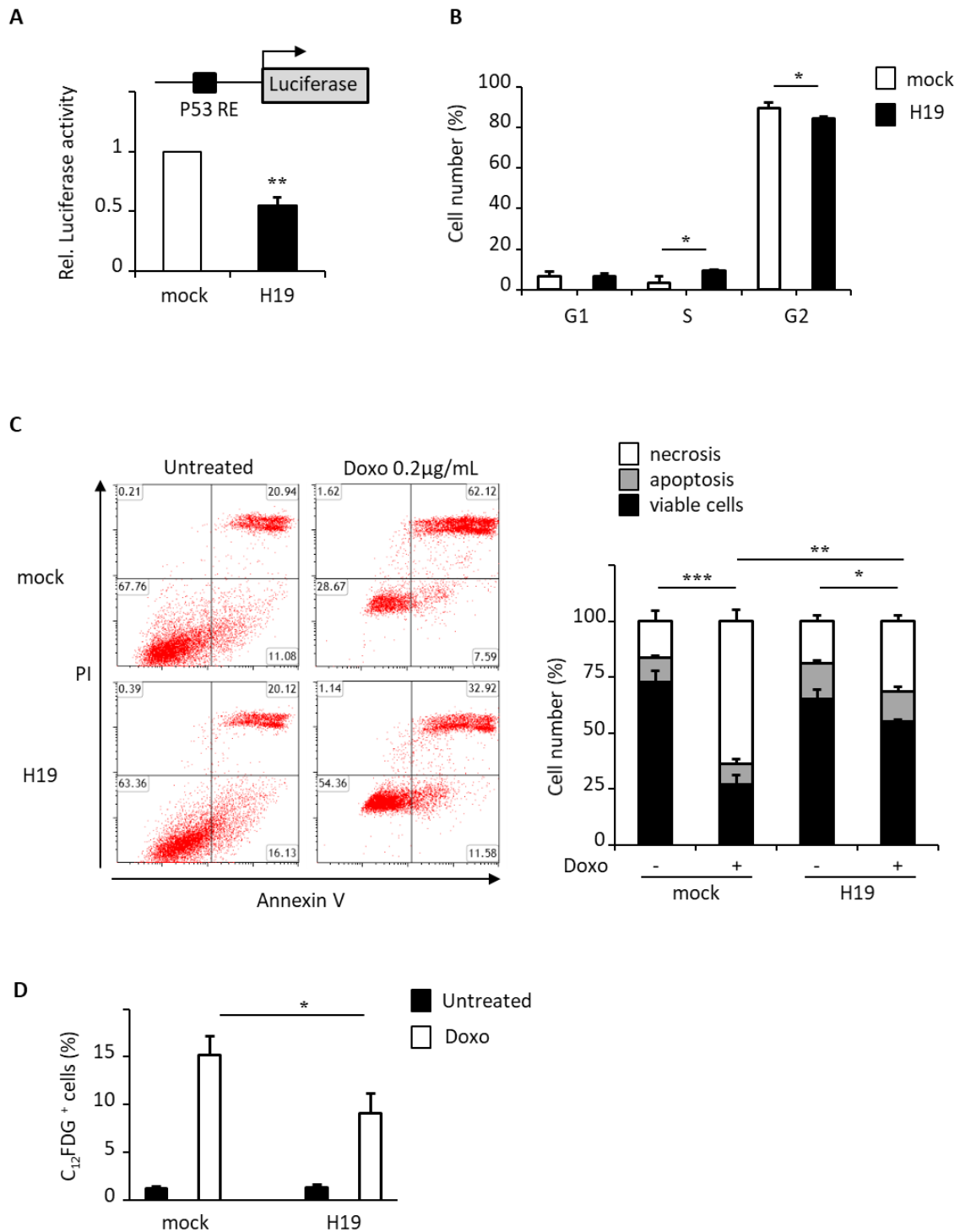


**Figure 2. Trimeric complex formation of H19/P53/MDM2.** **A.** MDM2 expression in MCF-7 control (mock) or overexpressing H19 (H19), determined by western blot. Actin was used as equiloading control. **B.** Co-immunoprecipitation of P53 and MDM2 proteins in T-47D (mock) or overexpressing H19 (H19) treated with MG132 for 6 h. Detection of immunoprecipitated proteins were performed by western blot. **C.** H19 RNA associated to MDM2 proteins were determined by qRT-PCR after MDM2 immunoprecipitation in MCF-7 control (mock) or overexpressing H19 (H19). IgG indexed to 1 was used as control. **D.** H19 RNA associated to P53 proteins were determined by qRT-PCR after P53 immunoprecipitation in breast cancer cells MCF-7 control (mock) or overexpressing H19 (H19). IgG indexed to 1 was used as control. **E.** in vitro P53/MDM2 binding. His-P53 synthesized by in vitro coupled transcription/translation was incubated in the presence of equal amount of GST or GST-MDM2 in the presence or not of H19 RNA. SRA RNA was used as control. Complexes were precipitated with Sepharose-glutathion

beads, resolved by 12% SDS PAGE and visualized by western blot analysis with antibodies against P53 or MDM2. Representative western-blot results are shown. F and G. H19 RNA associated to MDM2 (F) or P53 (G) proteins were determined by qRT-PCR after MDM2 (F) or P53 (G) immunoprecipitation in MCF-7 control (mock) or overexpressing H19 (H19) previously transfected with siRNA control (siCtrl) or siRNA raised against P53 (F) or MDM2 (G). IgG indexed to 1 was used as control. H. p53 and MDM2 expression after siRNA transfection was checked by western-blotting. Actin was used as equiloading control. I. H19 RNA associated to MDM2 proteins were determined by qRT-PCR after MDM2 immunoprecipitation in Calu-6 control (mock) or overexpressing H19 (H19). IgG indexed to 1 was used as control.

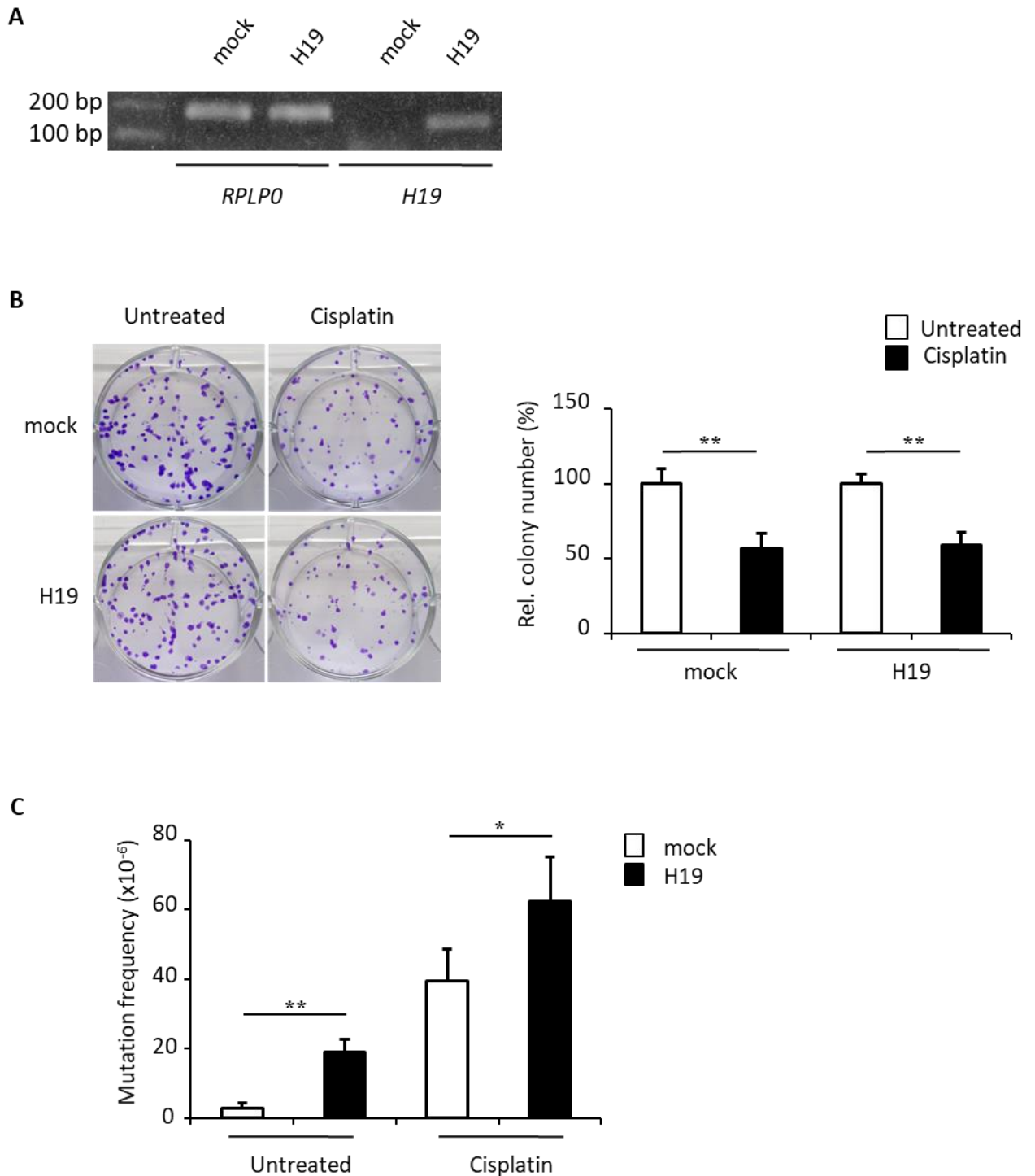


**Figure 3. H19 prevented P53 nuclear translocation upon doxorubicin treatment.** **A.** P53 localization determined by immunofluorescence in MCF-7 control (mock) or overexpressing H19 (H19). Cells were treated (+Doxo) or not (-Doxo) with 0.2µg/ml doxorubicin for 6 h or 24 h. **B.** P53 expression determined by western blot analysis after cell fractionation (c: cytoplasm, n: nucleus) in MCF-7 control (mock) or overexpressing H19 (H19) treated or not with doxorubicin (0.2µg/ml) during 2 h. Ras-GAP and PARP were used as cytosol and nucleus markers respectively. All experiments were performed at least three times.

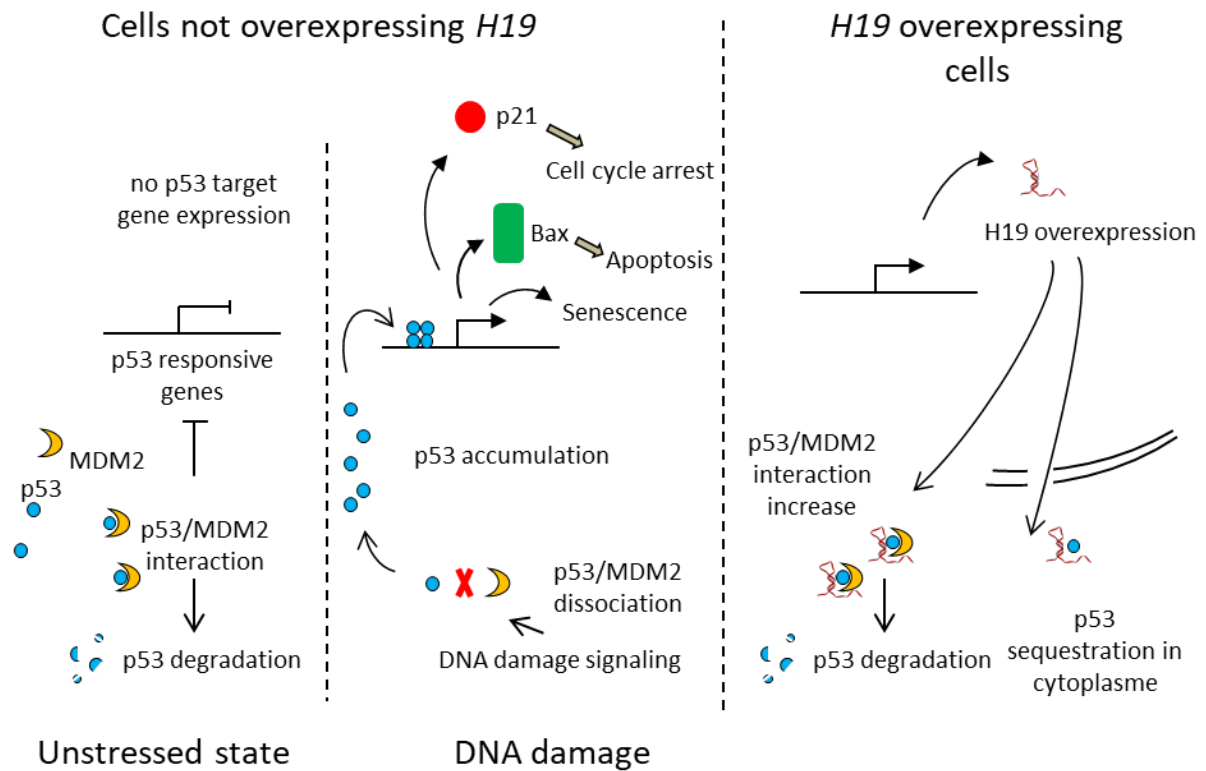


**Figure 4. H19 repressed P53 activities.** A. Relative luciferase activity of P53 reporter plasmid derived of pGL4 luc in MCF-7 control (mock) overexpressing H19 (H19). P53 RE indicates a P53 responsive element. B. Cell cycle analysis of MCF-7 (mock) or overexpressing H19 (H19) treated with doxorubicin (0.2 μM) after synchronization during 24 h with aphidicolin (6 μM). C. Annexin V- propidium iodide (PI) staining and flow cytometric analysis were performed to study the induction of apoptosis after doxorubicin (0.2 μg/ml during 24 h) treatment in MCF-7 control or overexpressing H19. The percentage of Annexin V- and PI-positive or -negative cells is given in each quadrant (representative experiment). Results shown in the graph were the mean of three independent experiments and error bar sem. Statistical analysis was performed on viable populations. D. Percentage of C<sub>12</sub>FDG-positive cells in MCF-7 control or overexpressing H19 treated (or not) with doxorubicin (0.067 μM during 72 h). Results shown in the graph were the

mean of three independent experiments and error bar sem. Statistical analysis was performed as described. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; ns not significant.



**Figure 5. H19 increased mutation frequency measured by HPRT assay.** A. H19 mRNA expression in CHO-K1 control (mock) or stably overexpressing H19 (H19), determined by semi-quantitative RT-PCR. RPLP0 was used as control. B. Effect of H19 overexpression on viability of CHO-K1 cells treated or not with cisplatin (15 $\mu$ M for 1 h) measured by clonogenic assays. Photographs represent a representative experiment and graph the mean of three independent experiments. C. Determination of mutation frequency after 6TG treatment in CHO-K1 control (mock) or stably overexpressing H19 (H19), treated or not by cisplatin (15 $\mu$ M for 1 h). Mutation frequency was measured as described in materials and methods.



**Figure 6. Graphical representation of H19 effects on p53 activities.** In unstressed cells, p53 is constitutively repressed by MDM2. After cellular stress inducing DNA damage, the dissociation of p53/MDM2 complex allows the p53 accumulation and transcriptional effects leading to cell cycle arrest, apoptosis or senescence. In H19 overexpressing cells, H19 strengthens the p53/MDM2 interaction, impairs the p53 nuclear localization and p53 activities.







## II. Le long ARN non codant *H19* et le miR-675 régule la population de cellules souches cancéreuses mammaires.

### Contexte

L'émergence du concept selon lequel des cellules souches cancéreuses seraient associées à la résistance aux traitements et aux récives, impose la recherche de nouveaux acteurs moléculaires permettant de mieux comprendre la biologie de ces cellules particulières. Il est maintenant admis que les longs ARN non codant (lncRNA) exercent des rôles essentiels pour la biologie de la cellule et qu'ils sont impliqués dans l'initiation et/ou le développement de nombreuses pathologies, dont le cancer. Cependant, le rôle des lncRNA dans les cellules souches cancéreuses reste peu connu. Notre équipe a déjà démontré le rôle du lncRNA *H19* dans les cellules cancéreuses mammaires et son implication dans l'émergence et le maintien d'une population de cellules souches prostatiques (Bauderlique-Le Roy et al., 2015). Mes travaux de thèse, initiés en master II, ont permis de montrer que non seulement le lncRNA *H19* est impliqué dans la régulation des cellules souches cancéreuses, mais également le miR-675-5p.

### Introduction

Le rôle oncogénique du lncRNA *H19* et du miR-675 est bien documenté dans de nombreux cancers. Dans les cancers du sein, notre équipe a déjà mis en évidence qu'*H19* est surexprimé dans 73% des cas et qu'il accroît les propriétés tumorigènes des cellules cancéreuses mammaires (Adriaenssens *et al.*, 1998 ; Lottin *et al.*, 2002). En effet, des xénogreffes de cellules MDA-MB-231 surexprimant *H19* dans des souris immunodéficientes induisent la formation de tumeurs qui se développeront plus rapidement et qui auront une taille plus importante que les tumeurs obtenues après l'injection de cellules témoins.

En 2015, notre groupe, en collaboration avec l'équipe du Dr. Roland Bourrette (UMR8161), a mis en évidence un rôle d'*H19* dans l'émergence et le maintien d'une population de cellules souches prostatiques (Bauderlique *et al.*, 2015). Des cellules prostatiques immortalisées, RWPE-1, transfectées avec un vecteur exprimant la protéine fluorescente GFP sous le contrôle du promoteur s-SHIP permettent l'isolement d'une population de cellules souches. Ce promoteur avait été décrit comme étant activé dans différentes populations de cellules souches (Bai & Rohrschneider, 2010 ; Bauderlique *et al.*, 2015). Cette sous-population

enrichie en cellules souches, surexprime *H19*. Le rôle d'*H19* dans la pluripotence de ces cellules souches a été démontré par des approches de surexpression et de silencing. Il a été établi une corrélation positive entre l'expression d'*H19*, *sox2*, *abcg2* et *oct3/4* dans différentes cellules mammaires cancéreuses ou non. Ces résultats suggéraient un rôle du gène *H19* dans la biologie des cellules souches cancéreuses mammaires. Mon travail de M2 a permis de démontrer pour la première fois que la surexpression du lncRNA *H19* augmente certaines capacités associées aux CSC mammaires. Cependant, l'apparition de publications démontrant l'implication du lncRNA *H19* dans la régulation des CSC mammaires au début de ma thèse ont impacté l'originalité de mon travail. Néanmoins, aucune de ces publications ne s'est intéressée au rôle du miR-675 dans les CSC mammaires. Nous avons donc décidé d'étudier l'effet du miR-675 sur les CSC mammaires. De plus, nous savons que de nombreux articles montrent l'implication du lncRNA *H19* dans les cancers et d'autres du miR-675. Cependant, aucun article ne fait la distinction claire entre les mécanismes d'action associés au lncRNA *H19* et les mécanismes associés au miR-675. Notre projet tend donc également à s'orienter sur l'identification des mécanismes de régulation du gène *H19* dépendant et indépendant de son microARN.

## **Matériels et méthodes**

### ***Lignées cellulaires utilisées***

Quatre lignées de cellules épithéliales cancéreuses mammaires sont utilisées. Les lignées MCF-7 et T-47D de type « luminal-like », qui expriment les récepteurs aux oestrogènes (ER) et à la progestérone (PR) mais pas HER2. Les lignées MDA-MB-231 et SUM159PT de type « basal-like » n'expriment pas les récepteurs aux œstrogènes, le récepteur à la progestérone et le récepteur de la famille d'EGFR, HER2. Les lignées T-47D, MDA-MB-231 et MCF-7 proviennent de patientes atteintes d'un adénocarcinome métastatique du sein. La lignée SUM159PT provient d'une tumeur primaire de cancer du sein.

### ***Entretien des lignées cellulaires***

Les cellules MCF-7 sont cultivées dans du Dulbecco's Modified Eagle's Medium (DMEM, Gibco®) supplémenté de 1% de sérum de veau foetal décomplémenté (SVF DC) et de 0,01% de Zell Shield® (Biolabs®). Les cellules T-47D et MDA-MB-231 sont cultivées dans du milieu Roswell Park Memorial Institute (RPMI-1640, Gibco®) supplémenté de 10% de SVF DC et de 0,01% de Zell Shield®. Les cellules SUM159PT sont cultivées dans du milieu F12 (Gibco®) supplémenté

avec 5% de SVF DC, et pour 500mL de milieu : 5 mL d'HEPES (1M, Gibco®), 2 mL de Pénicilline-Streptomycine (Pénicilline :10.000 unités/mL, Streptomycine :10mg/mL Gibco®), 500 µL d'insuline (#I9278, Sigma-Aldrich®), et 27,8µL d'hydrocortisone (50µM, Sigma-Aldrich®). Les cellules sont entretenues dans des flasques de 75cm<sup>2</sup> placées en atmosphère humide à 37°C et à 5% de CO<sub>2</sub>. A pré-confluence, les cellules sont dissociées avec de la Trypsine / 0,25% EDTA (Gibco®) et réensemencées au dixième.

### ***Transfection stable de la lignée SUM159PT***

Les cellules SUM159PT ont été transfectées par nucléofection (Nucleofector™ Amaxa) selon les instructions du fournisseur (Lonza®). La sélection des cellules a été réalisée en présence d'hygromycine B (500µg/mL, Euromedex®). Lorsque les colonies survivantes ont atteint la confluence, les cellules sont dissociées et réensemencées afin d'être amplifiées.

### ***Transfection de siRNA, mimic et antagomir***

Les cellules ont été transfectées avec les siRNAs par lipofection (JetPRIME® Polyplus transfection®) selon les instructions du fournisseur. Les cellules ont été transfectées avec les mimic et antagomir par lipofection (DharmaFECT™, dharmacon®) selon les instructions du fournisseur.

siNEG	SR-CL000-005 Eurogentec®
siH19-a (mix de 2siRNAs )	UAAGGUGUUCAGGAAGGCC AGCUUCACCUUCCAGAGCC
siH19-b	ACGAGGCCAGGUCUCCAGCUG

mimic control	CN-01000-01-05 (Dharmacon)
miR-675-5p mimic	C-301015-03-0020 (Dharmacon)
antagomir control	IN-001005-01-20 (Dharmacon)
antagomir miR-675	IH-301015-04-0020 (Dharmacon)

### ***Formation de mammosphères en milieu sans ancrage***

Les mammosphères sont générées sur des plaques 96 puits de faible adhésion (Corning® Costar®). Le milieu de culture utilisé est du milieu DMEM/F12 (Gibco®) supplémenté, pour 500mL de milieu, avec 50µL d'EGF (200µg/mL, Sigma-Aldrich®), 400µL de bFGF (25µg/mL, Sigma-Aldrich®), 5mL de BSA 30% (Sigma-Aldrich®) et 10 mL de supplément B27® (50X, Gibco®). Cent µL de milieu sont ajoutés dans les puits. Les cellules sont dissociées à la trypsine supplémentée par 0,25% d'EDTA et comptées. 20.480 cellules sont placées dans un Eppendorf, du milieu est ajouté pour atteindre un volume final de 1mL. 100µL de cellules sont ajoutés dans les puits de la première colonne. Après 5 aspirations-refoulements, 100µL du mélange de

la première colonne sont ajoutés à la deuxième colonne *via* une pipette multicanaux. L'opération est répétée jusqu'à la 11<sup>ème</sup> colonne. Les plaques sont placées dans l'incubateur et les sphères comptées après 5 jours d'incubation. Le pourcentage de mammosphères est estimé avec la formule suivante : (nombre de sphères / nombres de cellulesensemencées) x 100.

### ***Formation de colonies***

Les cellules sont dissociées à la trypsine et comptées. 100 ou 200 cellules sontensemencées en tripliquât dans des boites 6 puits. Après 10 jours, les cellules sont rincées au PBS puis fixées au paraformaldéhyde 4%. Les colonies sont ensuite colorées au crystal violet (0,5%) puis comptées. Le pourcentage de cellules capables de former des colonies est estimé avec la formule suivante : (nombre de colonies / nombres de cellulesensemencées) x 100.

### ***Cytométrie en flux et tri cellulaire***

Les analyses de cytométrie en flux et les tris cellulaires ont été réalisées à la plateforme de cytométrie de l'IRCL et de l'IBL (BioImaging Center of Lille). La protéine fluorescente mCherry est excitée à 561nm. Les analyses sont effectuées sur un cytomètre en flux BD LSRFortesa™ X-20 et les tris sont réalisés *via* le trieur BD FACSAria™ II (Special Order Research Products).

### **Mesure de l'invasion et de la migration cellulaire**

Lors des tests de migration par Transwell, 40.000 cellules sontensemencées dans des inserts de culture (0,8µm, BD Bioscience) préalablement recouverts par du collagène de type I (0,045mg/mL, millipore), dans des plaques 12puits. Après 8h, les cellules sont fixées au méthanol, cellules qui ont migré et envahi l'autre côté du filtre sont colorées au Hoescht 33258 puis comptées. Les tests d'invasion par Transwell sont réalisés avec les inserts de culture recouverts par du collagène de type I (3mg/mL) et les cellules sont fixées après 24h.

### ***Tests statistiques***

Le test statistique utilisé pour analyser les résultats est un t-test paramétrique, non apparié. A l'exception de la figure 3 qui utilise un test d'analyse de variance à deux critères de classification (Two-way ANOVA). La significativité des résultats est représentée de la manière suivante : \* : p-value < 0,05 ; \*\* : p-value <0,01 ; \*\*\* : p-value <0,001

## Résultats

### La surexpression du gène *H19* est corrélée avec les caractéristiques souches

Une méta-analyse sur des données de micro-array a été réalisée en collaboration avec le Pr Bertucci (Institut Paoli-Calmettes, Marseille). Dans cette étude, l'expression du gène *H19* a été recherchée dans plus de 5000 tumeurs du sein et comparée entre les tumeurs exprimant ou non des signatures géniques associées aux cellules souches cancéreuses mammaires. Les marqueurs utilisés étaient l'expression d'une signature génique ALDH1a1 ou la signature génique associée à la combinaison entre la surexpression de CD44 et la sous expression de CD24. Les tumeurs ont été classées en *H19*<sup>high</sup> et *H19*<sup>low</sup> en fonction de l'expression de *H19* supérieure ou inférieure respectivement à la médiane d'expression (Figure 21, page 100). Nous observons une plus forte expression d'*H19* dans les tumeurs exprimant les signatures associées aux CSC mammaires, ce qui peut laisser penser que le lncRNA *H19* pourrait avoir un rôle dans les cellules souches cancéreuses mammaires. Pour étudier cette possible régulation, nous avons étudié les capacités fonctionnelles associées aux cellules souches cancéreuses (formation de mammosphères en milieu sans ancrage et formation de colonie en milieu avec ancrage) dans des cellules MCF-7 et MDA-MB-231 qui surexpriment stablement le gène *H19*. Nos résultats montrent que ces deux propriétés fonctionnelles associées aux cellules souches sont plus importantes lorsque les cellules surexpriment le gène *H19*. Ces résultats montrent une corrélation entre les caractéristiques associées aux CSC mammaires et la surexpression du gène *H19*, à la fois dans les tumeurs *in vivo* mais également lorsque *H19* est surexprimé de manière ectopique dans les cellules *in vitro*.

### L'activité promotrice du gène *H19* permet d'isoler une population enrichie en cellules souches cancéreuses mammaires

Nous avons utilisé une lignée cellulaire SUM159PT transfectée stablement avec un vecteur permettant l'expression d'une protéine fluorescence mCherry exprimée sous le contrôle d'une séquence promotrice du gène *H19* de 1000pb (Figure 22, page 101). Cette lignée cellulaire nous permet de visualiser l'expression du gène *H19* et d'étudier les propriétés souches de la population qui exprime plus fortement le lncRNA *H19*, sans modifier de manière ectopique le niveau d'expression de *H19*. Nous avons trié les cellules exprimant ou non la protéine fluorescente mCherry et étudié leurs capacités à former des mammosphères et des

colonies. Nous observons que les cellules exprimant la protéine fluorescence mCherry ont une capacité à former des mammosphères et des colonies accrues.

**Ce résultat suggère que les cellules présentant une activité du promoteur du gène *H19* plus importantes sont des cellules avec un phénotype souche plus important.**

### ***H19* est nécessaire au maintien des propriétés souches des CSC**

Pour étudier le rôle de *H19* dans la régulation des CSC mammaires, nous avons également inhibé son expression avec deux siRNAs différents (Figure 23, page 102). La capacité des cellules à former des mammosphères et des colonies est altérée lorsque les cellules ont subi une perte d'expression d'*H19*, comparé au contrôle. La lignée MCF-7 ne possédant pas d'activité ALDH, nous avons réalisé le marquage de la population ALDH<sup>high</sup> uniquement dans la lignée SUM159PT. La répression d'*H19* diminue le pourcentage de population ALDH<sup>high</sup>. **Ces résultats suggèrent que l'expression du lncRNA *H19* est nécessaire au maintien des propriétés souches des CSC mammaires.** Cependant, à ce stade nous ne pouvons pas différencier l'implication du lncRNA *H19* de son microARN. En effet, la modulation d'*H19* via la surexpression ectopique du gène *H19* ou l'utilisation d'un siRNA entraîne une modulation de l'expression du lncRNA *H19* ainsi que du miR-675.

### **Le miR-675-5p est nécessaire au maintien des propriétés souche des CSC mammaires**

Pour étudier le rôle du miR-675-5p dans la régulation des CSC mammaires, nous avons utilisé un mimic du miR-675 pour mimer une surexpression de ce microARN et un antagomiR qui agit comme un antagoniste et bloque l'action de ce microARN (Figure 24, page 103).

Nous avons réalisé des expériences de formation de mammosphères en milieu sans ancrage, de formation de colonies en milieu avec ancrage et marqué la population de cellules ALDH<sup>high</sup>. L'utilisation d'un mimic du miR-675 induit une augmentation de capacité des cellules à former des mammosphères pour la lignée MCF-7 et de leurs capacités à former des colonies pour les lignées MCF-7 et SUM159PT. La répression du miR-675 par un antagomir diminue la capacité des cellules à former des mammosphères pour les lignées et la capacité des cellules à former des colonies uniquement pour les cellules MCF-7. Le pourcentage de cellules ALDH<sup>high</sup> est également régulé par le miR-675-5p. Lorsque les cellules sont transfectées avec le mimic, la population de cellules ALDH<sup>high</sup> est plus importante, alors que lorsque ce microARN est bloqué par l'antagomir, cette population est diminuée. **L'ensemble de ces**

**résultats suggèrent que la modulation du miR-675-5p a un impact sur les propriétés souches des CSC mammaires. La surexpression de ce microARN augmente leurs capacités alors que sa répression impacte le maintien de cette population souche.**

**La surexpression du gène *H19* et du miR augmente les capacités de migration et d'invasion des cellules cancéreuses mammaires.**

De nombreuses études ont fait un lien entre les CSC et la transition épithélio-mésenchymateuse (Shibue & Robert A. Weinberg, 2017). Nous avons donc décidé d'étudier l'impact d'*H19* et du miR-675 sur les capacités de migration et d'invasion des cellules cancéreuses mammaires. Nos résultats montrent que la surexpression du gène *H19* permet d'augmenter les capacités migratoires des cellules MCF-7 et MDA-MB-231, ainsi que leurs capacités d'invasion (Figure 25, page 104). La surexpression du miR-675 montre également une augmentation de leurs capacités d'invasion et de migration, cependant, dans les cellules MDA-MB-231 la surexpression du miR-675 augmente significativement moins fortement la capacité migratoire des cellules. Ce résultat pourrait indiquer que le miR-675 est bien impliqué dans l'accroissement des capacités de migration des cellules lorsque le gène *H19* est surexprimé mais qu'il y a également des mécanismes miR-675 indépendant qui régule ce processus.

## **Discussion**

Nos résultats ont montré que la modulation d'expression du gène *H19* régule les phénotypes associés aux cellules souches cancéreuses. Ces résultats confirment les observations préalablement réalisées par plusieurs études (Peng et al., 2017; Shima et al., 2018). En effet il a été démontré que le silencing du gène *H19* par siRNA diminue les capacités fonctionnelles associées aux cellules souches. Peng et collaborateurs ont notamment démontré que le lncRNA *H19* sert d'éponge aux microARN de la famille let-7 pour accroître les capacités souches des CSC. En effet, les microARN de la famille let-7 sont considérés comme des suppresseurs de tumeurs et ciblent notamment le facteur de transcription LIN28 qui est un régulateur des propriétés souches. La surexpression du gène *H19* permet donc de rétablir l'expression de LIN28, en servant d'éponge à let-7, et d'accroître la population de CSC mammaires. De manière intéressante, il a été montré que cette réexpression de LIN28 permet

en retour l'augmentation d'expression du lncRNA *H19*, via la régulation du microARN let-7 par LIN28.

Nous avons également démontré que l'expression d'*H19* dans les tumeurs de patientes est corrélée avec une signature d'expression génique associées à deux profils de CSC mammaires (*Aldh1a1* et *CD44+/CD24-*). Ce résultat permet de montrer la pertinence *in vivo* de nos résultats *in vitro* sur la régulation des CSC mammaires par *H19*. Shima et collaborateurs ont également démontré que les tumeurs qui exprimaient plus fortement le gène *Aldh1a1* sont des tumeurs qui surexpriment le gène *H19* (Shima et al., 2018).

L'utilisation de mimic et d'antagomir du miR-675-5p nous a permis de démontrer son implication dans la régulation des CSC mammaires. En effet, la surexpression artificielle du miR-675 via un mimic augmente les propriétés souches des cellules et son ciblage par antagomir permet de diminuer ces capacités. Nous sommes les premiers à mettre à jour l'implication du miR-675 dans les CSC mammaires. Ce résultat nous laisse suggérer qu'une partie des effets de la surexpression du gène *H19* sur les CSC mammaires est due à des effets miR-675 indépendant, comme la régulation de let-7 par exemple, mais qu'une autre partie est due au miR-675. Les cibles du miR-675 impliquées restent à identifier.

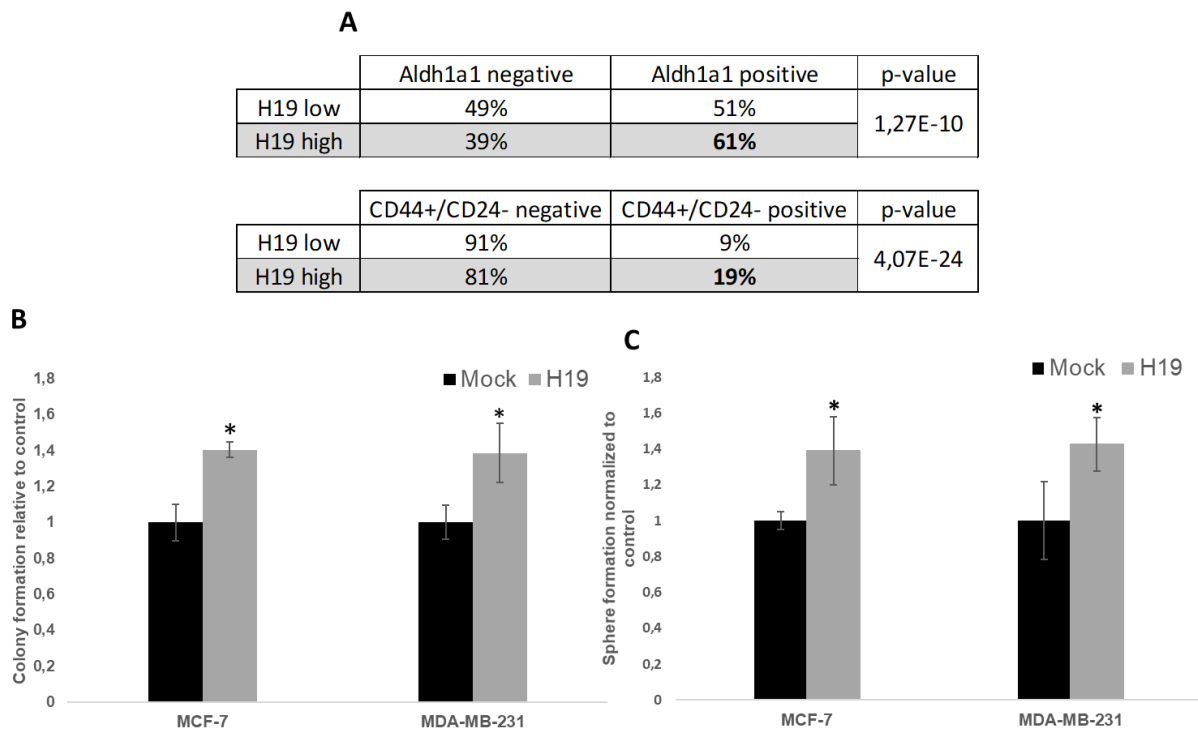
Nous avons également commencé à étudier l'impact du lncRNA *H19* et de son microARN sur l'EMT. Les résultats nous ont montré que la surexpression du lncRNA *H19* permet d'accroître les capacités migratoires et d'invasion des cellules cancéreuses mammaires via des mécanismes dépendants et indépendants de son microARN. Bien que ces résultats soient intéressants, nous avons besoin d'expériences supplémentaires pour conclure sur l'implication d'*H19* et de son microARN sur l'EMT. Nous allons notamment mesurer l'expression des protéines impliquées dans l'EMT dans nos différentes cellules. Ces expériences seront prises en charge par Evodie lors de sa thèse. Des expériences de xénogreffes en modèle poisson zèbre sont également en cours de réalisation au moment où je termine l'écriture de ma thèse. Nous allons également développer une lignée cellulaire exprimant de manière endogène un gène *H19* dépourvue du miR-675 afin de pouvoir discerner au mieux la part relative de chacun dans nos expériences. Cette stratégie sera détaillée dans la discussion de cette thèse.

Une des propriétés importantes des CSC mammaires est la capacité de ces cellules à résister aux thérapies anticancéreuses. La surexpression du gène *H19* est responsable de la résistance des cellules cancéreuses à diverses thérapies (Chen et al., 2017; Gao et al., 2017 ; Si et al., 2016). Par exemple, dans les cellules cancéreuses mammaires MCF-7, il a été décrit un

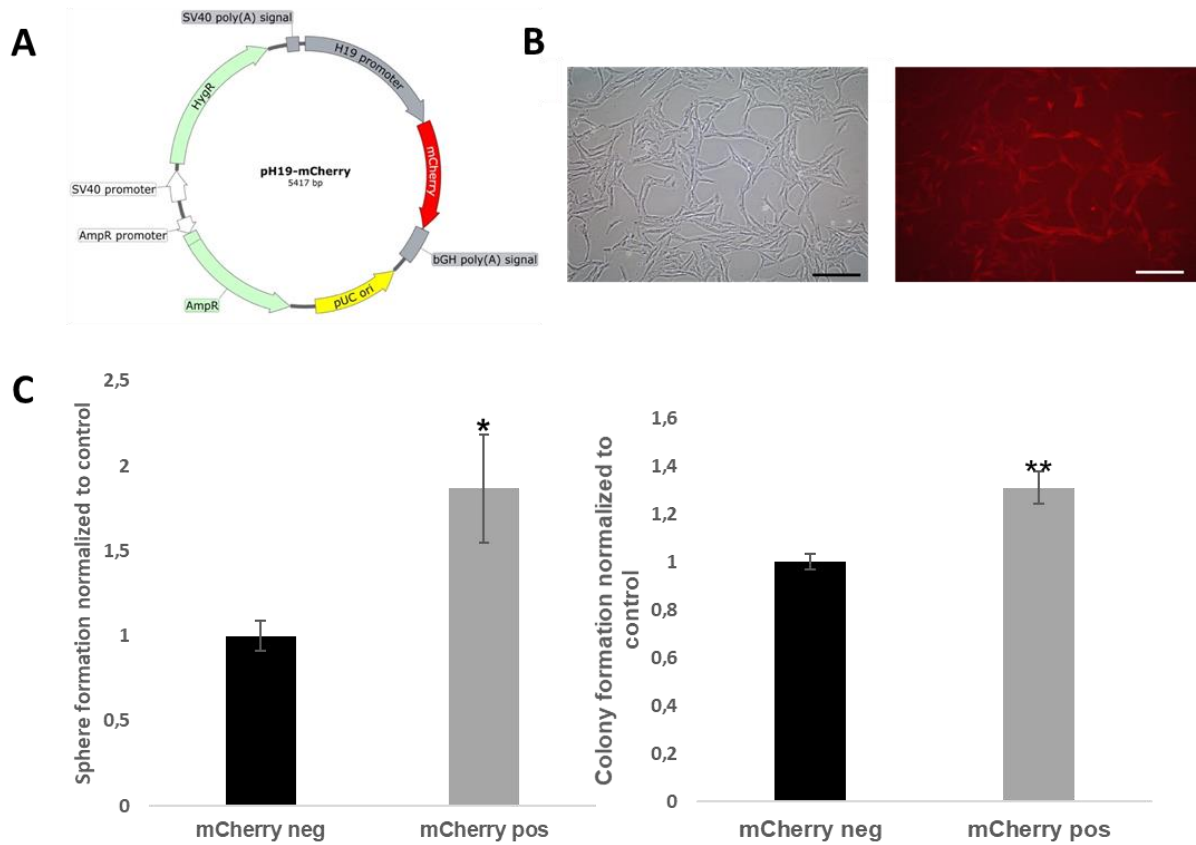


mécanisme de résistance au paclitaxel faisant intervenir l'interaction entre le lncRNA *H19* et la protéine de modification de la chromatine EZH2 (Si et al., 2016). Cette interaction permet le ciblage du complexe PRC2 au niveau de promoteurs cibles et d'induire une chimiorésistance. Cependant, nous ne savons pas si les chimiorésistances observées dans les cellules qui surexpriment le gène *H19* ne sont pas également en partie, dues au miR-675. La suite de ce projet va donc s'intéresser à l'implication du lncRNA *H19* et du miR-675 dans la résistance à certaines molécules utilisées dans le traitement des cancers du sein. De plus, bien que l'implication du lncRNA *H19* dans la résistance à de nombreuses drogues est bien décrite, seul un article montre un mécanisme de résistance associé au miR-675 (Chen et al., 2017). Dans les cellules cancéreuses du colon, le miR-675 cible la protéine VDR (récepteur de la vitamine D) et permet la résistance au 1,25(OH)<sub>2</sub>D<sub>3</sub> (calcitriol, utilisée en thérapie pour les cancers du colon).

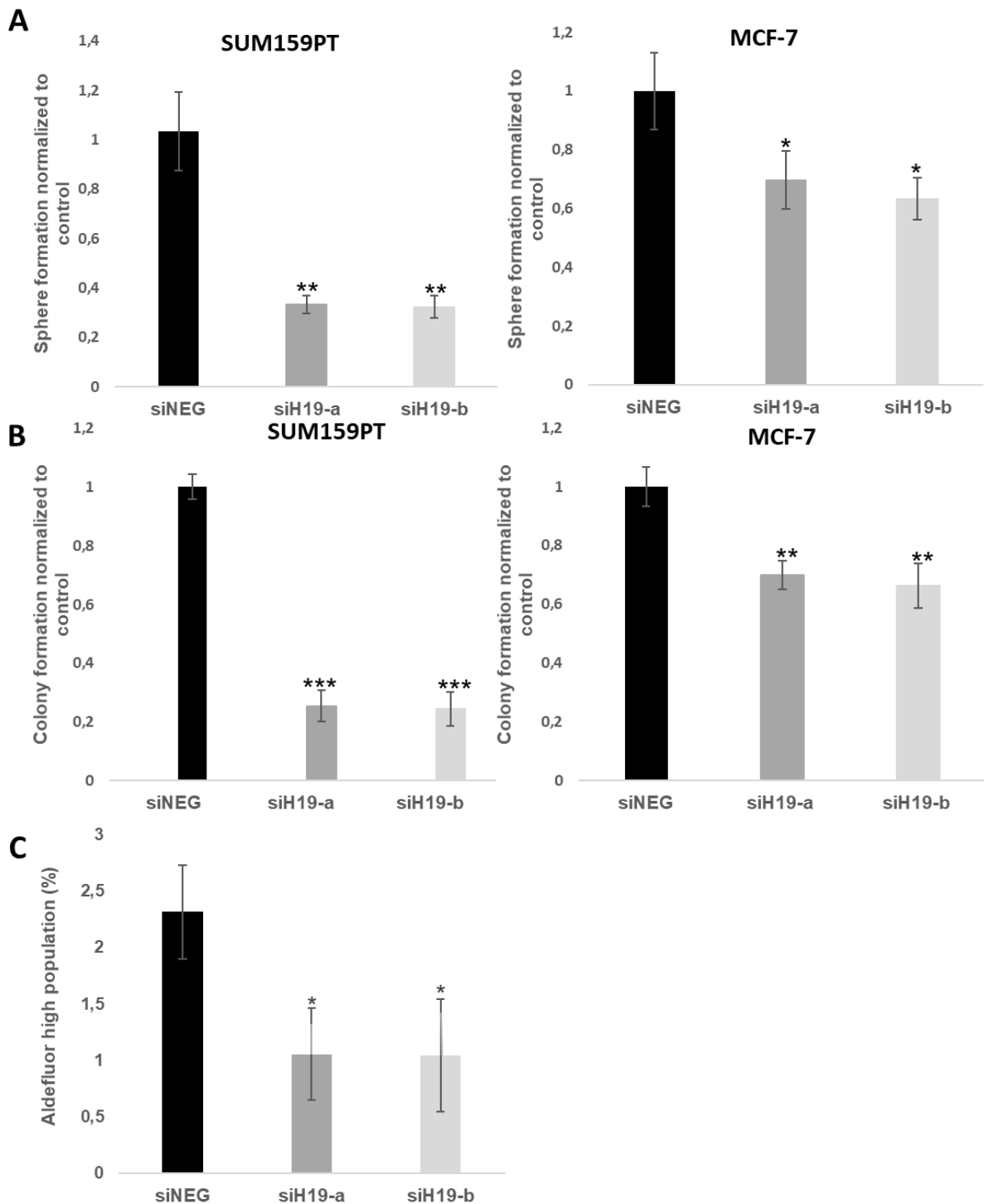
**Ces travaux ont permis la découverte de l'implication du miR-675 dans la régulation des CSC mammaires. La poursuite de ce projet va permettre une meilleure discrimination des implications relatives du lncRNA *H19* et du miR-675 dans les différents processus impliqués dans la tumorigenèse mammaire.**



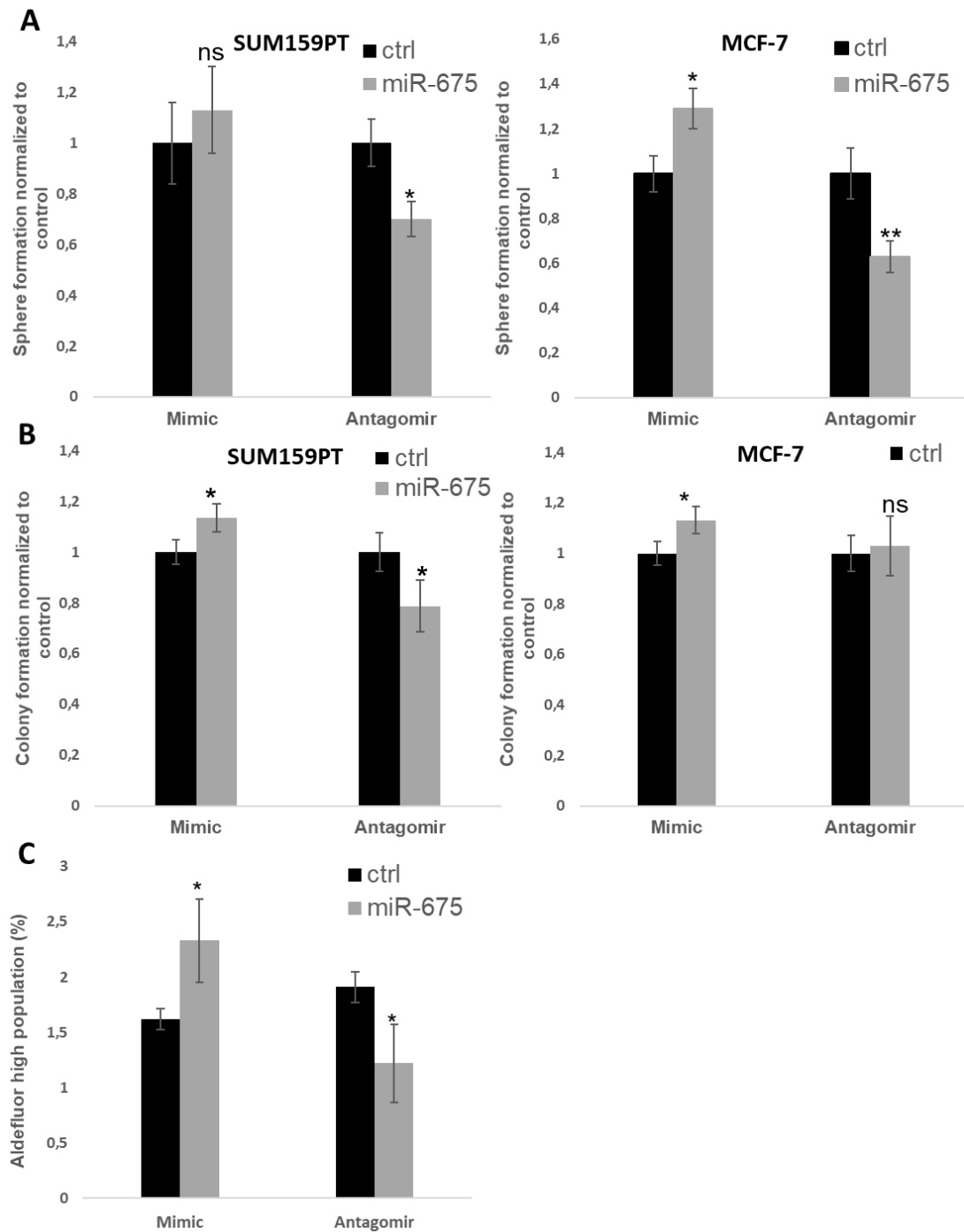
**Figure 21 : La surexpression du gène *H19* augmente les capacités souches des cellules.** A. Le tableau représente le pourcentage et le nombre de tumeurs exprimant (pos) ou non (neg) les signatures géniques associés aux marqueurs de cellules souches ALDH1a1 ou la combinaison CD44+/CD24- en fonction de l'expression du gène *H19*. B. Mesure de la capacité des cellules à former des colonies dans des cellules MCF-7 et MDA-MB-231 surexprimant stablement le gène *H19*. C. Mesure de la capacité des cellules à former des mammosphères dans des cellules MCF-7 et MDA-MB-231 surexprimant stablement le gène *H19*. Les tests statistiques ont été réalisés comme décrit, \* $p < 0.5$  ; \*\* $p < 0.01$ , \*\*\* $p < 0.001$



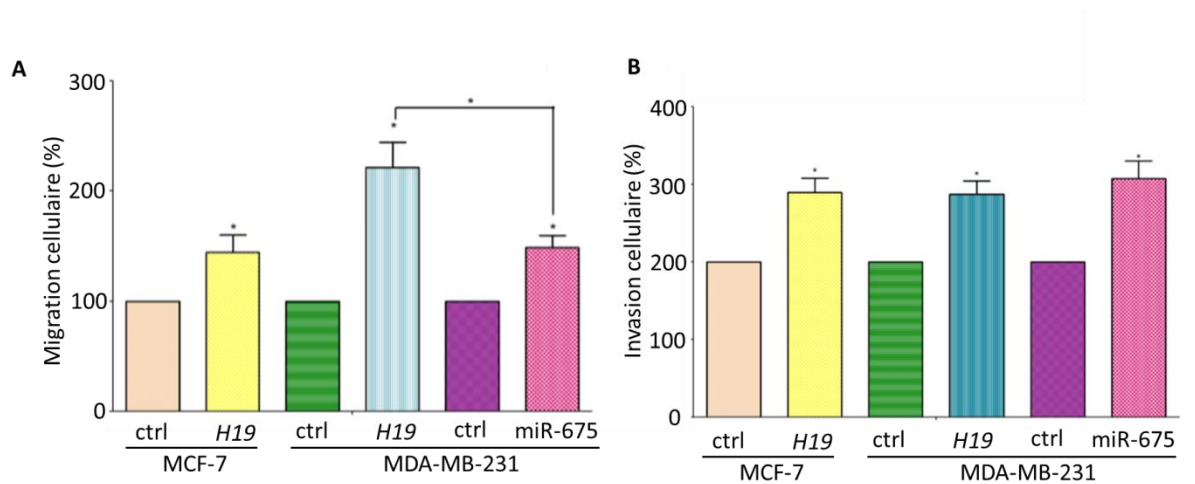
**Figure 22 : L'activité du promoteur du gène *H19* permet d'identifier une population enrichie en cellules souches cancéreuses mammaires.** A. Carte du vecteur pH19-mCherry. B. Photographie des cellules en lumière visible et en fluorescence *via* un filtre PE-Texas Red, barre : 200 $\mu$ m. C. Mesure de la capacité des cellules à former des mammosphères et des colonies, après un tri sur la base de l'expression du marqueur fluorescent mCherry.



**Figure 23 : Le lncRNA *H19* est nécessaire au maintien des cellules souches cancéreuses mammaires. A.** Mesure de la capacité des cellules à former des mammosphères après traitement avec deux siRNA ciblant *H19*, dans des cellules SUM159PT et MCF-7. **B.** Mesure de la capacité des cellules à former des colonies après traitement avec deux siRNA ciblant *H19*, dans des cellules SUM159PT et MCF-7. **C.** Observation par cytométrie en flux du marquage de la population ALDH<sup>high</sup>, après traitement avec deux siRNA ciblant *H19*, dans des cellules SUM159PT. Les tests statistiques ont été réalisés comme décrit, \* $p < 0.5$  ; \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .



**Figure 24 : Le miR-675-5p est nécessaire au maintien des cellules souches cancéreuses mammaires.** A. Mesure de la capacité des cellules à former des mammosphères après traitement avec un mimic du miR-675 ou un antagomir, dans des cellules SUM159PT et MCF-7. B. Mesure de la capacité des cellules à former des colonies après traitement avec un mimic du miR-675 ou un antagomir, dans des cellules SUM159PT et MCF-7. C. Observation par cytométrie en flux du marquage de la population ALDH<sup>high</sup>, après traitement avec un mimic du miR-675 ou un antagomir, dans des cellules SUM159PT. Les tests statistiques ont été réalisés comme décrit, \* $p < 0.5$  ; \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .



**Figure 25 : La surexpression du lncRNA *H19* et du miR-675-5p accroît les capacités de migration et d'invasion des cellules cancéreuses mammaires.** A. Mesure de la capacité migratoire de cellules MCF-7 surexprimant *H19* ou de cellules MDA-MB-231 surexprimant *H19* ou le miR-675. Cette mesure a été réalisée lors d'expériences de transwell. B. Mesure de la capacité d'invasion de cellules MCF-7 surexprimant *H19* ou de cellules MDA-MB-231 surexprimant *H19* ou le miR-675. Cette mesure a été réalisé lors d'expériences avec des transwell ayant subi un coating au collagène. Les tests statistiques ont été réalisés comme décrit, \* $p < 0.5$  ; \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .








### III. Mise en évidence des gènes cibles du lncRNA *H19* et du miR-675

#### III.1. Stratégie utilisée pour la mise en évidence des gènes cibles du lncRNA *H19* et de son microARN

Nous disposons de nombreux modèles cellulaires où l'expression d'*H19* ou de son microARN est modulée, ce qui nous a permis de déterminer leurs implications dans des phénotypes associés au cancer du sein. La détermination des cibles et voies moléculaires associées à ces phénotypes est cependant plus complexe à réaliser. Par exemple, les cibles du miR-675 publiées par le laboratoire (c-Cbl et Cbl-b) ont été trouvées *in silico* dans des banques de données permettant d'identifier des sites de fixations d'un miRNA d'intérêt sur l'ensemble des gènes humains (Targetscan par exemple). Cependant, de nombreuses bases de données existent et un gène cible putatif retrouvé *in silico* dans une banque peut être présent dans une deuxième banque et absente dans une troisième en fonction des algorithmes utilisés. De plus, pour un microARN donné, il peut y avoir plusieurs milliers de gènes cibles putatifs trouvés sur une base de données (Figure 26).



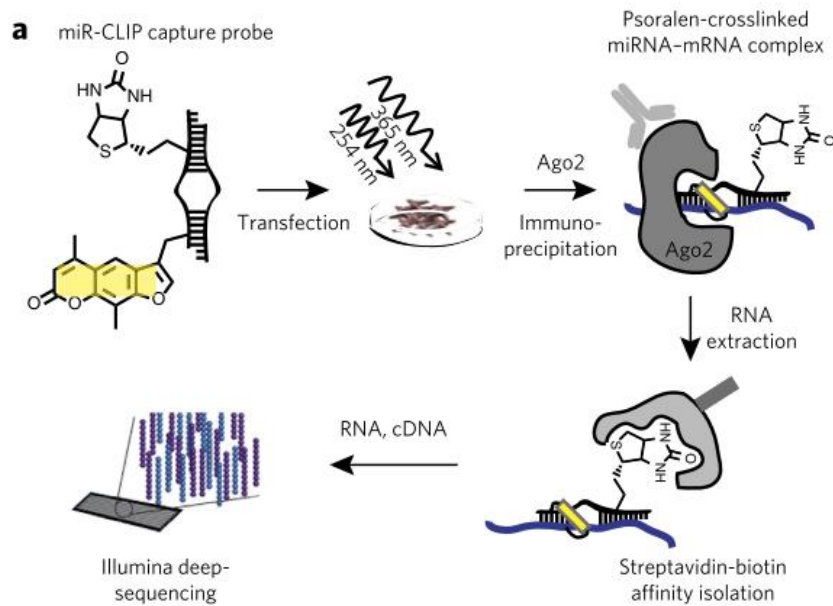
**Human | miR-675-5p/4466**  
 1342 transcripts with sites, containing a total of 1497 sites.  
 Please note that these predicted targets are primarily false positives. [\[Read more\]](#)  
 Table sorted by cumulative weighted context++ score  
 The table shows at most one transcript per gene, selected for being the most prevalent, based on 3P-seq tags (or the one with the longest 3' UTR, in case of a tie). [\[Download table\]](#)

Target gene	Representative transcript	Gene name	Number of 3P-seq tags supporting UTR + 5	Link to sites in UTRs	Site counts				Representative miRNA	
					total	8mer	7mer-m8	7mer-A1		6mer sites
NNAT	ENST00000062104.2	neuronatin	5	Sites in UTR	2	1	0	1	0	hsa-miR-675-5p
ICAM3	ENST00000160262.5	intercellular adhesion molecule 3	873	Sites in UTR	1	1	0	0	0	hsa-miR-675-5p
CAMK2B	ENST00000502837.2	calcium/calmodulin-dependent protein kinase II beta	11	Sites in UTR	2	2	0	0	0	hsa-miR-675-5p
FBXL16	ENST00000397621.1	F-box and leucine-rich repeat protein 16	5	Sites in UTR	2	1	1	0	0	hsa-miR-675-5p
GRAMD1A	ENST00000424536.2	GRAM domain containing 1A	121	Sites in UTR	1	1	0	0	0	hsa-miR-675-5p
LTBP4	ENST00000243562.9	latent transforming growth factor beta binding protein 4	32	Sites in UTR	2	1	0	1	1	hsa-miR-675-5p
MT-ND5	ENST00000361567.2	mitochondrially encoded NADH dehydrogenase 5	675	Sites in UTR	1	1	0	0	0	hsa-miR-675-5p
PHLDA1	ENST00000602540.1	pleckstrin homology-like domain, family A, member 1	3907	Sites in UTR	1	1	0	0	0	hsa-miR-675-5p
CDKN2A	ENST00000498124.1	cyclin-dependent kinase inhibitor 2A	354	Sites in UTR	1	1	0	0	0	hsa-miR-675-5p
G6PC3	ENST00000269097.4	glucose 6 phosphatase, catalytic, 3	429	Sites in UTR	1	1	0	0	0	hsa-miR-675-5p
RNASE7	ENST00000298690.4	ribonuclease, RNase A family, 7	5	Sites in UTR	1	1	0	0	0	hsa-miR-675-5p
PIP5K1B	ENST00000265382.3	phosphatidylinositol-4-phosphate 5-kinase, type I, beta	69	Sites in UTR	1	1	0	0	0	hsa-miR-675-5p
ZNF33A	ENST00000469037.2	zinc finger protein 33A	39	Sites in UTR	2	1	1	0	0	hsa-miR-675-5p
AP000867.1	ENST00000343767.3		5	Sites in UTR	1	1	0	0	0	hsa-miR-675-5p
ZNF446	ENST00000335841.4	zinc finger protein 446	44	Sites in UTR	2	1	1	0	0	hsa-miR-675-5p

**Figure 26 : Identification de 1342 cibles du miR-675-5p *in silico* dans la base de données TargetScan**

Identifier de nouvelles cibles du miR-675 par criblage des banques de données et tester aléatoirement (par qRT-PCR) les gènes cibles putatifs intéressants n'est donc pas pertinent. Il est également possible de vérifier des gènes cibles identifiés dans d'autres cancers ou pathologies afin de déterminer si *H19* régule ces gènes dans nos cellules cancéreuses mammaires. Cependant, étudier une cible déjà connue manque d'originalité et reste risqué

puisque d'autres équipes sont susceptibles de s'y être déjà intéressé dans notre modèle d'étude. Afin de déterminer sans *a priori* les gènes cibles du miR-675 il est possible de réaliser des techniques de miR-CLIP suivi d'un RNA-séq (Figure 27).



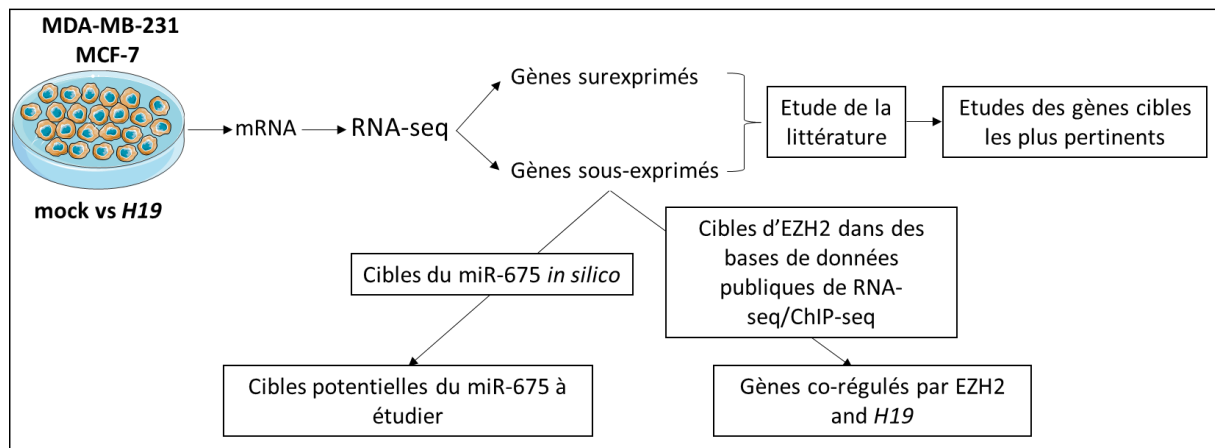
**Figure 27 : Représentation graphique du miR-CLIP**

Une sonde à miR portant un groupement psoralène et biotine sont transférées dans des cellules HeLa. Les cellules sont irradiées pour induire un crosslink entre les miRNA et les ARNm. Une immunoprécipitation d'Ago2 est réalisée 24h après transfection, suivie par une purification des sondes fixées aux ARNm par affinité de la streptavidine. Les ARNm sont extraits et convertis en cDNA avant d'être soumis au séquençage haut débit. Figure adaptée de (Imig et al., 2015)

Le miR-CLIP correspond à l'utilisation d'un mimic biotinylé de notre miR d'intérêt afin d'isoler spécifiquement les ARNm fixés au miR d'intérêt. Lors d'un congrès à Deauville, j'ai rencontré le Dr. Matthieu Meryet-Figuere (BioTICLA, Caen) qui avait déjà utilisé cette approche dans le cadre du cancer de l'ovaire pour le miR-491. Après des discussions sur les avantages et inconvénients de cette approche, nous nous sommes plutôt dirigés sur une approche par RNA-seq plutôt que miR-CLIP. En effet, dans notre modèle d'étude le miR-CLIP ne nous permettrait d'identifier que les cibles potentielles du miR-675 et non du lncRNA *H19*. De plus, le miR-CLIP permet d'identifier les ARNm qui interagissent avec notre miR d'intérêt mais pas forcément ceux dont l'expression est modulée lorsque le miR est surexprimé.

Nous avons donc réalisé une étude transcriptomique par RNA-séq, afin de mesurer les changements d'expression dans deux lignées cancéreuses mammaires MCF-7 et MDA-MB-231 induits par la surexpression du gène *H19*. Grâce à une stratégie d'analyse des résultats, nous

pouvons identifier les cibles putatives du miR-675 et les cibles directes ou indirectes du lncRNA *H19* (Figure 28).



**Figure 28 : Stratégie d’analyse des résultats obtenus par le RNAseq**

Nous avons réalisé un séquençage haut débits des ARNm de cellules MCF-7 mock vs *H19* et MDA-MB-231 mock vs *H19*, en n=2. Le séquençage a été réalisée par la société GATC Biotech (Allemagne) sur un séquenceur Illumina avec : un run de type paired-end avec 30 millions de read pairs garanties et une longueur de read de 2x 150bp.

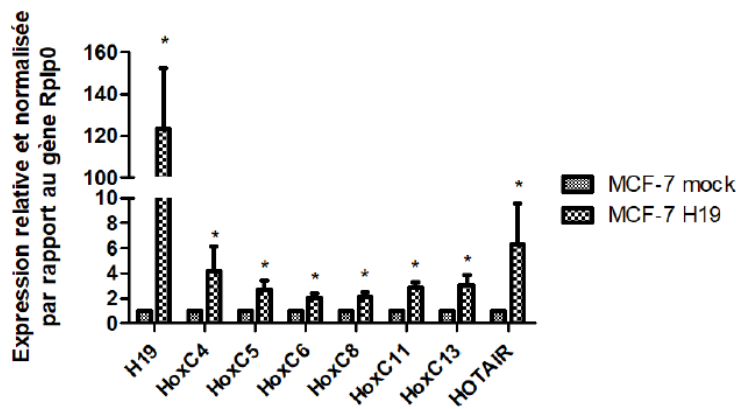
Afin de déterminer les gènes cibles du miR-675, Les gènes sous-exprimés ont été comparés avec la liste des gènes cibles putatifs du miR trouvés *in silico*. Une étude de la littérature nous permet d’identifier sur les gènes surexprimés pour identifier leurs implications possibles dans notre contexte. Un autre projet de notre équipe est l’implication de l’interaction entre *H19* et la protéine de modification épigénétique EZH2. Grâce à nos résultats, nous pouvons comparer la liste des gènes cibles sous-exprimés dans nos cellules surexprimant *H19* avec une liste de gènes cibles d’EZH2, obtenue dans des données de RNA-séq/ChIP-séq disponibles dans des bases de données publiques, afin d’avoir une liste de gènes cibles potentiellement corégulés par *H19* et EZH2.

## III.2. Résultats préliminaires.

### III.2.1. Confirmation de la validité du séquençage

Pour confirmer les résultats obtenus par le séquençage nous avons voulu vérifier certains résultats obtenus en séquençage par une approche de qRT-PCR dans nos cellules. Lors des premières analyses des données, nous avons identifié une surexpression de nombreux membres de la famille de gènes *HOX* dans nos cellules mais également le lncRNA *HOTAIR*, qui

est localisé dans un cluster de gènes *HOXC* sur le chromosome 12. Ces résultats ont aussi été obtenus par qRT-PCR et confirment la validité du séquençage (figure 29).



**Figure 29 : Analyse de l'expression de certains gènes *HOX* par qRT-PCR**

Expression relative et normalisée par rapport au gène *RPLP0* de certains gènes *HOX* mesurée par qRT-PCR. Expériences réalisées dans des cellules MCF-7 mock ou surexprimant le gène *H19*.

Une partie de ces gènes est connue comme étant impliquée dans la tumorigenèse mammaire (Hur et al., 2014 ; Liu et al., 2019). Il serait donc intéressant d'étudier les mécanismes associés à la régulation de ces gènes par *H19*.

### III.2.2. Identifications des cibles du miR-675 dans nos cellules

Pour identifier les cibles du miR-675, j'ai récupéré la liste des cibles *in silico* du miR-675 sur une base de données (mirWalk) et l'ai comparé à notre liste des gènes sous-exprimés dans nos cellules (Figure 30).

Gènes sous-exprimés dans les MCF7 <i>H19</i> comportant des sites de fixation pour le miR-675						Gènes sous-exprimés dans les MDA-MB-231 <i>H19</i> comportant des sites de fixation pour le miR-675					
Cible miR-675-3p			Cible miR-675-5p			Cible miR-675-3p			Cible miR-675-5p		
mirWalk						mirWalk					
5'UTR	CDS	3'UTR	5'UTR	CDS	3'UTR	5'UTR	CDS	3'UTR	5'UTR	CDS	3'UTR
ABCA12	ABCD3	AJAP1	TLE1	JUN	-	MLH1	IDE	GLRX3	JMJD1C	PDGFB	-
HRH1	WDR70	PAPSS2	SEMA4C	TLF1		SSBP1	C <i>Sorf51</i>	PCNP	UTRN		
KCNK2	C <i>Sorf51</i>	SGSM1	C <i>Sorf28</i>	BAG3		B3GALT5	UROS	ITGB5	TM4SF1		
	HERC6		PEU1	FZD7		DLG1	MLH1	SUCLG2			
	GREB1L		EGR3	EPHA2		NMD3	GULP1	UBA5	3		1 -
	SYTL4		PTGER3	MAFB		PRR16	DCUN1D1	TOP2B			
	SEMA3C		ARPP21	IRF7		SWT1	MITF	C3orf18			
	CA2			FAM178B		PDCD4	MLF1	FBXO9			
	DAP			PHLDA1		LIMA1	RSRC1	PRKCI			
	BMP5			DAB2			MBNL1	BDH1			
				STC1			WDR17	SYT16			
				FAM178B			ECT2	SLC2A14			
3	10	3	7	12	-	9	12	12			

**Figure 30 : Cibles potentielles du miR-675 identifiés par notre RNAseq**

Nous avons pu identifier 19 cibles du miR-675-5p dans les MCF-7 et 4 dans les MDA-MB-231. De manière étonnante, nous avons obtenu plus de cibles pour le miR-675-3p dans les MDA-MB-231 que pour le miR-675-5p. Bien que le miR-675-3p a un rôle dans certains contextes (Collette et al., 2017), le miR-675-5p est le plus souvent décrit comme étant le brin

actif du miR-675. De manière étonnante, nous ne retrouvons pas les gènes codant pour les protéines c-Cbl et Cbl-b dans notre étude. En effet, l'expression de ces gènes ne varie pas de manière significative dans notre RNAseq. Cependant, lorsque l'on étudie les données brutes du RNAseq, nous constatons que ces gènes sont bien sous-exprimés, mais que la significativité des résultats n'est pas assez forte. Ceci peut s'expliquer par le fait que nous n'avons réalisé que 2 réplicats biologiques pour effectuer notre analyse. Comme nous l'avons vu dans l'introduction, en fonction des algorithmes utilisés dans les différentes bases de données, les résultats seront différents. Pour identifier un maximum de cibles du miR-675 grâce à notre RNAseq nous allons devoir reproduire cette analyse en utilisant un maximum de base de données.

### ***III.2.3. Perspectives***

Nous avons déjà pu identifier que la famille des gènes *HOXC* est surexprimé lorsque le gène *H19* est surexprimé. L'étude de la régulation de ces gènes par le lncRNA *H19* pourrait mettre à jour un nouveau mode d'action inédit d'*H19*. Pour pouvoir identifier les cibles du miR-675, nous allons devoir continuer l'analyse des données en utilisant de nouvelles bases de données de cibles in silico du miR-675. La prochaine étape sera d'effectuer une analyse bibliographique des gènes cibles potentiellement identifiés et de sélectionner les gènes les plus pertinents dans notre contexte afin de les étudier.

Nous allons valoriser les résultats de cette étude dans un article prochainement. En effet, nous avons démontré que le miR-675-5p régule positivement l'ARNm de FADD. L'analyse de nos résultats de RNAseq a permis de confirmer cette régulation, et d'identifier des voies de signalisation impliquées dans la tumorigenèse et modulées lorsque le gène *H19* est surexprimé. Ces résultats seront détaillés dans la partie suivante.









## **IV. Identification d'un nouveau mode d'action des microARN : le miR-675 augmente l'expression de son ARNm cible FADD.**

### **Contexte**

Des travaux réalisés au laboratoire par Constance Vennin ont identifié un mode d'action inédit pour le miR-675. En effet, en se fixant sur sa séquence cible (seed sequence) située sur l'ARNm de FADD, le miR-675 va induire une surexpression de cette ARNm. Ce résultat est étonnant puisqu'il est admis que les microARN vont empêcher la traduction de leurs ARNm cibles ou induire leurs dégradations. Mon travail dans ce projet a consisté à réaliser les dernières expériences nécessaires à la publication de l'article associé à ces travaux. Pour permettre une meilleure compréhension du projet, j'ai préféré insérer tous les résultats de cette étude. Les expériences de Constance ont permis de mettre en évidence que la fixation de la protéine FADD sur une séquence proche de la seed sequence du miR-675, sur l'ARNm de FADD, est nécessaire à cette action inédite.

Des cellules surexprimant *H19* mais mutées sur le site de fixation de FADD ont été générées grâce à la technologie des TALENs pour étudier ce mécanisme. J'ai utilisé ces lignées pour réaliser une expérience de xénogreffe en souris immunodéficientes afin de mesurer l'effet de cette modification sur les propriétés tumorigènes *in vivo* des cellules. Nous avons récupéré les tumeurs chez les souris et j'ai mesuré les niveaux d'expression d'*H19* et de FADD par qRT-PCR. J'ai également étudié l'impact de cette mutation sur la résistance des cellules au TRAIL.

### **Matériels et méthodes**

#### **Culture cellulaire**

Les cellules MCF-7 sont cultivées dans du milieu DMEM (Gibco®) supplémenté de 1% de sérum de veau fœtal décomplémenté (SVF DC) et de 0,01% de Zell Shield® (Biolabs®). Les cellules T-47D et MDA-MB-231 sont cultivées dans du milieu RPMI-1640 (Gibco®) supplémenté de 10% de SVF DC et de 0,01% de Zell Shield®. Les cellules sont cultivées en atmosphère humide à 37°C et à 5% de CO<sub>2</sub>. Les cellules surexprimant le gène *H19* ont été obtenues comme décrit par Vennin et al. (Vennin et al., 2015).

#### **Tests de prolifération et de viabilité cellulaire**

La prolifération cellulaire a été déterminée en ensemençant 50.10<sup>3</sup> cellules en plaques 6puits et par un comptage journalier. Pour chaque test de viabilité, 15.10<sup>4</sup> cellules sont

ensemencées dans des plaques 96 puits. Les cellules ont été traitées avec du Trail (10µg/mL) pendant 6 heures et la viabilité cellulaire a été déterminée par un test MTT. Pour le test de viabilité des cellules TALENs nous avons ensemencés 1000 cellules en plaque 96 puits, traitées les cellules avec 10ng/mL ou 20ng/mL de TRAIL. Après 48heures, les cellules sont fixées et colorées au crystal violet 0,05%. Le crystal violet est ensuite dissous avec une solution de triton 0,1% et l'absorbance mesurée à 560nm.

### **Western blot**

Les cellules sont lysées avec du tampon RIPA contenant des inhibiteurs de protéases et de phosphatase (protease inhibitor P8340; phosphatase inhibitor cocktail 2 P5726, Sigma-Aldrich). Le protocole de western blot est décrit par Vennin et al. (Vennin et al. 2015) Les anticorps anti-FADD (sc-5559), anti-p-FADD (sc-12439) and anti-actin (sc-47778) ont été achetés chez Santa Cruz.

### **Extraction de l'ADNg**

L'ADNg a été extrait à partir de boîte 100mm à confluence dans du Tris-HCl 10mM pH7.4, NaCl 50mM, EDTA 10mM and SDS 0.2%. Après digestion des protéines par la proteinase K (Roche) sur la nuit à 55°C, l'ADNg a été précipité avec du NaCl 1,6M et de l'éthanol 100%. L'ADNg est dissous dans du Tris-HCl 10mM pH8, EDTA 0,1mM.

### **Extraction d'ARN, reverse transcription et RT-PCR en temps réelle**

Les extractions d'ARN et qRT-PCR ont été réalisées comme décrit précédemment (Bertaux et al., 2008). Les primers utilisés sont décrits dans le tableau S1. Les microARNs ont été extrait via le kit mirVANA™ (Life Technologies). Pour leurs detections, 50ng d'ARNs sont utilisés pour la transcription inverse (miRCURY LNA™ Universal RT, Exiqon). La qRT-PCR a été réalisée avec le SYBR Green master mix (Exiqon). L'ARN humain U6 est utilisé comme contrôle interne. Le fold d'expression des microARN est calculé par l'équation  $2^{-\Delta\Delta Ct}$ .

Pour le test de stabilité des ARN,  $14.10^4$  cellules sont ensemencées dans des plaques 6puits et sont traitées avec 5µg/L d'Actinomycine D (Sigma-Aldrich) pendant 4heures. Les ARNs sont extraits aux temps indiqués et quantifiés par qRT-PCR.

### **Analyse des données de RNAseq**

Un séquençage des ARN de cellules MDA-MB-231 et MCF-7 contrôle (mock) ou surexprimant *H19* a été réalisé par GATC Biotech. Après le contrôle qualité de la librairie, le

séquençage paire-end (2\*150) a été réalisé sur un Illumina HiSeq en dupliquât pour chaque condition. Le format fastq, le contrôle qualité et l'alignement sur le génome humain hg38 ont été réalisés par GATC Biotech. Les nombres de reads comptés ont été normalisés et l'analyse des gènes différentiellement exprimés a été réalisée en utilisation CuffDiff avec un ratio FDR de 0,05. Les voies KEEG, les fonctions biologiques et les gènes différentiellement exprimés (DEG) ont été réalisés en utilisant DAVID v6.8.

### **Construction des plasmides**

La séquence UTR-3' de l'ARNm de FADD a été clonée dans un vecteur *Renilla* luciférase pMIR-REPORT dans un site de restriction SpeI (Ambion). L'UTR-3' mute de FADD a été obtenu par PCR en utilisant les primers listé dans le tableau S1. Le vecteur de normalisation pRL-null est dépourvu de séquence promotrice driver l'expression du gène de la *Renilla* luciférase et a été acheté chez Promega.

### **Transfection des microARN et test d'activité luciférase**

Un total de  $1 \times 10^5$  cellules sont ensemencées en plaque 12 puits pendant 24 heures, Le mimic hsa-miR-675 ou l'antagomir (Dharmacon) sont transfectés avec le DharmaFECT-Dyo selon les instructions du fournisseur (Dharmacon). Brièvement, 75nM de mimic ou d'antagomir sont transfectés avec 500ng de plasmide *Firefly* luciférase et 12ng de plasmide pRL-null exprimant la *Renilla* luciférase pour mesurer l'efficacité de transfection. L'activité luciférase est mesurée 24 heures après la transfection en utilisant le système dual-luciférase reporter assay (Promega), en suivant les instructions du fournisseur. L'activité *Firefly* luciférase est normalisée avec la *Renilla* luciférase.

### **Xénogreffes et extraction des ARNm à partir des tumeurs**

Les expériences de xénogreffes ont été réalisées par le Pr. Eric Adriaenssens à l'animalerie de l'institut Pasteur (PLETHA). Deux millions de cellules (mock, H19 et TAL) ont été injectées dans des souris immunodéficientes SCID (mock =6 ; H19=9 ; TAL =9). Les souris ont été examinées régulièrement et la taille de la tumeur mesurée. Pour l'extraction des ARNm issus des tumeurs, les tumeurs ont été broyées avec un dissociateur de cellules dans du Trizol, les ARNm ont ensuite été purifiés par phénol-chloroforme.

### **Immunoprécipitation des ARN (RIP)**

Les ARN associés aux protéines Ago2 et FXR1 ont été immunoprécipités en utilisant le kit "Magna RIP-binding protein immunoprecipitation" (17-700, Millipore) Les anticorps anti-Ago2 (03-110) et anti-FXR1 (03-176) ont été achetés chez Millipore.

## **Analyses statistiques**

Les données sont exprimées en valeur moyenne  $\pm$  erreur standard d'au moins trois expériences indépendantes. Les analyses statistiques ont été réalisées avec un t-test de Student et une p.value  $<0,05$  est considérée comme significative.

## **Résultats**

### **H19, via le miR-675, augmente l'expression de l'ARNm de FADD**

La recherche *in silico* de cibles potentielles du miR-675-5p réalisée par Constance a permis d'identifier une cible potentielle qui est impliquée dans l'apoptose : FADD. Dans des cellules MDA-MB-231 qui surexpriment *H19*, l'ARNm de FADD est surexprimé et son expression protéique est plus importante (Figure 31, page 124). L'étude de la demi-vie de l'ARNm de FADD grâce à un traitement des cellules à l'Actinomycin D a permis de révéler que la hausse d'expression de FADD dans les cellules surexprimant *H19* peut être expliquée par une meilleure stabilité de son ARNm. Afin de valider que la régulation de FADD par le gène *H19* est bien dû à son microARN, la séquence UTR3' de FADD (contenant la seed sequence du miR) a été clonée en aval de la luciférase dans le vecteur rapporteur pMIR-REPORT luciférase. Une version mutée au niveau de la seed sequence du miR a également été clonée. La surexpression d'*H19* ou le traitement des cellules par mimic du miR-675-5p induisent une augmentation d'activité de la luciférase et révèle la fixation du miR-675-5p sur la seed sequence située dans l'UTR3' de FADD. Lorsque la seed sequence de FADD est mutée, il n'y a pas d'induction d'activité luciférase. Ces résultats indiquent que le miR-675-5p se fixe bien sur l'UTR3' de FADD au niveau d'une seed sequence spécifique du miR et révèle une mode de régulation inattendue du miR-675.

### **La protéine FXR1 régule le recrutement et la fonction du miR-675 sur l'ARNm de FADD**

Afin d'expliquer ce mode de régulation inattendue pour un microARN, nous avons étudié la possible modulation du miR-675 par la protéine FXR1. En effet, en 2007 Vasudevan et collaborateurs ont identifié un nouveau mécanisme d'action des microARN impliquant cette protéine. Dans leur étude, le miR-369 exerce un effet positif sur son ARNm cible, TNFalpha, lorsque les cellules sont privées en sérum (Vasudevan et al., 2007). Ils ont démontré l'existence de complexes RISC différents en fonction du statut prolifératif des cellules et notamment

l'association de la protéine FXR1 au complexe RISC lorsque les cellules sont arrêtées dans le cycle. C'est la fixation de FXR1 au RISC qui permet l'effet positif du miR-369 sur sa cible. Bien que nos cellules soient en phase de prolifération, l'hypothèse de la fixation de FXR1 sur l'ARNm de FADD a été testée puisque nous avons trouvé un site de fixation potentiel, qui est proche de la seed sequence du miR-675-5p (Figure 33, page 125). L'immunoprécipitation des ARNms associés aux protéines AGO2 (membre du RISC) et FXR1 a révélé que l'ARNm de FADD est bien associé à la protéine AGO2, et donc au complexe RISC, mais également à la protéine FXR1. Ce résultat pourrait indiquer que la régulation positive de l'ARNm de FADD par le miR-675-5p est régulée par la protéine FXR1. Afin de confirmer cette hypothèse, le site de fixation de FXR1 sur l'ARNm de FADD a été muté par TALENs (Transcription Activator Like-Effector Nuclease) dans des lignées de MDA-MB-231 contrôle ou surexprimant le gène *H19*. L'expression protéique de FADD est diminuée dans les lignées mutées sur le site FXR1. Le même résultat est obtenu lorsqu'un siRNA dirigé contre l'ARNm de FXR1 est utilisé. Ces résultats montrent que la fixation de FXR1 est bien nécessaire à l'action positive du miR-675-5p sur l'ARNm de FADD. Ce résultat a également été obtenu par une approche de vecteur rapporteur luciférase. Des expériences de RIP ont également été réalisées sur des cellules transfectées avec le vecteur rapporteur luciférase contenant la séquence 3'UTR de FADD ou sa version mutée sur le site FXR1. Le RIP Ago2 montre que cette protéine est fixée fortement sur l'UTR3 de FADD non muté lorsque les cellules sont traitées avec un mimic. Le RIP FXR1 montre également une fixation de FXR1 sur l'UTR3' non muté, comme nous l'avons vu précédemment. Cependant, lorsque l'UTR3' est muté sur le site de fixation à FXR1, il y a une forte perte de fixation d'Ago2. Ce résultat indique que la fixation de la protéine FXR1 est essentielle au recrutement du complexe RISC au niveau de l'UTR3' de FADD et donc à la fixation du miR-675-5p sur cet UTR3'.

### ***H19*, via la surexpression de FADD promeut la résistance à l'apoptose et la prolifération**

La protéine FADD est une protéine impliquée dans l'induction de l'apoptose, nous avons donc étudié la résistance des cellules surexprimant *H19* à plusieurs apoptogènes, comme le Trail (Figure 34, page 126). Nous observons que les cellules MDA-MB-231 qui surexpriment *H19* résistent mieux à l'apoptose induite par le Trail que les cellules contrôles, mais également avec un autre apoptogène comme la staurosporine. Pour expliquer ce phénomène, nous avons étudié la phosphorylation de FADD car cette phosphorylation empêche la formation du complexe d'apoptose (DISC) et favorise la progression du cycle cellulaire. Nous avons constaté que la surexpression d'*H19* induit non seulement la

surexpression de FADD mais également sa phosphorylation. Nous avons également étudié la croissance cellulaire de cellules surexprimant *H19*, mutées ou non sur le site de fixation de FXR1. Les cellules surexprimant *H19* poussent plus vite que les cellules mock et nous observons que les 3 clones mutés pour le site de fixation de FXR1 ont une croissance moindre que les cellules mock.

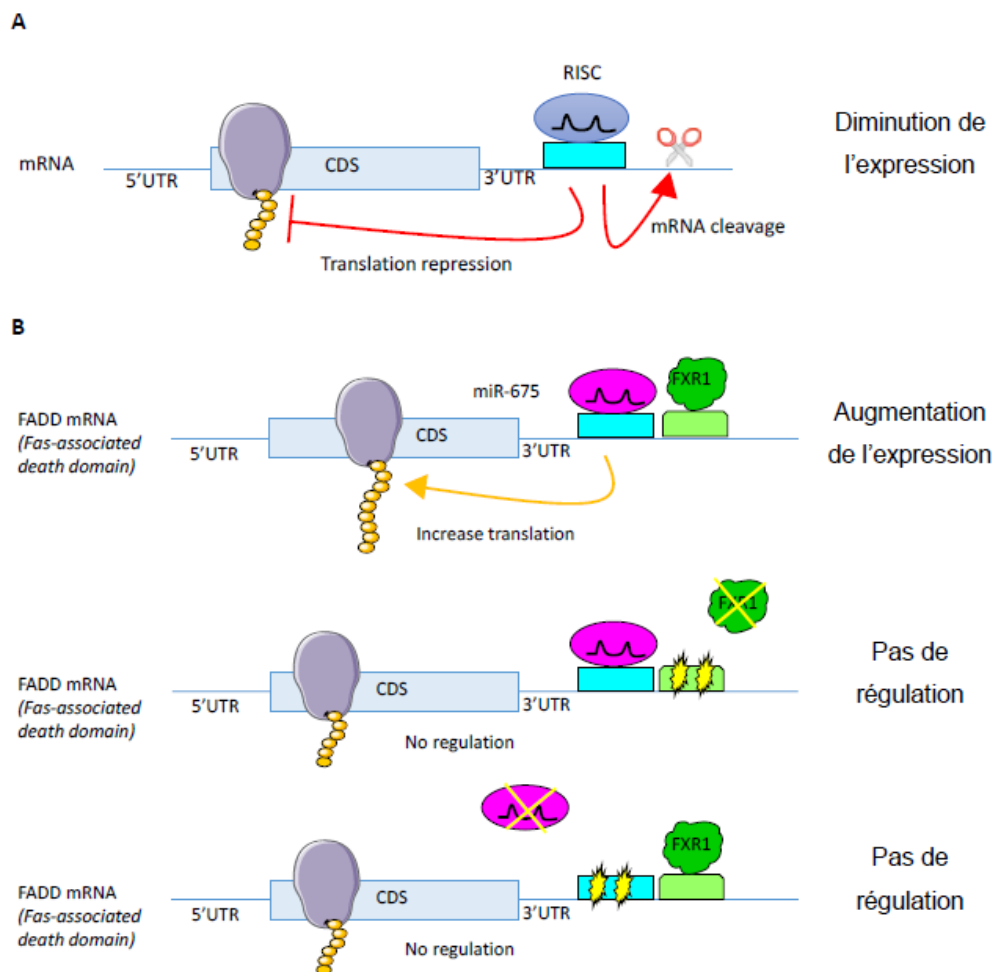
Les expériences précédentes ont été réalisées par Constance lors de sa thèse et de son post-doc au laboratoire, pour compléter ces travaux j'ai utilisé les cellules mock, *H19* et TALEN dans une expérience de xénogreffe *in vivo* afin de mesurer leurs capacités tumorigènes (Figure 35, page 127). Nous observons que les cellules surexprimant *H19* (*H19*) forment des tumeurs plus rapidement et de tailles plus importantes que les cellules contrôles (mock) et les cellules TALENs (TAL). Ces résultats indiquent qu'une partie des effets de la surexpression du gène *H19* sur les capacités tumorigènes des cellules cancéreuses est bien liée à la régulation de la protéine FADD par le miR-675.

Lors du sacrifice des souris, nous avons récupéré les tumeurs pour effectuer une extraction d'ARNm et mesurer le niveau d'expression de FADD. L'expression d'*H19* n'est pas significativement modifiée par la mutation du site FXR1. L'expression de l'ARNm de FADD n'est pas significativement plus forte dans les cellules surexprimant *H19*, comparé aux cellules contrôles (mock). Cependant, lorsque l'on compare le niveau d'expression de FADD entre les cellules surexprimant *H19* étant mutant ou non pour le site de fixation de FXR1 (TAL), nous observons que la mutation du site de FXR1 diminue de manière significative l'expression de FADD. Nous avons également voulu étudier l'impact de la mutation du site FXR1 sur la résistance des cellules au Trail (Figure 35, page 127). Comme nous l'avons déjà vu précédemment, les cellules surexprimant *H19* ont une résistance accrue à l'apoptose. Cependant, lorsque le site de fixation de FXR1 est muté, il n'y a plus de résistance à l'apoptose dans ces cellules comparé au contrôle mock. Ce résultat indique l'action du miR-675-5p sur l'ARNm FADD est essentiel à la résistance à l'apoptose induite par la surexpression d'*H19*.

Enfin, l'analyse des données du RNAseq des cellules MDA-MB-231 que nous avons réalisé a permis d'obtenir un volcano plot montrant la surexpression de l'ARNm de FADD dans les cellules surexprimant le gène *H19* (Figure 36, page 128). Nous avons également mis en évidence la surexpression de nombreux gènes impliqués dans des voies de signalisations impliquées dans les processus de tumorigénèse.

## Discussion

Dans cette étude, nous avons identifié une nouvelle cible du miR-675-5p dans les cellules cancéreuses mammaires : l'ARNm FADD. La surexpression du gène *H19* entraîne une meilleure stabilité de l'ARNm de FADD (Figure 32).



**Figure 32 : 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.

Cette action est due à la fixation du miR-675-5p sur l'UTR-3' de FADD, via un site prédit *in silico*. Nous avons montré que c'est la fixation de la protéine FXR1 sur un site proche du site de fixation du miR-675 qui permet la régulation positive de l'ARNm de FADD induite par le microARN. La protéine FXR1 est un membre de la famille FXR (Fragile-X related protein) dont les membres (FMRP, FXR1 et FXR2) sont principalement impliqués dans le développement neuronal, musculaire ou encore cardiaque. Ces protéines possèdent des domaines de fixations

aux ARNm qui permettent la modulation de la stabilité, de la traductibilité ou de la localisation cellulaire de ces ARNm (Khera et al.2010).

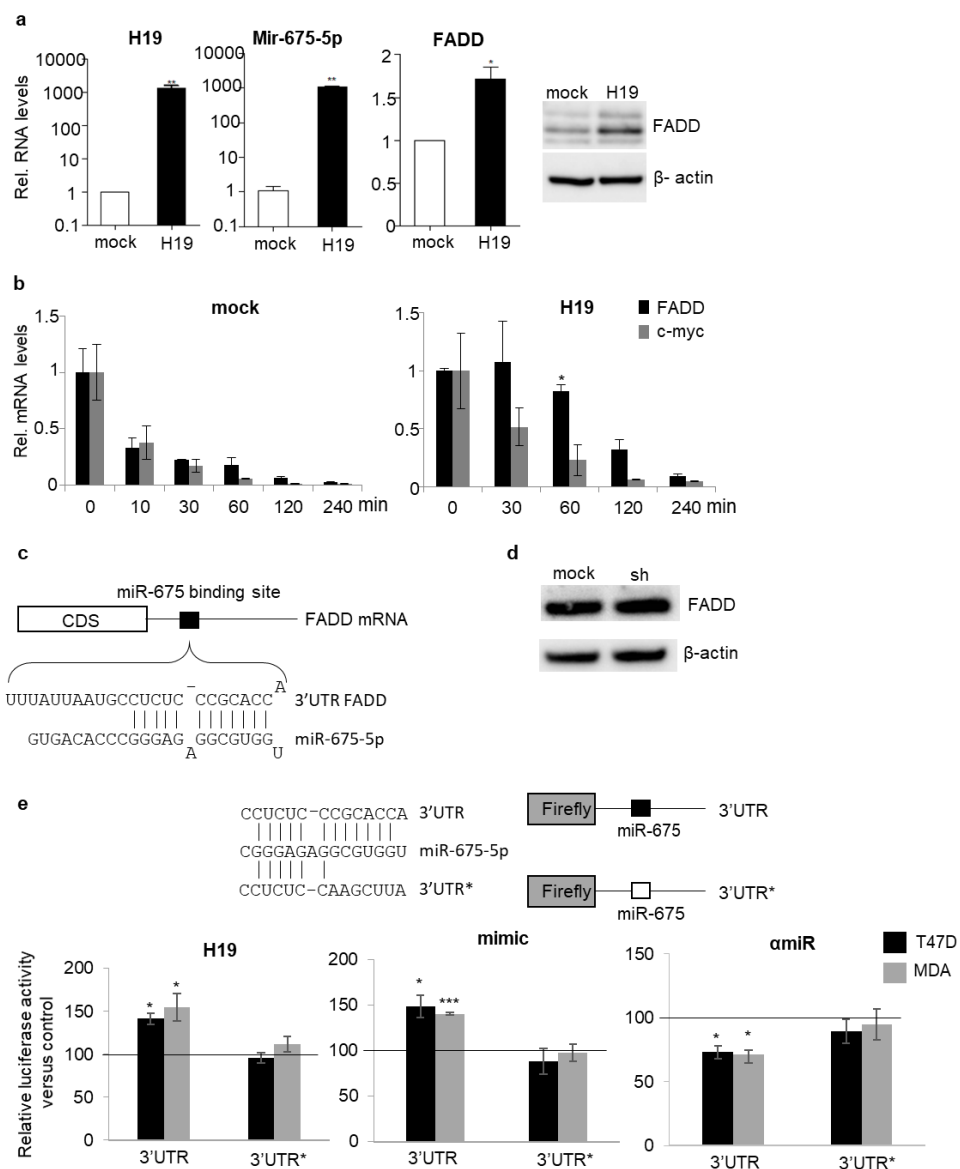
Notre étude révèle un mode d'action inédit des microARN qui permet de réguler positivement une cible du miR-675. En effet, il est admis que les microARNs se fixent, généralement, sur les UTR-3' des ARNm pour les réguler négativement. Tout comme l'a démontré Vasudevan et collaborateurs, ce mode d'action nécessite la fixation de la protéine FXR1 sur une séquence proche de la seed séquence du microARN (Vasudevan et al., 2007). L'action positive du miR-369 sur sa cible, le TNFalpha, n'a été observée qu'en absence de sérum. Ce mode d'action semble cellule-spécifique, en effet, des publications récentes ont également montré que le miR-675 régule négativement l'ARNm de FADD dans les cancers gastriques et les hépatoblastomes (Ge et al., 2019; Yan et al., 2017). Nous pensons qu'il existe d'autres acteurs de cette régulation et que leurs disponibilités, différentes selon les cellules, pourraient expliquer les modes d'actions différents observés pour le miR-675-5p sur l'ARNm de FADD. L'identification des acteurs moléculaires impliqués dans la spécificité d'action du miR-675 permettrait de mieux comprendre l'action des microARN sur leurs cibles.

La conséquence fonctionnelle de la régulation positive de l'ARNm de FADD par le miR-675 se traduit par prolifération et une résistance à l'apoptose accrue dans les cellules surexprimant le gène *H19*. Ce phénotype est dû à la surexpression d'une protéine FADD qui sera majoritairement phosphorylée. En effet, FADD est une protéine pro-apoptotique de 27kDa constituée de deux domaines : un domaine effecteur de mort (Death Effector Domain, DED) et un domaine de mort (Death Domain, DD). L'activation des récepteurs de mort cellulaire (TNFR, Fas receptor...) induit la fixation de FADD sur ces récepteurs grâce à son DD. La protéine FADD va permettre le recrutement et l'action des caspases 8 et 10 *via* son DED. Les caspases activées vont former le complexe DISC (Death-inducing signaling complex) qui va permettre l'induction de l'apoptose. La protéine FADD est donc importante dans la transduction du signal apoptotique de la voie extrinsèque. Cependant, des modifications post-traductionnelles peuvent modifier la fonction de cette protéine. La phosphorylation des sérines 194 et 200, situées proche du DED, empêche le recrutement et l'activation des caspases 8 et 10. (Vilmont et al., 2015; J. Zhang et al., 2004).

**Nous avons mis en évidence un nouveau mécanisme de recrutement et de régulation de l'action des microARN sur leurs ARNm cibles. Nous montrons que des protéines associées aux ARNm, FXR1 dans notre cas, sont nécessaires au recrutement de certains microARN sur**

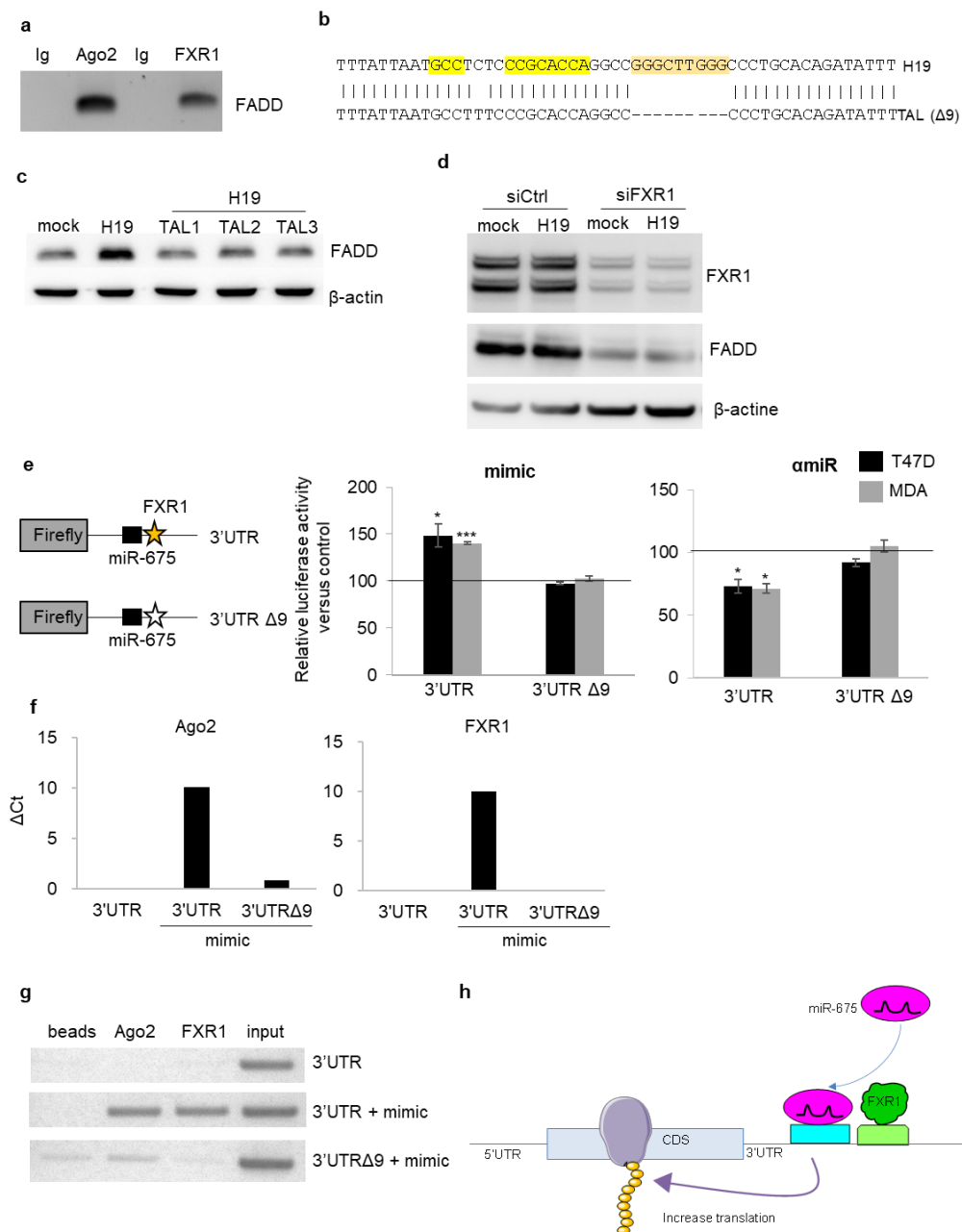


leurs cibles et régulent la fonction du complexe RISC sur cette cible. De plus, grâce à ce mécanisme d'action, l'ARN *H19* favorise la progression tumorale et la résistance aux drogues apoptotiques.



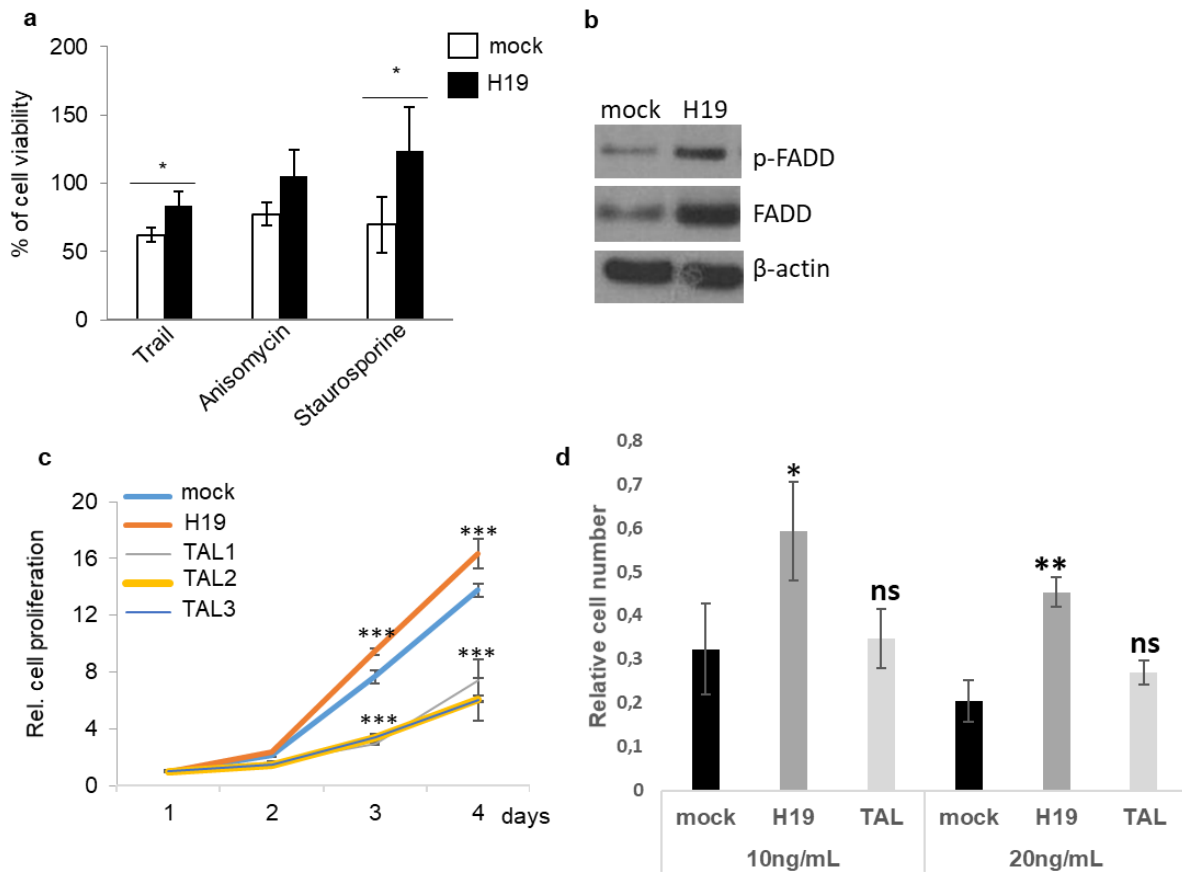
**Figure 31 : H19, via le miR-675, augmente l'expression de l'ARNm de FADD.**

**a**, Expression relative de *H19*, du miR-675 et de FADD déterminée par qRT-PCR (panel de gauche) et par western blot (panel de droite) dans des MDA-MB-23 transfectées avec un vecteur contrôle (mock) ou un vecteur de surexpression d'*H19* (*H19*). **b**, Niveau relative d'ARNm de FADD et c-myc dans des cellules contrôles (mock) ou surexprimant *H19* (*H19*) traitées par 5µg/mL d'actinomycine D. **c**, Site de fixation putatif du miR-675 sur l'UTR-3' de FADD, déterminée par les bases de données miRBase et TargetScan. **d**, Expression de FADD dans des cellules invalidées pour *91H* (sh) ou contrôles (mock) déterminée par western blot. **e**, Activité luciférase relative dans des cellules cancéreuses mammaires (T47D, MDA-MB-231) après co-transfections avec un plasmide rapporteur et *H19*, un mimic miR-675 ou un inhibiteur du miR-675 (amiR). Les graphiques représentent une moyenne de 3 expériences indépendantes. \* $p < 0.05$  \*\* $p < 0.01$  \*\*\* $p < 0.001$ .



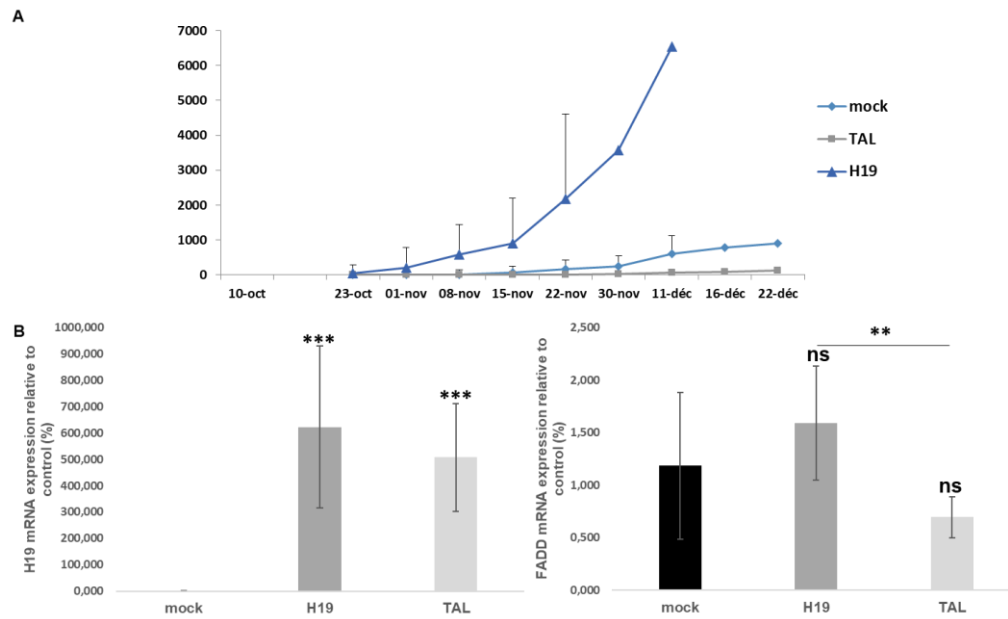
**Figure 33 : La protéine FXR1 régule le recrutement du miR-675 et sa fonction sur l'ARNm de FADD.**

a, Détection de l'association d'Ago2 et FXR1 sur l'ARNm de FADD par RIP. B, Analyse de la séquence de l'UTR-3' de FADD dans des cellules surexprimant *H19* (*H19*) et dans des cellules surexprimant *H19* mais transfectées avec TALENs (TAL). Jaune : site de fixation du miR-675 ; Violet : site de fixation de FXR1. C, Analyse par western blot de l'expression de FADD dans des cellules MDA-MB-231 contrôles (mock) ou surexprimant *H19* (*H19*) transfectées avec un siRNA contrôle (siCtrl) ou siRNA contre FXR1 (siFXR1). e, Activité luciférase relative dans des cellules cancéreuses mammaires (T47D, MDA-MB-231) après co-transfections d'un vecteur rapporteur et d'un mimic du miR-675 ou d'un inhibiteur du miR-675 (amiR). f,g, Détection par qPCR de l'UTR-3' de FADD par qPCR (f) ou gel d'agarose (g) après RIP d'Ago 2 (panel de gauche) ou FXR1 (panel de droite). Les cellules ont été co-transfectés par un vecteur rapporteur *wild type* (3'UTR) ou délété pour le site de fixation de FXR1 (3'UTRΔ9) et un mimic du miR-675. h, Nouveau mécanisme de recrutement du miR-675. \* $p < 0.05$  \*\*\* $p < 0.001$



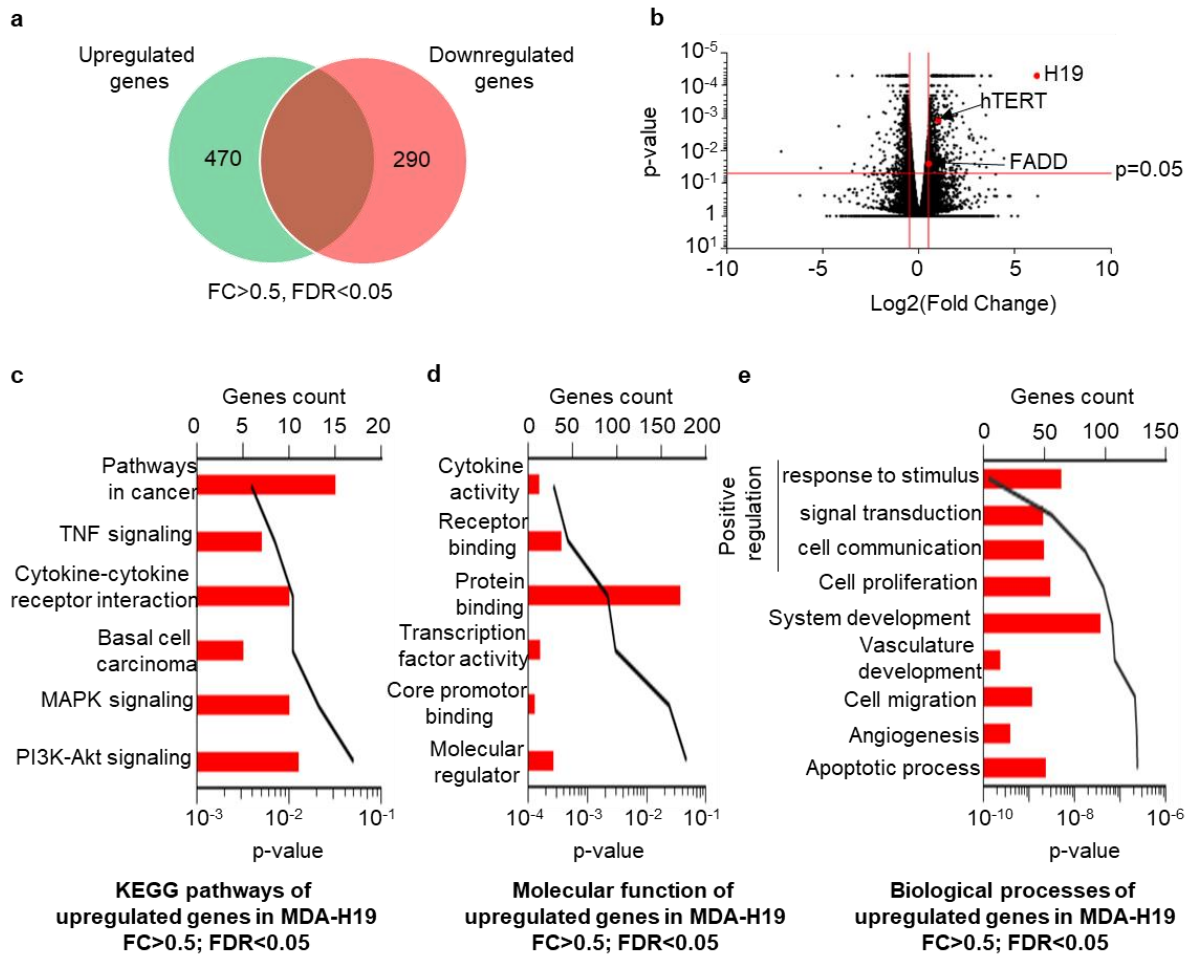
**Figure 34 : H19, via la surexpression de FADD, promeut la résistance à l'apoptose et la prolifération**

a, Viabilité cellulaire relative de cellules contrôles (mock) et surexprimant H19 (H19) ,traitées avec du Trail, de l'anisomycine ou de la staurosporine, déterminée par un test MTT. b, détection de FADD et p-FADD par western blot dans des cellules contrôles (mock) ou surexprimant H19 (H19). c, Prolifération cellulaire relative de cellules contrôles (mock), surexprimant H19 (H19) et surexprimant H19 transfectées avec le TALEN (TAL). d, Viabilité cellulaire relative des cellules contrôles (mock), surexprimant H19 (H19) et surexprimant H19 transfectées avec le TALEN (TAL,), traitées avec du Trail, déterminée 48heures après le traitement.\* $p < 0.05$  \*\*\* $p < 0.001$ .



**Figure 35. H19 promeut la croissance tumorale et expression de FADD dans les tumeurs**

A. Expériences de xénogreffes en souris immunodéficientes SCID. Les cellules mock, H19 et TAL ont été injectées le 10 Oct. 6 souris ont été injecté avec des cellules mock, 9 avec des cellules H19 et 9 avec des cellules TAL. L'axe des ordonnées représentent la taille de la tumeur (mm<sup>3</sup>). B. Expressions des ARNm d'*H19* et de *FADD* au sein de ces tumeurs. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .



**Figure 36 : H19 promeut la tumorigénèse en augmentant l'expression génique.**

a, diagramme de Venn montrant le nombres de gènes différentiellement exprimés dans les cellules MDA-MB-231 surexprimant le gène *H19*. b, Volcano plot montrant les gènes différentiellement exprimés dans ces cellules. Le plot indique la p-value des gènes (axe Y) contre leurs fold change (log2) respectifs. Le seuil de significativité a été placé à 0,05. Les points rouges indique les gènes *H19*, *hTERT* et *FADD* qui sont surexprimés dans les cellules H19. c, Voies KEGG, d, fonctions moléculaires, processus biologiques associés aux gènes surexprimés dans les cellules H19.

# Discussion





## Discussion

Les travaux réalisés au cours de ma thèse ont permis de mettre de nouveaux mécanismes d'actions du long ARN non codant *H19* et de son microARN. J'ai démontré que le long ARN non codant *H19* interagit avec les protéines p53 et MDM2 afin de former un complexe ternaire. La formation de ce complexe impacte négativement l'action de la protéine p53 et participe au rôle oncogénique du gène *H19* dans les cellules cancéreuses mammaires. La poursuite de mes travaux sur l'implication du gène *H19* dans les cellules souches cancéreuses (CSC) a permis de démontrer que non seulement le lncRNA *H19* est important dans le maintien des propriétés souches des CSC, mais également le miR-675. Enfin, la réalisation du profil transcriptomique des cellules cancéreuses mammaires surexprimant le gène *H19* a permis d'identifier des voies de signalisation régulées par le gène *H19* et permettra d'apporter de nombreux résultats intéressants pour les futurs projets de l'équipe. J'ai participé à la finalisation de l'étude de la régulation de la protéine FADD par le miR-675. Ces travaux démontrent l'action positive du miR-675-5p sur une de ces cibles, l'ARNm de FADD. Ce mode d'action atypique est très peu décrit dans la littérature et pourrait révolutionner nos connaissances sur l'action des microARN.

### **I. Le long ARN non codant *H19* régule négativement la protéine p53 et participe à l'instabilité génomique.**

#### **I.1. Découverte d'un nouveau mécanisme d'action du lncRNA *H19***

Mes travaux de thèse ont mis en évidence un nouveau mécanisme de régulation de la protéine p53 dans le cancer du sein. En réponse à des dommages à l'ADN, la protéine p53 va induire un arrêt dans le cycle cellulaire pour permettre la réparation des dégâts ou l'induction de l'apoptose si les dommages à l'ADN sont trop importants. La protéine p53 est une protéine essentielle pour l'intégrité du génome, qui empêche la transformation de cellules normales en cellules cancéreuses. Le gène *TP53* est muté dans de très nombreux cancers, dont les cancers du sein. La plupart de ses mutations sont des mutations perte de fonctions (LOF : Loss-of-function), ce qui classe le gène *TP53* comme un gène suppresseur de tumeurs d'après la définition de Haber et Harlow (Nat genet 1997). Le gène *TP53* est une protéine essentielle dans le maintien de l'intégrité des cellules de la glande mammaire.

Notre étude montre que la surexpression du gène *H19* dans des lignées de cancer du sein *wild-type* (MCF-7) ou mutés (T-47D et MDA-MB-231) pour la protéine p53 induit une

diminution d'expression de la protéine p53. Cette régulation négative de p53 dans les cellules MCF-7 empêche l'action de p53 en réponse à des dommages de l'ADN et diminue notamment le blocage du cycle cellulaire p53-dépendant induit par un traitement à la doxorubicine. Dans la littérature une fixation entre le lncRNA *H19* et la protéine p53 a déjà été identifiée dans les cancers gastriques. Cependant, le mécanisme d'action d'*H19* sur la protéine p53 est peu clair. Yang et al. ont démontré qu'*H19* interagit physiquement avec la protéine p53 et que cette interaction pourrait participer à l'inactivation de p53 (Yang et al., 2012). Pour la première fois, nous démontrons que le lncRNA *H19* interagit non seulement avec la protéine p53 mais également MDM2. Cette interaction tripartite a été observée *in cellulo* mais également *in vitro*. Nous pensons qu'*H19* pourrait servir d'ARN d'échafaudage pour l'interaction p53-MDM2 et favoriser la dégradation de p53 *via* le protéasome. Cette interaction empêcherait également le transport de p53 dans le noyau et son action en tant que facteur de transcription. L'interaction entre p53 et *H19* étant décrite dans les cancers gastriques, il serait intéressant d'étudier l'interaction entre *H19* et MDM2 dans ces cancers afin de généraliser le mécanisme que nous avons identifié. Dans notre étude, nous avons également réalisé un Hprt assay qui nous a permis de montrer que la surexpression du gène *H19* induit de l'instabilité génomique en l'absence de génotoxique. Cette expérience suggère que le gène *H19* pourrait être responsable d'une instabilité génomique chez des patientes non soumises à des génotoxiques. **Notre étude a donc permis de mettre en évidence un nouveau mécanisme d'action oncogénique du lncRNA *H19* dans les cancers du sein *via* la régulation de la protéine p53. Cependant, de nombreux points importants de l'impact de la régulation de p53 par *H19* dans les cellules et des possibles implications cliniques restent à discuter.**

## **I.2. Développement de nouvelles thérapies**

Les oncogènes apparaissent comme des cibles idéales pour le développement de thérapies anticancéreuses car ce sont souvent des variants de protéines hyperactifs. Ils semblent plus simples à cibler que des gènes suppresseurs de tumeurs, inactivés par des mutations non-sens ou des mutations indels. Les projets de développements thérapeutiques se focalisent donc sur des inhibiteurs de kinases ou des approches pour inhiber des oncoprotéines actives. Relativement peu de stratégies thérapeutiques visent à rétablir l'activité d'un gène suppresseur de tumeur. Certaines thérapies ont été développées pour rétablir l'expression ou l'activité suppresseur de tumeur de p53 dans les tumeurs où cette

protéine est non fonctionnelle. Par exemple, des études cliniques ont montré que l'utilisation d'inhibiteurs spécifiques de MDM2 (peptides, anticorps ou oligonucléotides antisens) permet de réduire la croissance des tumeurs lorsque p53 est fonctionnel (Toledo et al., 2007). Cependant, ces inhibiteurs présentent une grande toxicité.

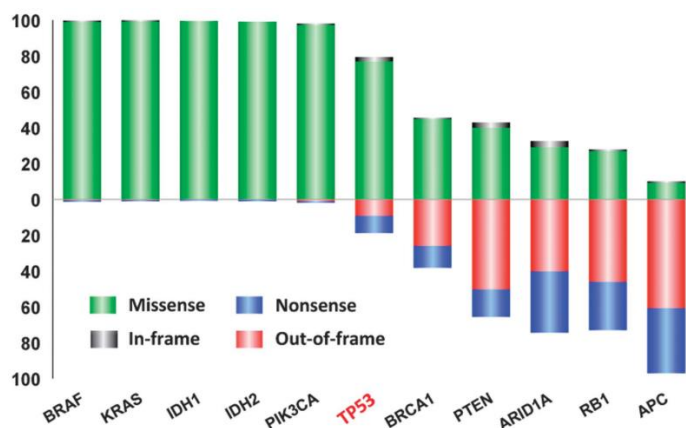
Nos résultats démontrant que le lncRNA *H19* régule négativement la protéine p53, nous pourrions imaginer que cibler le lncRNA *H19* dans les cancers du sein afin de rétablir l'activité de p53 serait une bonne piste thérapeutique. De plus, nous savons que la régulation de p53 n'est pas le seul mode d'action oncogène du gène *H19* dans les cancers du sein. Cibler *H19* semble être une piste thérapeutique intéressante puisque ce gène est surexprimé dans de nombreux cancers et contrairement à d'autres cibles thérapeutiques envisagées aujourd'hui, le gène *H19* est peu exprimé et ne participe pas à des processus vitaux à l'âge adulte. Les risques d'effets secondaires importants pour les patients semblent donc très limités. En 2002, une thérapie permettant de cibler les cellules qui surexpriment *H19* (sans pour autant cibler l'ARN *H19* directement) a été développée (Ohana et al., 2002). Un plasmide (H19-DTA ou BC-819) permettant l'expression de la toxine diphtérique sous le contrôle du promoteur du gène *H19* permet d'éliminer les cellules cancéreuses qui expriment *H19*. L'utilisation de ce plasmide fait l'objet d'un essai clinique de phase 2b dans le cancer de la vessie et de phase 1/2a pour le cancer de l'ovaire. Le résultat d'une étude de 2017 chez des patients atteints de cancer de l'ovaire et/ou péritonéal montre qu'une injection intrapéritonéale permet d'améliorer la prise en charge des patients. (Lavie et al., 2017). Cependant, la cohorte de patients utilisée étant faible, ces résultats restent à confirmer. A l'heure actuelle, ce plasmide n'a jamais été utilisé dans les cancers du sein.

Pour empêcher l'action d'*H19* il est très simple d'utiliser une approche par siRNA *in vitro*, cependant, l'utilisation de siRNA *in vivo* est très compliquée (Hu et al. 2019). La demi-vie d'un siRNA injecté dans le sang est inférieure à 5 minutes du fait de la présence de nombreuses nucléases présentes dans le plasma et de la clairance rénale. De plus les siRNA vont également devoir passer la membrane plasmique pour pouvoir être pris en charge par le RISC. Le développement de modifications spécifiques des siRNA pour augmenter leurs stabilités ainsi que des systèmes de vectorisation est devenu un enjeu majeur pour l'industrie pharmaceutique et les principales pistes envisagées sont résumées dans la revue de Hu et al. 2019 (Hu et al., 2019). Depuis le premier essai clinique d'un siRNA thérapeutique en 2004, une trentaine d'autres ont suivis et seulement une molécule a été approuvée par la FDA et la commission européenne en 2018. Le patisiran (ONPATRO™) est la première thérapie à base

de siRNA reconnue comme efficace et permet le traitement de l'Amylose de la transthyréine héréditaire (Rizk et al. 2019). Le développement des siRNAs thérapeutiques dans les années à venir pourrait permettre l'utilisation de siRNAs pour cibler l'expression d'*H19* dans les cancers. **La réactivation de p53, dans les cancers où le gène *TP53* n'est pas muté de manière perte de fonction, en ciblant le gène *H19* permettrait l'utilisation de molécules anticancéreuses nécessitant la présence d'une protéine p53 active pour tuer les cellules. Les thérapies ciblant le gène *H19* pourrait donc servir à sensibiliser les tumeurs à d'autres thérapies. Il a notamment été démontré que le gène *H19* permet la chimiorésistance à différentes molécules anticancéreuses dont le paclitaxel dans les cellules MCF-7, qui ont un gène *TP53* wild-type.**

### **I.3. Rôle ambigu du gène *TP53* dans les cancers : gain de fonction oncogénique de p53**

Pour discuter de la pertinence du statut mutationnel de *TP53* dans les cancers du sein, il nous faut discuter du rôle ambigu du gène *TP53* dans les cancers. Une des caractéristiques de *TP53*, comparé aux autres gènes suppresseurs de tumeurs, est son mode d'inactivation. Plus de 80% des altérations somatiques et germinales de *TP53* sont des mutations faux sens (mutation d'un nucléotide induisant un changement de codant, pour un codant qui ne sera pas un codant STOP) qui conduisent à la formation d'une protéine mutante stable qui s'accumule dans le noyau des cellules cancéreuses. Contrairement à la plupart des gènes suppresseurs de tumeurs qui sont inactivés par des mutations de type *indel* (insertions/délétions) ou non-sens (mutation d'un nucléotide induisant un changement de codant, pour un codant qui sera un codon STOP). Les gènes ayant un fort taux de mutations faux sens sont généralement des oncogènes et il est possible de prédire la fonction oncogène ou suppresseur de tumeur d'un gène grâce aux taux de mutations dans ce gène (Figure 36).



**Figure 37 : Spectre de mutation des gènes fréquemment mutés dans les cancers.**

BRAF: v-raf murine sarcoma viral oncogene homolog B1; KRAS, v-Ki-ras2 Kirsten rat sarcoma 2 viral oncogene homolog; IDH1, isocitrate dehydrogenase 1; IDH2, isocitrate dehydrogenase 2; PIK3CA, phosphoinositide-3- kinase, catalytic, alpha polypeptide; TP53, tumor protein p53; BRCA1, familial breast/ ovarian cancer gene 1; PTEN, phosphatase and tensin homolog gene; ARID1A, AT- rich interactive domain 1 A; RB1, retinoblastoma gene; APC, adenomatous polyposis coli. Figure de Soussi & Wiman, 2015.

Sur ce schéma nous observons que *TP53* possède un profil de mutation plus proche des oncogènes que des gènes suppresseurs de tumeurs comme *APC* ou *PTEN*. **De nombreuses études ont démontré que la forte sélection des mutations faux-sens de *TP53*, qui permettent l'expression d'une forme full-length de p53 dans les tumeurs suggèrent un rôle vital de cette protéine dans la transformation oncogénique (Milićević et al., 2014; Soussi & Wiman, 2015; Walerych et al., 2012).** En effet, de nombreux mutants de p53 ne vont pas uniquement perdre leurs fonctions de suppresseurs de tumeurs mais également gagner de nouvelles fonctions oncogéniques. Ces mutants de p53 « GOF (Gain-of-function) » sont considérés comme des oncogènes importants. Dans les cancers du sein, l'implication de ces mutants GOF est bien détaillée dans la littérature. Par exemple, en 2006 Hui et collaborateurs ont montré dans des cellules MDA-MB-231 que la protéine p53 mutée est stabilisée par un niveau élevé de Phospholipase D et permet la résistance à l'apoptose lorsque les cellules sont privées de sérum (Hui et al., 2006). Il y a une grande variété de mutants de p53, créé par des mutations au sein du gène *TP53*, qui peuvent acquérir des propriétés nouvelles qui leur confèrent un rôle oncogénique. Des variantes mutants de p53 (mutp53) peuvent reconnaître de nouveaux éléments de réponses sur l'ADN, bien qu'aucunes séquences consensus n'est décrit (Bellazzo et al., 2018 ). Certains mutp53 interagissent avec des facteurs de transcriptions comme E2F1 ou E2F4 pour activer l'expression d'oncogènes (ID4) ou diminuer l'expression de gènes impliqués dans la réparation de l'ADN afin de promouvoir la survie des cellules et la

tumorigenèse. Des mutsp53 peuvent également interagir avec d'autres membres de la famille de p53 comme p63 et p73 pour inhiber leurs fonctions suppresseurs de tumeurs (Bellazzo et al., 2018). Le Tableau 5 représente les voies de signalisation, les médiateurs et les gènes ciblés par certains mutp53.

Mutant p53 discovered	Mutant p53 validated	Pathway(s)	Mediator(s)	Downstream proteins/genes	Leading model(s)	Mutant p53-related phenotype
R248W	R273H, R175H	DNA damage response	MRE11	AKT	HUPKI mice/MEFs	Genomic instability
	R175H, R280K, L194F, R273L, R249S, R248Q, C242F	DNA damage response	ETS2 <sup>a</sup>	TDP2 <sup>a</sup>	Li-Fraumeni-derived cell line	Chemoresistance
R175H	R273L, R249S, R280K, R175H	Nucleotide homeostasis	ETS2 <sup>a</sup>	Nucleotide metabolism genes <sup>a</sup>	Li-Fraumeni and breast cancer cells	Cell proliferation
	L194F, R273H	IL-8 and GRO- $\alpha$ signaling	NFYA <sup>a</sup>	Cyclin A, B, E, CDK1, CDC25C <sup>a</sup>	Breast cancer cell lines	Cell proliferation
	R273H, R280K, R273H	Interleukin signaling, VDR signaling	ID4, VDR <sup>a</sup>	IL-8, GRO- $\alpha$ , IGFBP3, CYP24A1 <sup>a</sup>	Breast cancer cell lines	Angiogenesis Reduced apoptosis
	R273H, R280K	PDGF receptor $\beta$ signaling	p73 <sup>a</sup> , NFY complex <sup>a</sup>	PDGFR $\beta$ <sup>a</sup>	Pancreatic cancer mouse model and cell lines	Metastasis
R280K	H179R, G245S, R248Q, R249S, R273H	Phospholipid metabolism	ETS2 <sup>a</sup>	Pla2g16 <sup>a</sup>	KI mouse model, osteosarcoma cell line	Tumor growth and metastasis
	R273H	Cell cycle, cell movement	Pin1	Cyclin E2, BUB1, DEPDC1 <sup>a</sup>	Breast cancer cell lines	Cell proliferation, migration
	R175H	TGF $\beta$ -induced migration/invasion	SMAD/p63 <sup>a</sup>	SHARP-1, Cyclin G2	Breast cancer cell lines	Metastasis
R273H	R175H, R273H, M237I	TNF $\alpha$ -driven inflammation	DAB2IP	JNK, NF- $\kappa$ B, and their targets	Breast cancer cell lines	Cancer-related inflammation
	R280K	Steroid synthesis	SREBP1/2 <sup>a</sup>	MVK, FDFT1, TM7SF2, NSDHL <sup>a</sup>	Breast cancer cell lines	Tumor growth
	- R280K, L194F	HB-EGF signaling DNA replication, PAPP signaling	NRD1 -	- PARP, MCM4, PCNA	p53 null lung carcinoma Breast cancer cell lines	Invasion Cell proliferation
R175H, R273H (overexpressed)	R280K	EGFR/integrin signaling	p63	$\alpha$ 5 $\beta$ 1 integrin, EGFR	p53 null lung carcinoma, breast cancer cell lines	Cell motility, invasion
R175H, R273H, D281G (overexpressed)	-	NF- $\kappa$ B signaling	-	NFKB2 <sup>a</sup>	p53 null lung carcinoma	Chemoresistance
R175H, R248Q, R273H (overexpressed)	R175H, R273H	Glucose metabolism, Warburg effect	RhoA/ROCK	GLUT1	p53 null lung carcinoma, MEFs, breast cancer cell lines	Tumor growth
R175H, R248Q, R248W, R249S, R273H, R282W (overexpressed)	R273H, R280K	Membrane and secreted signaling factors	p63 <sup>a</sup>	DKK1, METTL7B, TFPI2 <sup>a</sup>	p53 null lung carcinoma, breast cancer cell lines	Invasion
V143A, R175H, R248W, R249S, R273H, R282W (overexpressed)	R175H, R248Q, R273C	Cell cycle, apoptosis	TopBP1 <sup>a</sup> , p63/p73 <sup>a</sup> , NFY <sup>a</sup>	Cyclin A, B, E, CDK1, CDC25C, BAX, NOXA <sup>a</sup>	p53 null lung carcinoma, breast cancer cell lines	Proliferation
R175H, H179R, G245S, R248Q, R273H (overexpressed)	R175H, R273H	Ras-mediated signaling	BTG2, NF- $\kappa$ B <sup>a</sup>	CXCL1, IL1B and MMP3 <sup>a</sup>	Human lung fibroblasts WI-38	-
R248Q, R249S, R273H (endogenous)	R175H, R248W	Chromatin epigenetic modification	ETS2 <sup>a</sup>	MLL1, MLL2, MOZ <sup>a</sup>	Breast cancer cell lines, MEFs, Li-Fraumeni cell lines	Proliferation and tumor growth

<sup>a</sup>Transcription-related mediators and transcriptionally regulated downstream mutant p53 targets.

## Tableau 6 : Actions connues de certains mutants gain de fonctions de p53 dans les cancers

Figure adaptée de Walerych et al. 2015

**Pour conclure, une mutation de TP53 dans les cancers peut mener à de nombreux phénotypes différents. Les protéines mutantes de p53 peuvent être inactivées (LOF) ou acquérir des propriétés oncogéniques (GOF).**

### I.4. Pertinence du statut mutationnel de TP53 dans les cancers du sein

Le gène TP53 est muté dans plus de 60% des cancers et dans près de 30% des cancers du sein. Pour les cancers du sein, la distribution des pourcentages de mutations va dépendre des sous-types moléculaires : on retrouve TP53 muté dans 26% des tumeurs lumineales (17%

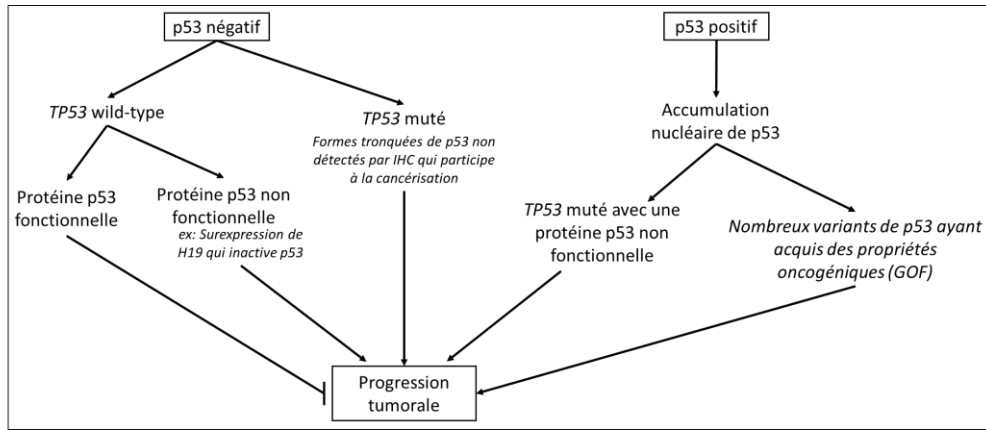
luminal A, 41% luminal B), dans 50% des tumeurs HER2+ et dans 88% des cancers basal-likes. En 2007, l'ASCO (American Society of Clinical Oncology) a confirmé ses précédentes recommandations de 1999 en ce qui concerne l'utilisation de p53 en tant que marqueur pour les cancers du sein : « Les données présentes sont insuffisantes pour recommander l'utilisation de la mesure de p53 dans le management des patientes avec un cancer du sein ». Encore aujourd'hui, le statut de p53 dans les tumeurs du sein n'est pas utilisé par les cliniciens.

L'intérêt pronostique des altérations de *TP53* dans les cancers du sein a fait l'objet de nombreuses études, avec des résultats conflictuels. Deux méthodes principales sont généralement utilisées : le séquençage de l'ADN et l'immunohistochimie (IHC). La protéine p53 wild-type est une protéine très rapidement dégradée dans les cellules normales. Les mutations de *TP53* sont souvent associées à la production d'une protéine p53 non fonctionnelle avec une demi-vie accrue, qui s'accumule dans le noyau. Un échantillon qui réagit positivement avec un anticorps ciblant p53 est donc considéré comme ayant une forme mutée de p53. L'IHC est fréquemment utilisé pour évaluer le statut mutationnel de p53 dans les tumeurs car c'est une méthode qui est peu coûteuse et rapide. Cependant, de nombreux désavantages de cette approche commencent à émerger dans la littérature (Milićević et al., 2014) :

- La variabilité des résultats de marquage p53 (de 20% à 60% de tumeurs p53 positives en fonction des études peut s'expliquer par l'utilisation de différents anticorps, de l'origine du matériel biologique utilisé, de la méthode de scoring ou encore de la sélection des groupes de patientes sélectionnés qui peut biaiser les études

- Nous savons maintenant que la corrélation entre l'accumulation de la protéine p53 mesuré par IHC et les mutations de *TP53* observé par séquençage est inférieure à 75% dans les cancers du sein. Ceci s'explique notamment par le fait que toutes les mutations de *TP53* ne conduisent pas à l'apparition d'une protéine plus stable et il est connu que certaines mutations vont produire des protéines tronquées qui ne sont pas détectées par IHC.

L'utilisation de l'IHC ne semble donc pas permettre d'établir clairement l'implication clinique de p53 dans les cancers du sein. Nous pouvons imaginer que dans certaines études, des tumeurs classées comme « P53 négative » et donc *TP53* wild-type sont en réalité des tumeurs où le gène *TP53* est muté puisqu'il a été montré que la formation d'isoformes courtes et mutantes de p53 ayant acquis des fonctions oncogéniques peuvent ne pas être détecté par IHC. **Nos résultats rajoutant encore une complexité supplémentaire à cette discussion, la Figure 38 illustre mon argumentation.**



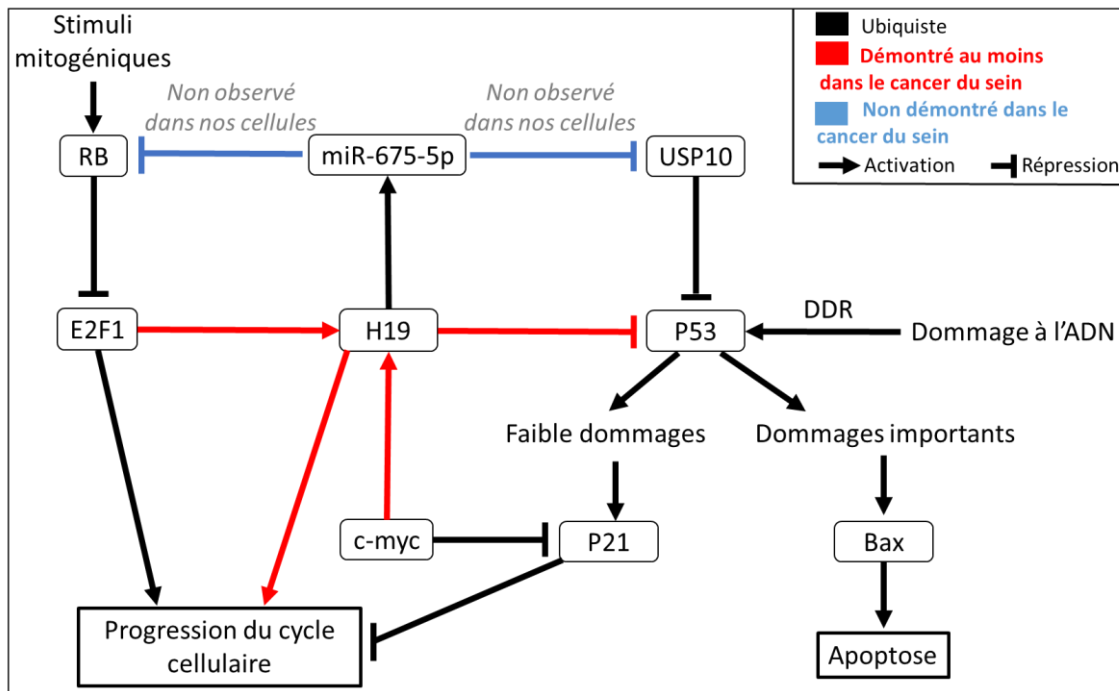
**Figure 38 : Hétérogénéité des tumeurs « p53 négatif / p53 positif » identifiées par immunohistochimie**

En effet, nous avons montré que dans des cellules comportant un TP53 wild-type (MCF-7), la surexpression d'*H19* est capable de diminuer l'expression protéique de p53 et son activité. Deux tissus cancéreux classés comme « p53 négatif » par IHC et donc considérés comme TP53 wild-type avec une protéine p53 fonctionnelle, pourraient refléter une activité de la protéine p53 différente en fonction de la surexpression du gène *H19*. En effet, si le lncRNA *H19* est surexprimé, nos travaux suggèrent que la protéine p53 est inactivée par *H19*. **Les tissus « p53 négatif » identifiés par IHC, mais qui surexpriment *H19* seraient donc des tissus où la protéine p53 wild-type ne joue plus sa fonction de suppresseur de tumeurs.** De plus, d'autres mécanismes sont décrits comme pouvant inactiver la fonction de la protéine p53 et pourraient également expliquer que deux tissus « p53 négatif » soient différents en termes d'activité de p53. Cette observation pourrait justifier la diversité des résultats obtenus par les études qui ont pour but de démontrer le rôle des mutations de *TP53* dans le diagnostic des cancers du sein.



## I.5. Rôle central de *H19* dans la régulation du cycle cellulaire

L'implication du lncRNA *H19* dans la régulation du cycle cellulaire est bien établie dans la littérature et résumé dans la Figure 39.



**Figure 39 : Régulation du cycle cellulaire par *H19***

En 2005, notre équipe a démontré que la surexpression du gène *H19* permet de faciliter la transition G1/S alors que sa downregulation par siRNA affecte l'entrée en phase S et la prolifération (Bertheaux et al., 2005). Par ailleurs, dans les cellules cancéreuses mammaires l'expression d'*H19* est activée par la fixation d'E2F1 (facteur clé dans la transition G1/S) sur deux sites consensus au sein du promoteur de *H19* et négativement régulée par la protéine RB. De manière intéressante, *H19* est capable de downreguler l'expression de RB via le miR-675-5p afin d'augmenter la croissance cellulaire. Cette observation réalisée dans des lignées de cancers colorectaux n'a pas été constatée dans nos lignées (Vennin et al., 2015). Baryte-Lovejoy et al. ont démontré que l'oncogène c-Myc se fixe spécifiquement sur l'allèle maternel d'*H19* pour promouvoir sa transcription, ce qui mène à une prolifération accrue des cellules cancéreuses mammaires et pulmonaires (Baryte-Lovejoy et al., 2006). Nos résultats ont montré qu'*H19* est capable de réprimer l'arrêt dans le cycle induit par l'activation de p53 dans les cellules cancéreuses mammaires. En effet, lorsque les cellules MCF-7 (préalablement synchronisées) sont traitées avec de la doxorubicine, nous avons observé que la surexpression d'*H19* diminue l'arrêt dans le cycle en phase G1/S induit par p53. Bien que la régulation de la protéine p53 par *H19* avait déjà été démontrée dans les cancers gastriques par Yang et al.,

nous sommes les premiers à avoir découvert cette régulation dans le cancer du sein (Yang et al., 2012). De plus, notre étude a permis d'identifier les mécanismes moléculaires impliqués dans cette régulation. En 1998, notre équipe a démontré que la protéine p53 est capable de réprimer l'expression du gène *H19* (Dugimont et al., 1998). L'étape à laquelle intervient la surexpression du gène *H19* dans les cancers du sein n'est pas connue à ce jour, cependant nous savons que les mutations du gène p53 ont tendance à être un élément précoce de la cancérisation. En effet, il a été démontré que les altérations de TP53 sont importantes dans les étapes précoces du développement cancéreux de la glande mammaires (Milićević et al., 2014). **Nous pouvons donc émettre l'hypothèse que la perte de fonctionnalité de p53 dans une cellule cancéreuse au début d'un processus de cancérisation puisse favoriser la surexpression du gène *H19* et amplifier le processus. A contrario, si une cellule possédant une protéine p53 fonctionnelle commence à surexprimer le gène *H19*, le lncRNA *H19* pourrait impacter l'action de p53 et favoriser l'instabilité génomique et un processus de cancérisation.** En plus du rôle du lncRNA *H19* dans la régulation du cycle cellulaire, il a été démontré que le miR-675 régulait négativement la protéine p53 en ciblant USP10 (protéine qui régule la localisation et la stabilité de p53) dans des cellules progénitrices cardiaques (Cai et al., 2016). Cependant, dans nos cellules nous n'avons pas retrouvé de régulation de USP10 (données non montrées). **L'ensemble de ces résultats démontre que le gène *H19* joue un rôle important dans la régulation du cycle cellulaire dans le cancer.**

**Notre étude a démontré que le lncRNA *H19* inactive la protéine p53 dans les cellules cancéreuses mammaires. Cette régulation fait intervenir la formation d'un complexe tripartite entre *H19*-p53-MDM2, jamais décrit dans la littérature jusqu'à présent. Comme nous l'avons vu, la découverte de ce mécanisme pourrait remettre en perspective les résultats contradictoires obtenus par les études de l'implication du statut mutationnel de p53, observé par IHC, dans les cancers du sein. L'implication du lncRNA *H19* dans l'instabilité génomique démontrée dans cette étude a donné naissance à un nouveau projet au sein du laboratoire. Mon successeur, Clément Lecerf travaille actuellement sur les mécanismes moléculaires impliqués dans la régulation de l'instabilité génomique par *H19*.**

## **II. Le lncRNA *H19* a des actions dépendantes et indépendantes de son microARN dans les cancers**

### **II.1. Le lncRNA *H19* et son microARN sont impliqués dans l'émergence et le maintien des cellules souches cancéreuses mammaires**

Mes travaux montrent que le lncRNA *H19* et le miR-675-5p sont importants dans le maintien des propriétés souches des cellules souches cancéreuses (CSC) mammaires. Au commencement de mon projet, nous étions les seuls à avoir émis l'hypothèse d'un rôle du lncRNA *H19* dans la régulation des CSC mammaires. Depuis, quelques études ont démontré l'implication du lncRNA *H19* dans cette régulation. Peng et collaborateurs ont montré que le lncRNA *H19* est surexprimé dans des populations enrichies en CSC mammaires (Peng et al., 2017). Le mécanisme d'action d'*H19* dans cette régulation fait intervenir la fixation du miR let-7 par *H19*, qui *via* un effet éponge va empêcher l'action de let-7 sur ses cibles. Ce miR est décrit comme suppresseur de tumeur et a notamment un effet négatif sur les populations de CSC. Une autre étude a montré que par ce même mécanisme, *H19* est capable de stimuler les divisions symétriques des CSC mammaires, ce qui augmente leur auto-renouvellement. Shima et al. ont également démontré que la surexpression du lncRNA *H19* est associée au phénotype souches dans les cancers ALDH1 positif (Shima et al., 2018).

J'ai confirmé l'implication du lncRNA *H19* dans les CSC mammaires via l'utilisation de deux siRNAs ciblant *H19*. Cependant, la nouveauté de mon étude est l'implication du miR-675-5p dans les CSC mammaires. En effet, le ciblage du miR-675-5p par un antagomir ou sa surexpression *via* un mimic permet de réguler les phénotypes associés aux CSC. A l'heure actuelle, aucune étude ne montre le rôle du miR-675 dans la régulation des CSC. J'ai également utilisé une approche innovante pour étudier le phénotype des cellules qui expriment *H19* dans la population globale des cellules cancéreuses en culture. En effet, grâce à un système rapporteur de l'expression du gène *H19*, j'ai montré que les cellules où le promoteur d'*H19* est actif sont des cellules avec un phénotype souche plus important. Dans la littérature il est connu que les CSC sont fortement liées à la transition épithélio-mésenchymateuse (EMT), nous avons décidé d'étudier l'implication d'*H19* et de son miR ce processus. Les résultats préliminaires de cette étude, réalisée par Evodie Peperstraete, montre que la surexpression d'*H19* et de son miR augmente les capacités de migration et d'invasion des cellules cancéreuses mammaires. Un autre processus pathologiques associés aux CSC est la résistance aux thérapies. Nous savons également qu'*H19* participe à la résistance des cellules cancéreuses à de nombreuses

chimiothérapies. Nous allons donc étudier l'impact de la modulation du lncRNA *H19* et de son miR sur la chimiorésistance. Je vais participer à cette étude jusqu'à mon départ du laboratoire et ces travaux seront terminés par Evodie par la suite.

**Ce projet nous a donc permis de montrer l'implication du lncRNA *H19* et du miR-675 dans la biologie des CSC mammaires. La poursuite de ce projet permettra d'éclaircir les mécanismes d'actions mis en jeu dans cette régulation.**

## **II.2. Identification de la part relative entre le lncRNA *H19* et son microARN.**

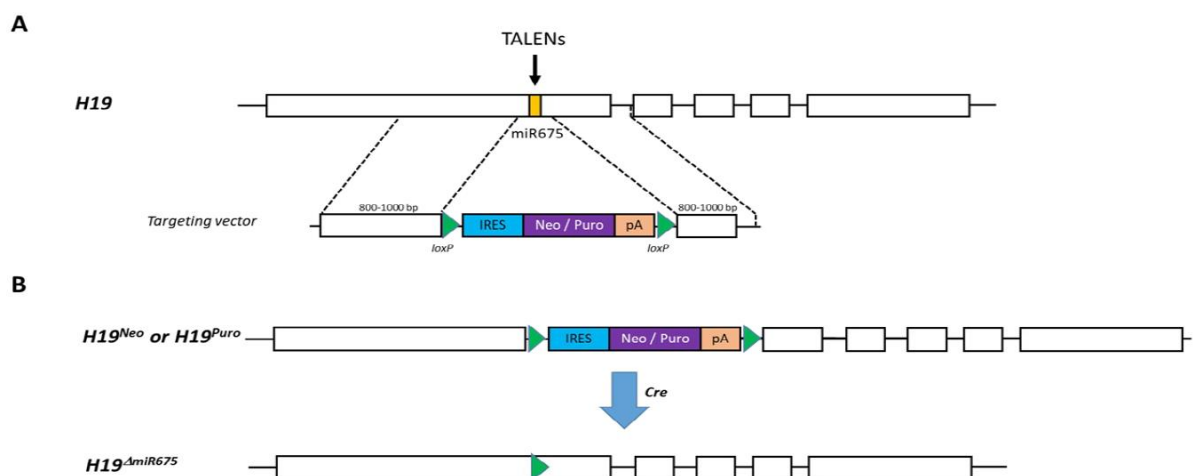
Depuis la découverte de la production d'un microARN à partir de l'exon 1 du gène *H19* en 2007, nous savons que les effets biologiques médiés par *H19* sont en partie dus à ce microARN. Dans le cancer du sein, notre équipe a montré que le miR-675-5p cible les protéines Cbl-b et c-Cbl pour accroître la prolifération cellulaire et les capacités migratoires des cellules *in vivo*. Mes travaux de thèse ont également démontré qu'*H19* va agir en tant que lncRNA en servant d'échafaud à l'interaction p53-mdm2. De nombreux autres résultats de la littérature montrent donc que le gène *H19* a une action via son lncRNA et/ou son microARN. Cependant, il est difficile d'étudier la part relative du lncRNA et du microARN produit à partir d'*H19*.

**Plusieurs approches sont déjà utilisées dans notre laboratoire pour moduler l'expression du lncRNA *H19* et/ou du miR-675 :**

- La surexpression ectopique stable (ou transitoire) du gène codant *H19* grâce à un plasmide, qui va induire la surexpression du lncRNA et du miR-675.
- La surexpression ectopique du miR-675 grâce à un plasmide, qui va induire une surexpression du miR-675 sans modifier l'expression du lncRNA *H19*.
- L'utilisation de siRNAs ciblant le lncRNA *H19*, qui va impacter l'expression du miR-675 également
- L'utilisation de mimic ou d'antagomir pour moduler la disponibilité du miR-675, sans modifier l'expression du lncRNA *H19*.
- L'utilisation d'un vecteur shRNA ciblant le lncRNA *H19* est également envisageable, cependant, l'expression du miR-675 serait également impactée.

Cependant, ces approches ne nous permettent pas de différencier l'implication du lncRNA *H19* et/ou du miR-675 de manière parfaite. En effet, ces méthodes permettent d'étudier l'effet à long terme de la surexpression ectopique du miR-675 sans moduler *H19*, cependant, il est impossible de mesurer l'impact de l'absence du miR-675 au long terme, sans modifier

l'expression du lncRNA *H19*. Pour contrer ce problème, il serait possible de réaliser une mutagenèse dirigée sur le plasmide de surexpression du gène *H19* que nous utilisons, afin de détruire le site de synthèse du miR-675 ou sa *seed* sequence. Cette approche nous permettrait de voir l'effet de la surexpression du lncRNA *H19* sans moduler l'expression du miR-675. Cependant, ces approches sont des approches de surexpressions exogènes et il serait intéressant de pouvoir moduler l'expression du miR-675 sans moduler l'expression endogène du lncRNA *H19*. Une stratégie a été développée au laboratoire pour travailler sur des cellules portant des allèles du gène *H19* déficient en miR-675. Ces cellules nous permettront d'étudier l'impact de la perte d'expression du miR-675 dans les cellules cancéreuses mammaires, sans moduler artificiellement l'expression du lncRNA *H19* et de discriminer l'action du gène *H19* en tant que lncRNA de son action en tant que miR-675. Cette stratégie est illustrée sur la Figure 40.



**Figure 40 : Génération d'un allèle *H19* déficient en miR-675.**

(A) Représentation schématique du locus *H19* et de la séquence de ciblage. La localisation de la séquence de reconnaissance de TALEN au miR-675 (jaune) dans l'exon 1 est indiquée. La cassette de ciblage, flanquée de sites loxP (triangles verts) est composée d'un site d'entrée de ribosome interne (IRES), d'un gène de résistance aux antibiotiques (Neo ou Puro) et d'une séquence de polyadénylation (pA). (B) L'insertion de la cassette entraîne l'insertion du gène de résistance dans l'allèle *H19* conduisant à un arrêt prématuré dans sa transcription. L'inversion du Knock-out est obtenue par l'excision engendrée par Cre du gène marqueur sélectionnable conduisant à l'obtention d'un allèle *H19* déficient en miR-675.

Pour cela, une cassette sans promoteur composée du site « internal ribosome entry site » (IRES) de l'EMCV, un gène de résistance aux antibiotiques (Neo ou Puro) et la séquence de polyadénylation SV40 flanquée de sites loxP et de bras d'homologie avec le gène *H19* seront ciblés sur la séquence du miR-675 préalablement coupée par TALENs dans le gène *H19*. Comme indiqué sur la Figure 38.A, la réparation par homologie du gène *H19* ciblé par les

TALENs, sera responsable de la délétion de la séquence miR-675 et de l'insertion de la cassette sélectionnable permettant l'identification des clones résistants modifiés. Ces clones abriteront un knockout d'un allèle *H19*. L'utilisation à la fois d'une cassette contenant Neo et d'une cassette Puro permettra de cibler les deux allèles *H19*. Ensuite, l'élimination de la cassette sélectionnable par l'expression transitoire du vecteur exprimant Cre va générer un allèle *H19* déficient en miR-675 (Figure 38.B). Les expériences seront effectuées sur des cellules MDA-MB-231 et T47D qui possèdent deux allèles du locus *H19*. L'obtention de ces lignées cellulaires nous permettra de pouvoir réaliser l'étude de la part relative du lncRNA *H19* et de son microARN. Cette stratégie s'intégrera dans un projet plus grand d'analyse transcriptomique de nos lignées cellulaires et les résultats obtenus dans ces lignées seront comparés aux lignées cellulaires utilisant d'autres stratégies afin d'identifier clairement le rôle du lncRNA *H19* et de son microARN dans les cancers du sein.

### **II.3. Identifications des gènes cibles du lncRNA *H19* et du miR-675**

Comme nous l'avons vu dans la troisième partie des résultats, il est très difficile d'identifier *in silico* de nouvelles cibles valides du miR-675. Nous avons donc réalisé une approche de séquençage haut débit des ARNm de cellules surexprimant *H19* afin d'identifier les gènes régulés par *H19* et son microARN. Cette approche nous a déjà permis d'identifier de nombreux gènes surexprimés qui font partir de voies de signalisations communes (Signalisation du TNF, MAPK, PI3K-Akt...). Nous avons également identifié de nombreux gènes de la famille HOXC régulés dans nos cellules. Il serait intéressant d'étudier les mécanismes associés à cette régulation. Par ailleurs, j'ai débuté la comparaison entre les cibles du miR-675 retrouvées *in silico* et nos gènes cibles identifiés par RNAseq. La poursuite de cette étude devrait nous permettre de trouver des cibles intéressantes du miR-675 dans nos cellules.

**Pour conclure, l'ensemble de mes travaux de thèse ont permis d'identifier de nouveaux mécanismes d'actions du lncRNA *H19* et du miR-675 dans les cancers du sein. Le lncRNA *H19* et son microARN régulent de nombreux aspects de la tumorigenèse mammaire et possèdent un rôle oncogénique indiscutable dans les cancers du sein.**

# Bibliographie





## BIBLIOGRAPHIE

- Adriaenssens, E, Lottin, S., Dugimont, T., Fauquette, W., Coll, J., Dupouy, J. P., & Boilly, B. (1999a). *Steroid hormones modulate H19 gene expression in both mammary gland and uterus*. *18*, 4460–4473.
- Adriaenssens, E, Lottin, S., Dugimont, T., Fauquette, W., Coll, J., Dupouy, J. P., & Boilly, B. (1999b). *Steroid hormones modulate H19 gene expression in both mammary gland and uterus*. 4460–4473.
- Adriaenssens, Eric, Dumont, L., Lottin, S., Bolle, D., Leprêtre, A., Delobelle, A., ... Curgy, 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*(5), 1597–1607.
- Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morisson, S. J., & Clarke, M. F. (2003). *Prospective identification of tumorigenic breast cancer cells*. *100*(11).
- Angrand, P. O., Vennin, C., Le Bourhis, X., & Adriaenssens, E. (2015). The role of long non-coding RNAs in genome formatting and expression. *Frontiers in Genetics*, *6*(APR), 1–12.
- Antoniou, A., Pharoah, P. D. P., Narod, S., Risch, H. A., Eyfjord, J. E., Hopper, J. L., ... Easton, D. F. (2003). Average Risks of Breast and Ovarian Cancer Associated with BRCA1 or BRCA2 Mutations Detected in Case Series Unselected for Family History: A Combined Analysis of 22 Studies. *American Journal of Human Genetics*, *72*, 1117–1130.
- Ariel, I., Lustig, O., Tamar, S., D, G. P. M., Mally, S., Nathan, D., & Abraham, H. (1995). The imprinted H19 gene as a tumor marker in bladder carcinoma. *Urology*, *45*(2), 335–338.
- Ariel, I., Weinstein, D., Voutilainen, R., Tamar, S., Lustig-Yariv, O., Nathan, D., & Hochberg, A. (1997). *Genomic Imprinting and the Endometrial Cycle. The Expression of the Imprinted Gene H19 in the Humal Femal Reproductive Organs*. (pp. 17–25). pp. 17–25.
- Aubrey, B. J., Kelly, G. L., Janic, A., Herold, M. J., & Strasser, A. (2017). How does p53 induce apoptosis and how does this relate to p53-mediated tumour suppression? *Nature Publishing Group*, *25*(1), 104–113.
- Babiarz, J. E., Ruby, J. G., Wang, Y., Bartel, D. P., & Blelloch, R. (2008). Mouse ES cells express endogenous shRNAs, siRNAs, and other microprocessor-independent, dicer-dependent small RNAs. *Genes and Development*, *22*(20), 2773–2785. <https://doi.org/10.1101/gad.1705308>
- Balzeau, J., Menezes, M. R., Cao, S., & Hagan, J. P. (2017). The LIN28/let-7 pathway in cancer. *Frontiers in Genetics*, *8*(MAR), 1–16.
- Barsyte-Lovejoy, D., Lau, S. K., Boutros, P. C., Khosravi, F., Jurisica, I., Andrulis, I. L., ... Penn, L. Z. (2006). The c-Myc oncogene directly induces the H19 noncoding RNA by allele-specific binding to potentiate tumorigenesis. *Cancer Research*, *66*(10), 5330–5337.
- Basak, P., Chatterjee, S., Weger, S., Bruce, M. C., Murphy, L. C., & Raouf, A. (2015). Estrogen regulates luminal progenitor cell differentiation through H19 gene expression. *Endocrine-Related Cancer*, *22*(4), 505–517.
- Battle, E., & Clevers, H. (2017). Cancer stem cells revisited. *Nature Medicine*, *23*(10), 1124–1134.
- Bauderlique-Le Roy, H., Vennin, C., Brocqueville, G., Spruyt, N., Adriaenssens, E., & 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 and Development*, *24*(10), 1252–1262.
- Bell Daphne W., Varley Jennifer M., Szydlo Tara E., Deborah H. Kang, \* D. C. R. W., \* K. E. S., ... Daniel A. Haber. (1999). Heterozygous Germ Line hCHK2 Mutations in Li-Fraumeni Syndrome. *Science*, *286*(5449), 2528–2531.
- Bellazzo, A., Sicari, D., Valentino, E., Sal, G. Del, & Collavin, L. (2018). Breast Cancer-Targets and Therapy Dovepress Complexes formed by mutant p53 and their roles in breast cancer. *Breast Cancer-Targets and Therapy*, 10–101.

- Berteaux, Nathalie; Séverine, L., Monte, D., Quatannens, B., Coll, J., Hondermarck, H., Cury, J., ... Adriaenssens, E. (2005). *H19 mRNA-like Noncoding RNA Promotes Breast Cancer Cell Proliferation through Positive Control by E2F1*. *280(33)*, 29625–29636.
- Berteaux, N., Aptel, N., Cathala, G., Genton, C., Coll, J., Daccache, A., ... Adriaenssens, E. (2008). A Novel H19 Antisense RNA Overexpressed in Breast Cancer Contributes to Paternal IGF2 Expression. *Molecular and Cellular Biology*, *28(22)*, 6731–6745.
- Berteaux, N., Lottin, S., Adriaenssens, E., Van Coppennolle, F., Leroy, X., Coll, J., ... Cury, J. J. (2004). Hormonal regulation of H19 gene expression in prostate epithelial cells. *Journal of Endocrinology*, *183(1)*, 69–78.
- Berteaux, Nathalie, Lottin, S., Monté, D., Pinte, S., Quatannens, B., Coll, J., ... Adriaenssens, E. (2005). H19 mRNA-like noncoding RNA promotes breast cancer cell proliferation through positive control by E2F1. *Journal of Biological Chemistry*, *280(33)*, 29625–29636.
- Bertheau, P., Lehmann-che, J., Varna, M., Dumay, A., Poirot, B., Porcher, R., ... Thé, H. De. (2013). p53 in breast cancer subtypes and new insights into response to chemotherapy. *The Breast*, *22*, S27–S29.
- Bhatia, M., Wang, J. C., Kapp, U., Bonnet, D., & Dick, J. E. (1997). Purification of primitive human hematopoietic cells capable of repopulating immune-deficient mice. *Proceedings of the National Academy of Sciences of the United States of America*, *94(10)*, 5320–5325.
- Binder-Foucard, A. F., Belot, A., Delafosse, P., Remontet, L., Woronoff, A.-S., & Bossard, N. (2013). Estimation nationale de l'incidence et de la mortalité par cancer en France Mots-clés : incidence, mortalité, cancer, France, registres de population. *Rapport INCA*.
- Brannan, C. I., Dees, E. C., Ingram, R. S., & Tilghman, S. M. (1990). The product of the H19 gene may function as an RNA. *Molecular and Cellular Biology*, *10(1)*, 28–36.
- Burch, L. R. Y., Midgley, C. A., Currie, R. A., Lane, D. P., & Hupp, T. R. (2000). *Mdm2 binding to a conformationally sensitive domain on p53 can be modulated by RNA*. *472*, 93–98.
- Cai, B., Ma, W., Bi, C., Yang, F., Zhang, L., Han, Z., ... Lu, Y. (2016). Long noncoding RNA H19 mediates melatonin inhibition of premature senescence of c-kit + cardiac progenitor cells by promoting miR-675. *Journal of Pineal Research*, 82–95.
- Cai, X., & Cullen, B. R. (2007). The imprinted H19 noncoding RNA is a primary microRNA precursor. *Rna*, *13(3)*, 313–316.
- Chao, C. C. (2014). Mechanisms of p53 degradation. *Clinica Chimica Acta*.
- Chen, S., Bu, D., Ma, Y., Zhu, J., Chen, G., Sun, L., ... Wang, P. (2017). H19 Overexpression Induces Resistance to 1,25(OH)2D3 by Targeting VDR Through miR-675-5p in Colon Cancer Cells. *Neoplasia (United States)*, *19(3)*, 226–236.
- Collette, J., Le Bourhis, X., & Adriaenssens, E. (2017). Regulation of Human Breast Cancer by the Long Non-Coding RNA H19. *International Journal of Molecular Sciences*.
- Cordero, F., Ferrero, G., Polidoro, S., Fiorito, G., Campanella, G., Sacerdote, C., ... Naccarati, A. (2015). Differentially methylated microRNAs in prediagnostic samples of subjects who developed breast cancer in the european prospective investigation into nutrition and cancer (EPIC-Italy) cohort. *Carcinogenesis*, *36(10)*, 1144–1153.
- Cui, P., Zhao, Y., Chu, X., He, N., Zheng, H., Han, J., ... Chen, K. (2018). SNP rs2071095 in lincRNA H19 is associated with breast cancer risk. *Breast Cancer Research and Treatment*, *171(1)*, 161–171.
- Deleo, A. B., Jayt, G., Appellat, E., Duboist, G. C., Lawt, L. W., & Old, L. J. (1979). *Detection of a transformation-related antigen in chemically induced sarcomas and other transformed cells of the mouse Immunology : Abibro-*. *76(5)*, 2420–2424.
- Dippold, W. G., Jayt, G., Deleo, A. B., Khouryt, G., & Old, L. J. (1981). *p53 transformation-related protein : Detection by monoclonal antibody in mouse and human cells*. *78(3)*, 1695–1699.
- Dragomir, M., Mafra, A. C. P., Dias, S. M. G., Vasilescu, C., & Calin, G. A. (2018). Using microRNA networks to understand cancer. *International Journal of Molecular Sciences*, *19(7)*, 1–18.
- Dugimont, T., Montpellier, C., Adriaenssens, E., Lottin, S., Dumont, L., lotsova, V., ... Cury, J. J. (1998). The H19 TATA-less promoter is efficiently repressed by wild-type tumor suppressor gene product

- p53. *Oncogene*, 16(18), 2395–2401.
- El Hajj, J., Nguyen, E., Liu, Q., Bouyer, C., Adriaenssens, E., Hilal, G., & Ségal-Bendirdjian, E. (2018). Telomerase regulation by the long non-coding RNA H19 in human acute promyelocytic leukemia cells. *Molecular Cancer*, 17(1), 1–13.
- Elenbaas, B., Dobbstein, M., Roth, J., Shenk, T., & Levine, A. J. (1996). *The MDM2 Oncoprotein Binds Specifically to RNA through Its RING Finger Domain*. 2(4).
- Eyfiã, J. E., Steinarsdottir, M., Anamthawat-jonsson, K., & Helga, M. Æ. (1995). *p53 Abnormalities and Genomic Instability in Primary Human Breast Carcinomas*. 646–651.
- Farmer, P., Bonnefoi, H., Becette, V., Tubiana-Hulin, M., Fumoleau, P., Larsimont, D., ... Iggo, R. (2005). Identification of molecular apocrine breast tumours by microarray analysis. *Oncogene*, 24(29), 4660–4671.
- Feng, Y., Spezia, M., Huang, S., Yuan, C., Zeng, Z., Zhang, L., ... Ren, G. (2018). Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. *Genes and Diseases*, 5(2), 77–106.
- Freeman, J. A., & Espinosa, J. M. (2012). *The impact of post-transcriptional regulation in the p53 network*. 12(1), 46–57.
- Fukuzawa, R., Umezawa, A., Ochi, K., Urano, F., Ikeda, H., & Hata, J. I. (1999). High frequency of inactivation of the imprinted H19 gene in “sporadic” hepatoblastoma. *International Journal of Cancer*, 82(4), 490–497.
- Gao, H., Hao, G., Sun, Y., Li, L., & Wang, Y. (2018). Long noncoding RNA H19 mediated the chemosensitivity of breast cancer cells via wnt pathway and EMT process. *OncoTargets and Therapy*, 11, 8001–8012.
- Ge, L., Zhang, X., Hu, S., Song, Y., Kong, J., Zhang, B., & Yang, X. (2019). H19 suppresses the growth of hepatoblastoma cells by promoting their apoptosis via the signaling pathways of miR-675/FADD and miR-138/PTK2. *Journal of Cellular Biochemistry*, 120(4), 5218–5231.
- Ginestier, C., Hur, M. H., Charafe-Jauffret, E., Monville, F., Dutcher, J., Brown, M., ... Dontu, G. (2007). ALDH1 Is a Marker of Normal and Malignant Human Mammary Stem Cells and a Predictor of Poor Clinical Outcome. *Cell Stem Cell*, 1(5), 555–567.
- Goehe, R. W., Di, X., Sharma, K., Bristol, M. L., Henderson, S. C., Valerie, K., ... Gewirtz, D. A. (2012). *The Autophagy-Senescence Connection in Chemotherapy : Must Tumor Cells ( Self ) Eat Before They Sleep ?* □. 343(3), 763–778.
- Haiyan, S., Ping, C., Hongtao, L., & Xiang, W. (2019). Long non-coding RNA H19 regulates cell growth and metastasis via the miR-22-3p/Snail1 axis in gastric cancer. *International Journal of Oncology*, 54(6), 2157–2168.
- Hammond, S. M. (2015). An overview of microRNAs. *Adv Drug Deliv Rev.*, 87, 3–14.
- Hanahan, D., & Weinberg, R. A. (2000). The Hallmarks of Cancer. *Netherlands Heart Journal*, 100, 57–70.
- Hanahan, D., & Weinberg, R. A. (2011). Review Hallmarks of Cancer : The Next Generation. *Cell*, 144(5), 646–674.
- Hashad, D., Elbanna, A., Ibrahim, A., & Khedr, G. (2016). Evaluation of the Role of Circulating Long Non-Coding RNA H19 as a Promising Novel Biomarker in Plasma of Patients with Gastric Cancer. *Journal of Clinical Laboratory Analysis*, 30(6), 1100–1105.
- Hayes, J., Peruzzi, P. P., & Lawler, S. (2014). MicroRNAs in cancer : biomarkers , functions and therapy. *Trends in Molecular Medicine*, 20(8), 460–469.
- He, D., Wang, J., Zhang, C., Shan, B., Deng, X., Li, B., ... Duan, C. (2015). Down-regulation of miR-675-5p contributes to tumor progression and development by targeting pro-tumorigenic GPR55 in non-small cell lung cancer. *Molecular Cancer*, 14(1), 1–14.
- Hernandez, J. M., Elahi, A., Clark, C. W., Wang, J., Humphries, L. A., Centeno, B., ... Shibata, D. (2013). MiR-675 mediates downregulation of twist1 and rb in afp-secreting hepatocellular carcinoma. *Annals of Surgical Oncology*, 20(3 SUPPL.), 625–635.
- Higgs, P. G., & Lehman, N. (2015). The RNA World: Molecular cooperation at the origins of life. *Nature*

*Reviews Genetics*, 16(1), 7–17.

- Hu, B., Weng, Y., Xia, X., Liang, X., & Huang, Y. (2019). Clinical advances of siRNA therapeutics. *The Journal of Gene Medicine*, e3097.
- Huang, Z., Du, G., Huang, X., Han, L., Han, X., Xu, B., ... Lu, C. (2018). The enhancer RNA Inc-SLC4A1-1 epigenetically regulates unexplained recurrent pregnancy loss (URPL) by activating CXCL8 and NF- $\kappa$ B pathway. *EBioMedicine*, 38, 162–170.
- Hui, L., Zheng, Y., Yan, Y., Bargonetti, J., & Foster, D. A. (2006). Mutant p53 in MDA-MB-231 breast cancer cells is stabilized by elevated phospholipase D activity and contributes to survival signals generated by phospholipase D. *Oncogene*, 25(55), 7305–7310.
- Hur, H., Lee, J. Y., Yun, H. J., Park, B. W., & Kim, M. H. (2014). Analysis of HOX gene expression patterns in human breast cancer. *Molecular Biotechnology*, 56(1), 64–71.
- Imig, J., Brunschweiler, A., Brümmer, A., Guennewig, B., Mittal, N., Kishore, S., ... Hall, J. (2015). MiR-CLIP capture of a miRNA targetome uncovers a lincRNA H19-miR-106a interaction. *Nature Chemical Biology*, 11(2), 107–114.
- Jiao, Z.-Y., Tian, Q., Li, N., Wang, H.-B., & Li, K.-Z. (2018). Plasma long non-coding RNAs (lncRNAs) serve as potential biomarkers for predicting breast cancer. *European Review for Medical and Pharmacological Sciences*, 22(7), 1994–1999.
- Kreso, A., & Dick, J. E. (2014). Evolution of the cancer stem cell model. *Cell Stem Cell*, 14(3), 275–291.
- Kung, J. T. Y., Colognori, D., & Lee, J. T. (2013). Long Noncoding RNAs: Past, Present, and Future. *Genetics*, 193(March), 651–669.
- Lanz, R. B., Mckenna, N. J., Onate, S. A., Albrecht, U., Wong, J., Tsai, S. Y., ... This, P. R. (1999). A Steroid Receptor Coactivator, SRA, Functions as an RNA and Is Present in an SRC-1 Complex. 97, 17–27.
- Lavie, O., Edelman, D., Levy, T., Fishman, A., Hubert, A., Segev, Y., ... Hochberg, A. (2017). A phase 1/2a, dose-escalation, safety, pharmacokinetic, and preliminary efficacy study of intraperitoneal administration of BC-819 (H19-DTA) in subjects with recurrent ovarian/peritoneal cancer. *Archives of Gynecology and Obstetrics*, 295(3), 751–761.
- Lecerf, C., Le Bourhis, X., & Adriaenssens, E. (2019). The long non-coding RNA H19: an active player with multiple facets to sustain the hallmarks of cancer. *Cellular and Molecular Life Sciences*, (0123456789).
- Leighton, P. A., Saam, J. R., Ingram, R. S., Stewart, C. L., & Tilghman, S. M. (1995). An enhancer deletion affects both *H19* and *Igf2* expression. *Genes & Development*, 9(17), 2079–2089.
- Li, J. P., Xiang, Y., Fan, L. J., Yao, A., Li, H., & Liao, X. H. (2019). Long noncoding RNA H19 competitively binds miR-93-5p to regulate STAT3 expression in breast cancer. *Journal of Cellular Biochemistry*, 120(3), 3137–3148.
- Li, Z., Li, Y., Li, Y., Ren, K., Li, X., Han, X., & Wang, J. (2017). Long non-coding RNA H19 promotes the proliferation and invasion of breast cancer through upregulating DNMT1 expression by sponging miR-152. *Journal of Biochemical and Molecular Toxicology*, 31(9).
- Liu, J., Kahri, A. I., Heikkilä, P., Ilvesmäki, V., & Voutilainen, R. (1995). H19 and Insulin-like Growth Factor-II Gene Expression in Adrenal Tumors and Cultured Adrenal Cells. *Journal of Clinical Endocrinology and Metabolism*, (November), 492–496.
- Liu, L. C., Wang, Y. L., Lin, P. Le, Zhang, X., Cheng, W. C., Liu, S. H., ... Wang, S. C. (2019). Long noncoding RNA HOTAIR promotes invasion of breast cancer cells through chondroitin sulfotransferase CHST15. *International Journal of Cancer*, 00(6), 1–11.
- Liu, L., Yang, J., Zhu, X., Li, D., Lv, Z., & Zhang, X. (2016). Long noncoding RNA H19 competitively binds miR-17-5p to regulate YES1 expression in thyroid cancer. *FEBS Journal*, 283, 2326–2339.
- Lottin, S., Adriaenssens, E., Dupressoir, T., Berteaux, N., Montpellier, C., Coll, J., ... Cury, J. J. (2002). Overexpression of an ectopic H19 gene enhances the tumorigenic properties of breast cancer cells the phenotype of human breast epithelial cells, we investi- expressed from MDA-MB-231 breast cancer cells stably. *Carcinogenesis*, 23(11), 1885–1895.
- Luo, Q., & Chen, Y. (2016). Long noncoding RNAs and Alzheimer's disease. *Clinical Interventions in Aging*, 867–872.

- Marise S. Bartolomei, Sharon Zemel, S. M. T. (1991). Parental imprinting of the mouse H19 gene. *Nature*, *351*, 153–155.
- Milićević, Z., Bajić, V., Živković, L., Kasapović, J., Andjelković, U., & Spremo-Potparević, B. (2014). Identification of p53 and Its Isoforms in Human Breast Carcinoma Cells. *The Scientific World Journal*, *2014*, 1–10.
- Mott, J. L., & Mohr, A. M. (2015). Overview of microRNA biology. *Seminars in Liver Disease*, *35*(1), 3–11.
- Nakada, H., Nakagomi, H., Hirotsu, Y., Amemiya, K., Mochizuki, H., Inoue, M., ... Omata, M. (2017). A study of tumor heterogeneity in a case with breast cancer. *Breast Cancer*, *24*(3), 483–489.
- Ohana, P., Bibi, O., Matouk, I., Levy, C., Birman, T., Ariel, I., ... Hochberg, A. (2002). Use of H19 regulatory sequences for targeted gene therapy in cancer. *International Journal of Cancer*, *98*(5), 645–650.
- Okutsu, T., Kuroiwa, Y., Kagitani, F., & Kai, M. (2000). Expression and Imprinting Status of Human Paternally Expressed Antisense Transcript from the IGF2 Wilms' Tumors. *J. Biochem*, *127*(3), 475–483.
- Pachnis, V., Belayew, A., & 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*, *81*(17), 5523–5527.
- Palazzo, A. F., & Lee, E. S. (2015). Non-coding RNA: What is functional and what is junk? *Frontiers in Genetics*, *5*(JAN), 1–11.
- Peng, F., Li, T. T., Wang, K. L., Xiao, G. Q., Wang, J. H., Zhao, H. D., ... Liu, Q. (2017). H19/let-7/LIN28 reciprocal negative regulatory circuit promotes breast cancer stem cell maintenance. *Cell Death and Disease*, *8*(1).
- Perou, C. M., Sùrlie, T., Eisen, M. B., Rijn, M. Van De, Jeffrey, S. S., Rees, C. A., ... Grant, S. (2000). *Molecular portraits of human breast tumours*. *533*(May), 747–752.
- Pohlmann, P. R., Mayer, I. A., & Mernaugh, R. (2009). Resistance to trastuzumab in breast cancer. *Clinical Cancer Research*, *15*(24), 7479–7491.
- Poirier, F., Chan, C. T., Timmons, P. M., Robertson, E. J., Evans, M. J., & 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 (Cambridge, England)*, *113*(4), 1105–1114.
- Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W. J., & Pandolfi, P. P. (2010). A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature*, *465*(7301), 1033–1038.
- Prat, A., Fan, C., Karginova, O., Perou, C. M., Livasy, C., Herschkowitz, J. I., ... Parker, J. S. (2010). Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Research*, *12*(5).
- Pu, H., Zheng, Q., Li, H., Wu, M., An, J., Gui, X., ... Lu, D. (2015). CUDR promotes liver cancer stem cell growth through upregulating TERT and C-Myc. *Oncotarget*, *6*(38).
- Rahman, N., Seal, S., Thompson, D., Kelly, P., Renwick, A., Elliott, A., ... Stratton, M. R. (2007). PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene. *Nature Genetics*, *39*(2), 165–167.
- Reik, W., Brown, K. W., Schneid, H., Bouc, Y. Le, Bickmore, W., & Maher, E. R. (1995). Imprinting mutations in the Beckwith-Wiedemann syndrome suggested by an altered imprinting pattern in the IGF2-H19 domain Wolf. *Human Molecular Genetics*, *4*(12), 2379–2385.
- Rich, J. N. (2016). Cancer stem cells: understanding tumor hierarchy and heterogeneity. *Medicine*.
- Rotter, V. (1983). p53, a transformation-related cellular-encoded protein, can be used as a biochemical marker for the detection of primary mouse tumor cells. *Proceedings of the National Academy of Sciences*, *80*(May), 2613–2617.
- Safari, M. R., Mohammad Rezaei, F., Dehghan, A., Noroozi, R., Taheri, M., & Ghafouri-Fard, S. (2019). Genomic variants within the long non-coding RNA H19 confer risk of breast cancer in Iranian population. *Gene*, *701*(March), 121–124.
- Santer, L., Bär, C., & Thum, T. (2019). Circular RNAs: A Novel Class of Functional RNA Molecules with a

- Therapeutic Perspective. *Molecular Therapy*, 27(8), 1–14.
- Shibue, T., & Robert A. Weinberg. (2017). EMT, CSCs, and drug resistance: the mechanistic link and clinical implications Tsukasa. *Nat Rev Clin Oncol*, 14(10), 611–629.
- Shima, H., Kida, K., Adachi, S., Yamada, A., Sugae, S., Narui, K., ... Endo, I. (2018). Lnc RNA H19 is associated with poor prognosis in breast cancer patients and promotes cancer stemness. *Breast Cancer Research and Treatment*, 170(3), 507–516.
- Shimizu, H., Burch, L. R., Smith, A. J., Dornan, D., Wallace, M., Ball, K. L., & Hupp, T. R. (2002). *The Conformationally Flexible S9 – S10 Linker Region in the Core Domain of p53 Contains a Novel MDM2 Binding Site Whose Mutation Increases Ubiquitination of p53 in Vivo* \*. 277(32), 28446–28458.
- Si, X., Zang, R., Zhang, E., Liu, Y., Shi, X., Zhang, E., ... Sun, Y. (2016). LncRNA H19 confers chemoresistance in ER $\alpha$ -positive breast cancer through epigenetic silencing of the proapoptotic gene BIK. *Oncotarget*, 7(49), 81452–81462.
- Silva, M. J., Costa, P., Dias, A., Valente, M., Louro, H., Boavida, M. G., ... Gene, C. De. (2005). Comparative Analysis of the Mutagenic Activity of Oxaliplatin and Cisplatin in the Hprt Gene of CHO Cells. *Environmental and Molecular Mutagenesis*, 115(May), 104–115.
- Sørli, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J. S., Nobel, A., ... Botstein, D. (2003). *Repeated observation of breast tumor subtypes in independent gene expression data sets*. 100(14), 8418–8423.
- Soussi, T., & Wiman, K. G. (2015). TP53: An oncogene in disguise. *Cell Death and Differentiation*, 22(8), 1239–1249.
- St. Laurent, G., Wahlestedt, C., & Kapranov, P. (2015). The Landscape of long noncoding RNA classification. *Trends in Genetics*, 1–13.
- Sugihara, E., & Saya, H. (2013). Complexity of cancer stem cells. *International Journal of Cancer*, 132(6), 1249–1259.
- Sun, H., Wang, G., Peng, Y., Zeng, Y., Zhu, Q. N., Li, T. L., ... Zhu, Y. S. (2015). H19 lncRNA mediates 17 $\beta$ -estradiol-induced cell proliferation in MCF-7 breast cancer cells. *Oncology Reports*, 33(6), 3045–3052.
- Tang, Q., & Hann, S. S. (2018). HOTAIR: An oncogenic long non-coding RNA in human cancer. *Cellular Physiology and Biochemistry*, 47(3), 893–913.
- Tian, D., Sun, S., & Lee, J. (2008). The long noncoding RNA, Jpx, is a molecular switch for X-chromosome inactivation. *Bone*, 23(1), 1–7.
- Toledo, F., Bluteau, O., & Simeonova, I. (2007). Réactivation de p53 dans les tumeurs. Une stratégie antitumorale prometteuse. *Médecines/Sciences*, (Figure 1), 565–567.
- Torres, L., Ribeiro, F. R., Pandis, N., Andersen, J. A., Heim, S., & Teixeira, M. R. (2007). Intratumor genomic heterogeneity in breast cancer with clonal divergence between primary carcinomas and lymph node metastases. *Breast Cancer Research and Treatment*, 102(2), 143–155.
- Treiber, T., Treiber, N., & Meister, G. (2019). Regulation of microRNA biogenesis and its crosstalk with other cellular pathways. *Nature Reviews Molecular Cell Biology*, 20(January).
- Tsang, W. P., Ng, E. K. O., Ng, S. S. M., Jin, H., Yu, J., Sung, J. J. Y., & Kwok, T. T. (2010). Oncofetal H19-derived miR-675 regulates tumor suppressor RB in human colorectal cancer. *Carcinogenesis*, 31(3), 350–358.
- Ungerleider, N. A., Rao, S. G., Shahbandi, A., Yee, D., Niu, T., Frey, W. D., & Jackson, J. G. (2018). *Breast cancer survival predicted by TP53 mutation status differs markedly depending on treatment*. 1–8.
- Vargo-Gogola, T., & Rosen, J. M. (2007). Modelling breast cancer: One size does not fit all. *Nature Reviews Cancer*, 7(9), 659–672.
- Vasudevan, S. (2012). Posttranscriptional Upregulation by MicroRNAs. *WIREs RNA*, 3(June), 311–330.
- Vasudevan, S., Tong, Y., & Steitz, J. A. (2007). *Switching from Repression to Activation: MicroRNAs Can Up-Regulate Translation Shobha*. 318(December), 1931–1934.
- Vennin, C. (2016). *TUMORIGENESE DES ARN NON CODANTS TRANSCRITS AU LOCUS H19 / IGF2 : H19 ET 91H*.

- Vennin, C., Spruyt, N., Dahmani, F., Julien, S., Bertucci, F., Finetti, P., ... Adriaenssens, E. (2015). The long non coding RNA-derived miR-675 enhances tumorigenesis and metastasis of breast cancer cells by downregulating c-Cbl and Cbl-b. *Oncotarget*, 6(30), 29209–29223.
- Vennin, C., Spruyt, N., Robin, Y. M., Chassat, T., Le Bourhis, X., & Adriaenssens, E. (2017). The long non-coding RNA 91H increases aggressive phenotype of breast cancer cells and up-regulates H19/IGF2 expression through epigenetic modifications. *Cancer Letters*, 385, 198–206.
- Vilmont, V., Filhol, O., Hesse, A. M., Couté, Y., Hue, C., Rémy-Tourneur, L., ... Chiocchia, G. (2015). Modulatory role of the anti-apoptotic protein kinase CK2 in the sub-cellular localization of Fas associated death domain protein (FADD). *Biochimica et Biophysica Acta - Molecular Cell Research*, 1853(11), 2885–2896.
- Visvader, J. E., & Lindeman, G. J. (2008). Cancer stem cells in solid tumours: Accumulating evidence and unresolved questions. *Nature Reviews Cancer*, 8(10), 755–768.
- Walerych, D., Napoli, M., Collavin, L., & Del Sal, G. (2012). The rebel angel: Mutant p53 as the driving oncogene in breast cancer. *Carcinogenesis*, 33(11), 2007–2017.
- Wang, R., Chen, X., Xu, T., Xia, R., Han, L., Chen, W., ... Shu, Y. (2016). MiR-326 regulates cell proliferation and migration in lung cancer by targeting phox2a and is regulated by HOTAIR. *American Journal of Cancer Research*, 6(2), 173–186.
- Wiles, M. V. (1988). Isolation of differentially expressed human cDNA clones: similarities between mouse and human embryonal carcinoma cell differentiation. *Development*, 104(3), 403–413.
- Wu, C., & Alman, B. A. (2008). Side population cells in human cancers. *Cancer Letters*, 268(1), 1–9.
- Wu, G., Cai, J., Han, Y., Chen, J., Huang, Z., Chen, C., ... Huang, H. (2014). *LincRNA-p21 Regulates Neointima Formation, Vascular Smooth Muscle Cell Proliferation, Apoptosis and Atherosclerosis by Enhancing p53 Activity*.
- Wu, W., Hu, Q., Nie, E., Yu, T., Wu, Y., Zhi, T., ... You, Y. (2017). Hypoxia induces H19 expression through direct and indirect Hif-1 $\alpha$  activity, promoting oncogenic effects in glioblastoma. *Scientific Reports*, 7(October 2016), 1–13.
- Wu, Z. R., Yan, L., Liu, Y. T., Cao, L., Guo, Y. H., Zhang, Y., ... Wu, Z. B. (2018). Inhibition of mTORC1 by lncRNA H19 via disrupting 4E-BP1/Raptor interaction in pituitary tumours. *Nature Communications*, 9(1).
- Xie, M., Li, M., Vilborg, A., Lee, N., Shu, M. Di, Yartseva, V., ... Steitz, J. A. (2013). Mammalian 5'-capped microRNA precursors that generate a single microRNA. *Cell*, 155(7), 1568–1580.
- Xing, Y. H., & Chen, L. L. (2018). Processing and roles of snoRNA-ended long noncoding RNAs. *Critical Reviews in Biochemistry and Molecular Biology*, 53(6), 596–606.
- Yan, J., Zhang, Y., She, Q., Li, X., Peng, L., Wang, X., ... Zhang, G. (2017). Long Noncoding RNA H19/miR-675 Axis Promotes Gastric Cancer via FADD/Caspase 8/Caspase 3 Signaling Pathway. *Cellular Physiology and Biochemistry*, 42(6), 2364–2376.
- Yang, F., Bi, J., Xue, X., Zheng, L., Zhi, K., Hua, J., & Fang, G. (2012). *Up-regulated long non-coding RNA H19 contributes to proliferation of gastric cancer cells*. 1–7.
- Yoshimizu, T., Miroglio, A., Ripoche, M.-A., Gabory, A., Vernucci, M., Riccio, A., ... Dandolo, L. (2008). The H19 locus acts in vivo as a tumor suppressor. *Proceedings of the National Academy of Sciences*, 105(34), 12417–12422.
- Youness, R. A., & Gad, M. Z. (2019). Long non-coding RNAs: Functional regulatory players in breast cancer. *Non-Coding RNA Research*, 4(1), 36–44.
- Yu, H., Wang, S., Zhu, H., & Rao, D. (2019). LncRNA MT1JP functions as a tumor suppressor via regulating miR-214-3p expression in bladder cancer. *Journal of Cellular Physiology*, 234(9), 16160–16167.
- Zhai, L. L., Wang, P., Zhou, L. Y., Yin, J. Y., Tang, Q., Zhang, T. J., ... Deng, Z. Q. (2015). Over-expression of miR-675 in formalin-fixed paraffin-embedded (FFPE) tissues of breast cancer patients. *International Journal of Clinical and Experimental Medicine*, 8(7), 11195–11201.
- Zhang, A., Zhou, N., Huang, J., Liu, Q., Fukuda, K., Ma, D., & Lu, Z. (2012). The human long non-coding RNA-RoR is a p53 repressor in response to DNA damage. *Cell Research*, 23(3), 340–350.
- Zhang, J., Zhang, D., & Hua, Z. (2004). FADD and its phosphorylation. *IUBMB Life*, 56(7), 395–401.

- Zhang, Y., Yang, L., & Chen, L. L. (2014). Life without A tail: New formats of long noncoding RNAs. *International Journal of Biochemistry and Cell Biology*, *54*, 338–349.
- Zhou, S., Schuetz, J. D., Bunting, K. D., Colapietro, A. M., Sampath, J., Morris, J. J., ... Sorrentino, B. P. (2001). The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. *Nature Medicine*, *7*(9), 1028–1034.
- Zhou, W., Ye, X. L., Xu, J., Cao, M. G., Fang, Z. Y., Li, L. Y., ... Xie, D. (2017). The lncRNA H19 mediates breast cancer cell plasticity during EMT and MET plasticity by differentially sponging miR-200b/c and let-7b. *Science Signaling*, *10*(483).
- Zhou, X., Yin, C., Dang, Y., Ye, F., & Zhang, G. (2015). Identification of the long non-coding RNA H19 in plasma as a novel biomarker for diagnosis of gastric cancer. *Scientific Reports*, *5*(March), 1–10.



# Annexe



# Regulation of Human Breast Cancer by the Long Non-Coding RNA *H19*

Jordan Collette <sup>1,2</sup>, Xuefen Le Bourhis <sup>1,2</sup> and Eric Adriaenssens <sup>1,2,\*</sup>

<sup>1</sup> INSERM U908, 59655 Villeneuve d'Ascq, France; jorkan62@gmail.com (J.C.); xuefen.lebourhis@univ-lille1.fr (X.L.B.)

<sup>2</sup> University of Lille, 59655 Villeneuve d'Ascq, France

\* Correspondence: eric.adriaenssens@univ-lille1.fr; Tel.: +333-204-340-38

Received: 3 October 2017; Accepted: 30 October 2017; Published: 3 November 2017

**Abstract:** Breast cancer is one of the most common causes of cancer related deaths in women. Despite the progress in early detection and use of new therapeutic targets associated with development of novel therapeutic options, breast cancer remains a major problem in public health. Indeed, even if the survival rate has improved for breast cancer patients, the number of recurrences within five years and the five-year relative survival rate in patients with metastasis remain dramatic. Thus, the discovery of new molecular actors involved in breast progression is essential to improve the management of this disease. Numerous data indicate that long non-coding RNA are implicated in breast cancer development. The oncofetal lncRNA *H19* was the first RNA identified as a riboregulator. Studying of this lncRNA revealed its implication in both normal development and diseases. In this review, we summarize the different mechanisms of action of *H19* in human breast cancer.

**Keywords:** lncRNA; *H19*; breast cancer; non-coding RNA

---

## 1. Introduction

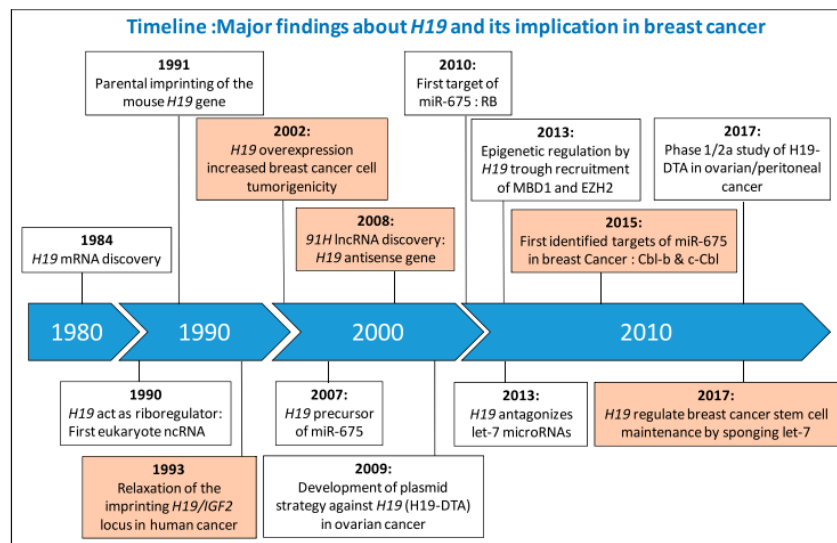
Breast cancer is the most common tumor in women and caused 508,000 deaths worldwide in 2011 [1]. Recent advances in molecular classification of this pathology allowed personalized-treatment of patients and better outcomes, like the use of Herceptin in patients with overexpression of human epidermal growth factor receptor 2 (Her-2) [2]. However, some classes of breast cancer such as triple-negative, which is characterized by neither expression of progesterone receptor (PR), estrogen receptor, nor Her-2, remains a poor prognostic for patients. The discovery of new molecular actors involved in the regulation of breast cancer development is essential to improve the management of this disease. During the last decades, plenty of non-coding RNAs have been involved in breast cancer development [3]. The study of non-coding RNAs could lead to the development of new therapeutic strategies and better outcomes for patients with triple negative breast cancer, and more generally to patients with cancer.

The Encyclopedia of DNA Elements (ENCODE) consortium revealed that up to 80% of the human genome is transcribed into functional RNAs, but only 2% of the genome codes for proteins [4–6]. RNAs without coding potential are referred to as non-coding RNAs (ncRNAs). Based on their lengths, they can be classified into two classes: small ncRNAs (<200 nt) and long ncRNAs (>200 nt). Small ncRNAs include microRNAs (miRs), small interfering RNAs (siRNAs), PIWI-interacting RNAs (piRNAs) or small nucleolar RNA (snoRNAs). miRs, siRNAs, and piRNAs were shown to mainly act as negative regulators of gene expression, whereas snoRNA serves as a guide to induce chemical modification of other RNAs [7]. Recently, Hon et al. identified 19,175 potentially functional lncRNAs in the human genome [8]. The majority of lncRNAs shared similarities with mRNA: they are transcribed by RNA polymerase II; 5' capped; 3' polyadenylated; and could be subjected to splicing [9,10]. The lncRNAs are also transcriptionally regulated by transcription factor and epigenetic modifications and their expression is cell-type/tissue specific [8].

Implication of lncRNA is well documented in different pathologies, including neurological diseases [11,12], diabetes [13], and cancers [14]. lncRNAs exert their functions through diverse molecular mechanisms such as acting as scaffolds for RNP complex, decoys for transcriptional factors or microRNAs, RNA interference, targeting of transcriptional factor or chromatin modifier protein to specific genomic loci, transcriptional regulation in *cis* or *trans* [15]. In this review, we will focus on the implication of *H19*, the first identified lncRNA, in human breast cancer.

## 2. *H19* Gene Locus

The *H19/IGF2* locus located at position 11p15.5 is subject to genomic imprinting and is encoded for several transcripts. One of these transcripts, *H19*, was proposed to act as a riboregulator by Brannan et al. in 1990 and was the first identified lncRNA [16]. Major findings about *H19* and its implication in breast cancer are summarized below in a timeline manner (Figure 1).



**Figure 1.** Major finding about *H19* and its implication in breast cancer. Information concerning the implication of *H19* in breast cancer are colored orange. IGF2, Insulin-like Growth Factor 2; lncRNA, Long Non-Coding RNA; RB, Retinoblastoma; MBD1, Methyl-CpG Binding Domain, EZH2, Enhancer of Zeste Homolog 2; Cbl, Casitas B-lineage Lymphoma; *H19*-DTA, Plasmid encoding the A chain of diphtheria toxin (DT-A) driven by the regulatory sequences of human *H19*.

The *H19* gene is transcribed by the RNA polymerase II to give raise a polyadenylated, capped and spliced 2.3 kb RNA. *H19* is paternally imprinted and maternally expressed [17]. It is expressed during embryonic development and repressed after birth expect in a few tissues like mammary gland and uterus [18]. Aberrant expression of *H19* due to imprinting modification is responsible for developmental diseases. The loss of *H19* expression and a biallelic expression of IGF2 are responsible for the Beckwith–Wiedmann syndrome characterized by fetal and postnatal overgrowth and predisposition to pediatric cancers such as Wilm’s tumors [19]. Biallelic *H19* expression and loss of IGF2 expression can lead to Silver–Russel syndrome characterized by intrauterine and postnatal growth retardation combined with others symptoms [19]. In cancer, *H19* is frequently overexpressed and it is associated to many aspect of cancer development [20,21].

In 2007, Cai & Kullen demonstrated that *H19* is a precursor of miR-675 [22]. The two strands of miR-675, miR-675-5p, and miR-675-3p, have been involved in disease development and notably in cancer development (Section 3.1). In the locus *H19/IGF2*, we have identified the presence of a new lncRNA antisense to *H19* gene, named *91H* [23]. This lncRNA is also associated with cancer development in solid-tumors such as breast cancer or osteosarcoma [23,24]. The implication of *H19* and *91H* in cancer is well described and is mediated by different mechanisms characteristics of that observed for other non-coding RNAs. In this review, we will focus on the role of *H19* in human breast cancer.

### 3. *H19* Gene Locus in Human Breast Cancer

The implication of *H19* in tumorigenesis has been reported and *H19* is overexpressed in many solid tumors such as prostate, bladder or breast cancers [25–27]. We showed that *H19* is overexpressed in 73% of breast cancer tissues when compared to healthy tissues [28]. Several studies showed that *H19* is controlled by steroid hormones in normal and cancerous mammary gland, uterus and prostate [29–31]. In breast cancer, the expression of *H19* is higher in Estrogen receptor (ER $\alpha$ ) positive cells, but in the ER $\alpha$  negative MDA-MB-231 cell line, ectopic overexpression of *H19* is associated with increased proliferation [27,30]. Collectively, these data indicate that *H19* favor breast cancer development probably through different mechanisms discussed below.

#### 3.1. *H19*: Precursor of miR-675-5p and miR-675-3p

The *H19*-derived miR-675 gives rise to two functional microRNA, miR-675-5p, and miR-675-3p with different validated targets. The implication of miR-675 in cancer was firstly shown by Tsang WP et al. in colorectal cancer (CRC). By targeting the tumor suppressor retinoblastoma (RB) protein, miR-675-5p regulates the CRC development [32]. In human breast cancer, we have identified two ubiquitin ligase E3, c-Cbl, and Cbl-b, as direct targets of miR-675-5p [33]. The overexpression of miR-675-5p in breast cancer cells lines induced the downregulation of c-Cbl and Cbl-b proteins and increased the stability and the activation of Epidermal growth factor receptor (EGFR) and c-Met. Steady activation of Akt and Erk pathway enhanced the proliferation of human breast cancer and their metastasis abilities in xenograft experiments.

Zhai et al. investigated the expression of miR-675-5p in formalin-fixed paraffin-embedded (FFPE) tissues of 100 breast cancer patients [34]. The authors showed that miR-675 is significantly up-regulated in breast cancer patients compared with controls, but this up-regulation is not correlated with clinical and pathological status including ER and PR expression, age, and lymph node stage. The frequency of miR-675 overexpression was higher in the patients with low histological grade (I and II). Cordero et al. analyzed DNA methylation levels of 517 microRNA encoding genes in prediagnostic peripheral white blood cells of subjects who have developed colorectal cancer or breast cancer (BC) and subjects who remained clinically healthy [35]. They found that eight microRNAs, including miR-675-5p, were differentially methylated in subjects who went on to develop breast cancer. In those subjects, miR-675-5p was significantly hypomethylated suggesting that miR-675-5p could be used as biomarker for breast cancer.

All the known targets of miR-675-5p and miR-675-3p implicated or not in neoplasia are resumed in Table 1.

**Table 1.** Validated targets of miR-675-5p and miR-675-3p.

Targets of miR-675-5p	Cellular Context	Proteins Function	References
c-Cbl & Cbl-b	Breast cancer	Ubiquitin ligase E3	[33]
HDAC 4/5/6	Bone Marrow Mesenchymal Stem Cells	Histone deacetylase	[36]
CaMKII $\delta$	Cardiomyocyte hypertrophy	Serine threonine protein kinase	[37]
USP10	c-kit+ cardiac progenitor cells	Ubiquitin-specific protease	[38]
RB	Colorectal cancer	Cell cycle regulator	[32]
DDB2	Colon cancer cells	Transcriptional repressor	[39]
VDR	Colon cancer cells	Vitamin D receptor	[40]
VDAC1	Diabetic cardiomyopathy	Required for mitochondria-mediated apoptosis	[41]
REPS2 cell carcinoma	Esophageal squamous	Repressor of cell proliferation and migration	[42]

**Table 1. Cont.**

<b>Targets of miR-675-5p</b>	<b>Cellular Context</b>	<b>Proteins Function</b>	<b>References</b>
CALN1	Gastric cancer	Calcium-binding protein	[43]
RUNX1	Gastric cancer	Transcription factor	[44,45]
FADD	Gastric cancer	Apoptotic adaptor that recruits caspase 8 or 10	[46]
Cadherin 13	Glioma cell	Atypical cadherin lacking the cytoplasmic domain	[47]
RB & TWIST1	Hepatocellular carcinomas	Twist1: Transcription factor	[48]
GPR55	Non-small cell lung cancer	G protein-coupled receptor	[49]
TGF- $\beta$ 1	Osteoblast differentiation	Growth factor	[50]
TGF- $\beta$ 1	Prostate cancer	Growth factor	[25]
NOMO1	Placental trophoblast cell	Nodal signaling pathway	[51]
ATP8A2	Skeletal cell flippase complex	Catalytic component of a P4-ATPase	[52]
CDC6	Skeletal muscle	Essential for the initiation of DNA replication	[53]
VDR	Ulcerative Colitis	Vitamin D receptor	[54]
<b>Targets of miR-675-3p</b>	<b>Cellular Context</b>	<b>Proteins Function</b>	<b>References</b>
Cadherin 11	Melanogenesis	Type II classical cadherin	[55]
MITF	Melanogenesis	Transcription factor	[56]
IGF1R	Placenta	Insulin-like growth factor 1 receptor	[57]
TGF- $\beta$ 1	Osteoblast differentiation	Growth factor	[50]
SMAD1 & SMAD5	Skeletal muscle transcriptional modulator	Intracellular signal transducer and	[53]

Some of these targets could explain the oncogenic role of *H19* in breast cancer. For example, miR-675 stimulates migration and invasion by targeting TGF- $\beta$ 1 in prostate cancer cells, Cadherin13 in glioma cells, or RUNX1 in gastric cancer cells [42,44,47]. We found that *H19* and miR-675 expression enhances breast cancer cell migration [33,58]. This could be mediated by targeting the above cited molecules, even if miRNAs targets remain tissue specific. By example, miR-675 was shown to downregulate the expression of RB in human colorectal cancer to promote tumor development [32]. RB was also demonstrated as a target of miR-675-5p in hepatocellular carcinomas [48]. However, we showed that miR-675-5p doesn't interact with RB mRNA in human breast cancer cell lines [33].

### 3.2. Competing Endogenous RNAs (ceRNAs): Sequestration of miRs by *H19*

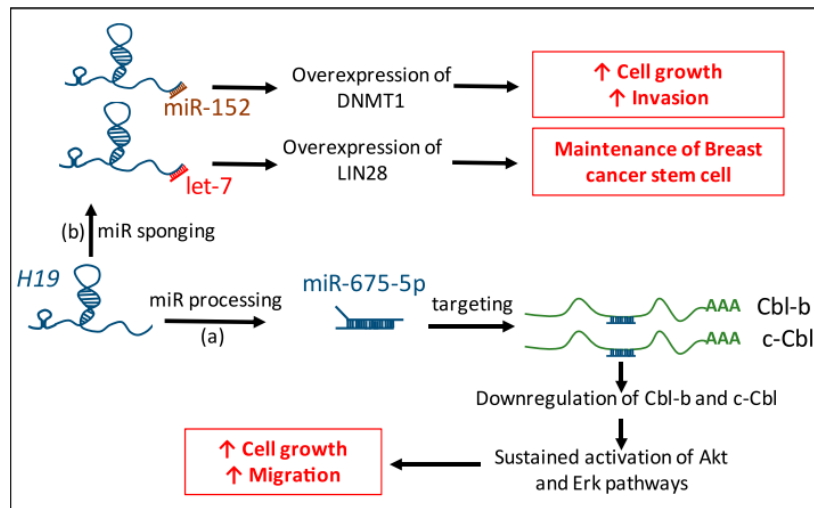
Tay et al. reported that lncRNA can be in competition with mRNA for common microRNAs and termed such lncRNA transcript as competing endogenous RNAs (ceRNAs) [59]. *H19*, as numerous lncRNAs, could act by this mechanism. Recently, Peng et al. demonstrated the implication of *H19* in maintenance of breast cancer stem cells through the sequestration of let-7 [60]. The lower availability of let-7 increases the expression of its target, the core pluripotency factor LIN28; LIN28 in turn blocks mature let-7 production and enhances the expression of *H19* in breast cancer stem cells. In human breast cancer cells lines, *H19* upregulates the DNA methyltransferase DNMT1 by sponging miR-152, leading to enhancement of cell proliferation and invasion of the cells [61]. The authors also revealed a correlation between the overexpression of *H19* and DNMT1 and the downregulation of miR-152 in human breast tumor tissues. In 2017, Zhou et al. demonstrated that *H19* regulates Epithelial-mesenchymal transition (EMT) and Mesenchymal-epithelial transition (MET) by differentially acting as a sponge for the microRNA miR-200b/b and let-7b using a mouse model of spontaneous metastatic breast cancer [62]. Other microRNAs sequestered by *H19* are indicated in Table 2

**Table 2.** Validated miRNAs sponged by *H19*.

miRNAs Sponged by <i>H19</i>	Cellular Context	References
hsa-miR-19a/b	Acute myelocytic leukemia	[63]
miR-29b-3p	Bladder cancer	[64]
miR-152	Breast cancer	[61]
let-7	Breast cancer stem cells	[60]
miR-455	Cardiac fibrosis	[65]
let7	Endometriosis	[66]
miR-181-d	Glioblastoma	[67]
let-7	HEK293	[68]
miR-106-a & miR-17-5p	Hela Cells, myoblast	[69]
let-7	Muscle cells	[70]
let-7b & miR-200b/c	Mouse breast cancer	[62]
miR 141 miR 22	Osteoblast	[71]
mir-200s	Osteosarcoma	[72]
miR-17-5p	Thyroid cancer	[73]

The impact of *H19* on metastasis abilities of the human breast cancer cell could be due to the sponging of these microRNAs. By example, *H19* regulates EMT in bladder cancer by sponging miR-29b-3p [64]. In thyroid cancer, *H19* promotes proliferation, migration, and invasion of cancer cells through the sponging of miR-17-5p [73]. *H19* also regulates tumorigenic abilities of glioblastoma cells by sponging miR-181-d, in acute myelocytic leukemia by sponging has-miR-19a/b and in osteosarcoma by serving as competing endogenous RNA for the miR-200s family [63,67,72].

In conclusion, the lncRNA *H19* interacts with miRs pathways not only by being the precursor of miR-675, but also by physically interacting with other miRs to regulate the expression of their targets (Figure 2).



**Figure 2.** Connection between *H19* and microRNAs. *H19* is the precursor of miR-675-5p which targets Cbl-b and c-Cbl mRNA in breast cancer. Downregulation of Cbl-b and c-Cbl protein expression induces sustained activation of Akt and Erk pathways that lead to increased cell growth and migration potential (a). *H19* physically interacts with miR-152 and let-7 and impairs their bioavailability to induce the overexpression of their targets, DNMT1 and LIN28, and participate in tumorigenic properties and maintenance of stemness in breast cancer cells (b). Red arrows indicate an increased phenotype.

### 3.3. Epigenetics Modification Induced by *H19*

Long non-coding RNA could interact with chromatin modifier protein and contribute to epigenetic regulation of gene expression. *H19* was shown to interact with the histone methyl transferase Enhancer

of zeste homolog 2 (EZH2) and epigenetically silenced *E-cadherin* in bladder cancer and *DIRAS3* in diabetic cardiomyopathy [74,75]. Si et al. identified *H19* as a factor associated with paclitaxel resistance in ER $\alpha$ -positive breast cancer cells. *H19* decreases cell apoptosis induced by paclitaxel treatment by inhibiting the transcription of pro-apoptotic genes *BIK* and *NOXA*. The recruitment of EZH2 by *H19* and its targeting onto the promoter of *BIK* induce its downregulation. As described in Section 3.2, *H19* impairs availability of miR-152 and increases the expression of the epigenetic regulator DNMT1 and so increases proliferation of breast cancer cells lines [61]. However, the epigenetic modification induced by DNMT1 remains unknown. In the embryo, *H19* has been shown to physically interact with MBD1 and induce its recruitment at several imprinted genes including *IGF2*, *PEG1*, and *SLC38A4* [76]. In 2015, Zhou et al. demonstrated that *H19* binds to and inhibits *S*-adenosylhomocysteine hydrolase (SAHH) which carries the synthesis of *S*-adenosylmethionine (SAM), the only source of methyl for methyltransferases and other processes that are methyl-dependent [77]. Modulation of *H19* can modify genome wide DNA methylation and knockdown of *H19* increased DNMT3B-mediated methylation of the lncRNA-encoding gene *Nctc1* within the *Igf2-H19-Nctc1* locus. Taken together, these data revealed that *H19* is an important regulator of epigenetic status of target genes. The epigenetic regulation driven by *H19* could be done by physical interaction with chromatin modifier protein or indirectly by regulating their expression.

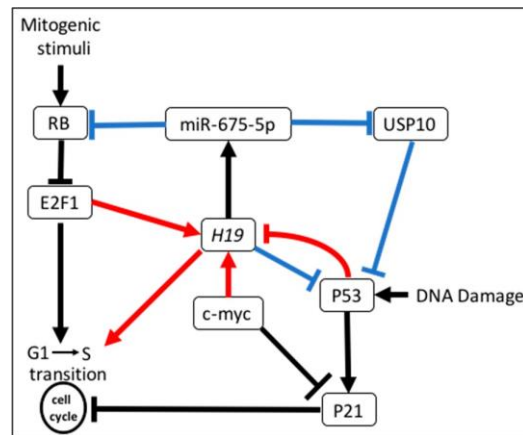
#### 3.4. *91H: H19 Antisense Transcript*

In 2008, we identified a new non-coding transcript within the *IGF2/H19* locus which is an antisense gene to *H19* named *91H* [23]. This transcript of 120 kb in humans is an lncRNA expressed in human and mice from the maternal allele. We also showed that *91H* is overexpressed in breast tumors. Further studies demonstrate the implication of *91H* in osteosarcoma [24] and colorectal cancer [78]. Recently, we showed that in breast cancer cells that *91H* exerts oncogenic properties by promoting cell growth, migration, and invasion [79]. In *91H*-knockdown cell lines the expression of *H19* and *IGF2* is reduced through epigenetic modifications on *H19/IGF2* locus. These data suggest that *91H* plays an essential role in breast cancer development, but it is possible that *91H* regulation of tumorigenicity requires other factors than *H19* or *IGF2*. Differentiate the implication of *91H* in biological process independently of *H19* is not an easy task, but well thought experiments are needed to address the issue.

#### 3.5. *Regulation of Cell Cycle*

Besides the involvement of *H19* in EMT, migration, metastasis, and carcinogenesis through the mechanisms described above, *H19* plays a key role in the regulation of the cell cycle. The overexpression of *H19* in breast cancer cells lines facilitates cell cycle transition G1/S while downregulation of *H19* by RNA interference impedes S-phase entry and proliferation [80]. *H19* is activated by E2F1 binding (a key factor in the G1/S transition) to two consensus sites on *H19* promoter and negatively regulated by E2F6 and RB protein. Interestingly, in human colorectal cancer cell lines, *H19* through microRNA downregulates RB protein and increases cell growth [32]. Barsyte-Lovejoy et al. demonstrated that the oncogene c-Myc binds specifically to *H19* maternal allele to promote its transcription, leading to proliferation of breast and lung cancerous cells [81]. We have also shown that the tumor suppressor protein and cell cycle regulator p53 negatively regulates *H19* in tumor cells [82]. The interaction between *H19* and p53 was also described in other cancers. In gastric cancer cells *H19* physically interacts with p53 to induce p53 inactivation [83]. The *H19*-derived miR-675 negatively regulates p53 through an unknown target in bladder cancer cell [26]. In C-kit<sup>+</sup> cardiac progenitor cells, Cai et al. showed that miR-675 negatively regulates p53 through the targeting of USP10 [38]. Even if the regulation of p53 by *H19* in human breast cancer is not yet described, these data, collectively, demonstrate that *H19* and miR-675 play a pivotal role in the regulation of cell cycle in cancer as illustrated in Figure 3.

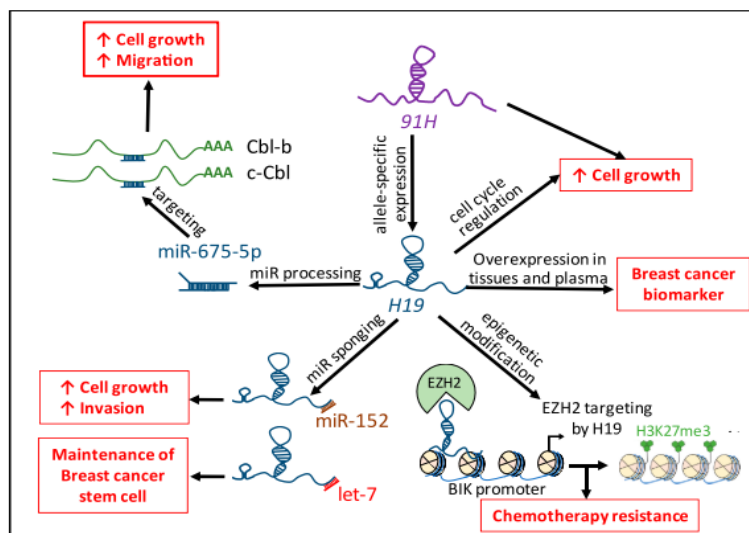




**Figure 3.** *H19* and cell cycle regulation. *H19* favors the cell cycle progression through different mechanisms. Mechanisms demonstrated in breast cancer are indicated in red, in other cellular context are indicated in blue. General mechanisms of cell cycle regulation are indicated in black. Arrows indicate positive regulations whereas lines with bars correspond to inhibition.

#### 4. Conclusions

*H19* is involved in human breast cancer through interaction with protein, microRNAs, *H19*-derived miR-675-5p, and *H19* antisense lncRNA (*91H*) but the study of these ncRNAs is not only restrained in cancer cell behavior. Increasing studies have reported their interests as cancer biomarkers and therapeutic targets. In breast cancer, Zhang et al. demonstrated that the expression of *H19* is significantly increased in cancer biopsies and plasma compared with healthy controls, plasma *H19* levels were significantly correlated with progesterone and estrogen receptors and lymph node metastasis [84]. The plasma level of *H19* is higher in patients with gastric cancer compared to normal controls. Lower *H19* expression is found in postoperative samples compared to preoperative ones [85]. Higher levels of miR-675 are also found in both tumor samples and gastric juice of patients suffering from gastric cancer [86]. The different implication of *H19* in human breast cancer is illustrated in Figure 4.



**Figure 4.** Implication of *H19* in human breast cancer. The *H19* lncRNA favors breast tumorigenicity by regulating the cell cycle, through the processing of miR-675-5p and the sponging of miR-152 and let-7, and regulating chemotherapy resistance through epigenetic modification. The *H19* antisense, *91H*, allows allele-specific expression of *H19* and participates in breast cancer cell biology. *H19* and miR-675 are overexpressed in breast cancer tissues and plasma of patients and could be used as tumor biomarkers. H3K27me3: Trimethylation of lysine 27 on histone H3. Red arrows indicate an increased phenotype.

From a therapeutic perspective, a plasmid-based strategy (DTA-H19/BC-819) to target *H19* is presently in a phase 2b clinical trial for bladder cancer and in phase 1/2a for ovarian and peritoneal cancer [87,88]. Although further studies are needed, the targeting of *H19* and miR-675 could provide novel opportunities in the treatment of cancer patients.

**Acknowledgments:** This work was supported by INSERM, INCA (PLBio 2010-180), Ligue contre le cancer and Cancéropole Nord-Ouest.

**Author Contributions:** Eric Adriaenssens provided direction during the preparation of this manuscript; Jordan Collette conducted the literature review and drafted the manuscript; Xuefen Le Bourhis reviewed the manuscript and made significant revisions on the drafts.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

RNA	Ribonucleic Acid
DNA	Deoxyribonucleic Acid
ENCODE	Encyclopedia of DNA
ncRNAs	Non-coding RNAs
lncRNAs	Long non-coding RNAs
miRs	MicroRNAs
siRNAs	Small non-coding RNAs
piRNAs	PIWI-interacting RNAs
snoRNAs	Small nucleolar RNA
Her-2	Human epidermal growth factor receptor 2
PR	Progesterone receptor
ER	Estrogen receptor
IGF2	Insulin growth factor 2
BC	Breast cancer
CRC	Colorectal cancer
RB	Retinoblastoma
DNMT1	DNA Methyltransferase 1
EMT	Epithelial-mesenchymal transition
MET	Mesenchymal-epithelial transition
BIK	BCL2 Interacting Killer
NOXA	NADPH Oxidase Activator 1
MBD1	Methyl-CpG Binding Domain Protein 1
PTX	Paclitaxel
PEG1	Paternally-Expressed Gene 1 Protein
SAHH	S-adenosylhomocysteine hydrolase
E2F1	E2F Transcription Factor 1
CPCs	c-kit+ cardiac progenitor cells
USP10	Ubiquitin Specific Peptidase 10

## References

1. World Health Organization. *Global Health Estimates*; World Health Organization: Geneva, Switzerland, 2013.
2. Dean-Colomb, W.; Esteva, F.J. Her2-positive Breast Cancer: Herceptin and Beyond. *Eur. J. Cancer* **2008**, *44*, 2806–2812. [[CrossRef](#)] [[PubMed](#)]
3. Cerk, S.; Schwarzenbacher, D.; Adiprasito, J.B.; Stotz, M.; Hutterer, G.C.; Gerger, A.; Ling, H.; Calin, G.A.; Pichler, M. Current status of long non-coding RNAs in human breast cancer. *Int. J. Mol. Sci.* **2016**, *17*, 1485. [[CrossRef](#)] [[PubMed](#)]
4. Birney, E.; Stamatoyannopoulos, J.A.; Dutta, A.; Guigó, R.; Gingeras, T.R.; Margulies, E.H.; Weng, Z.; Snyder, M.; Dermitzakis, T.; Thurman, R.E.; et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **2007**, *447*, 799–816. [[CrossRef](#)] [[PubMed](#)]

5. Kapranov, P.; Cheng, J.; Dike, S.; Nix, D.A.; Duttagupta, R.; Willingham, A.T.; Stadler, P.F.; Hertel, J.; Hackermüller, J.; Hofacker, I.L.; et al. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **2007**, *316*, 1484–1488. [[CrossRef](#)] [[PubMed](#)]
6. Qu, H.; Fang, X. A Brief review on the human encyclopedia of DNA elements (ENCODE) project. *Genom. Proteom. Bioinform.* **2013**, *11*, 135–141. [[CrossRef](#)] [[PubMed](#)]
7. Scott, M.S.; Motoharu, O. From snoRNA to miRNA: Dual function regulatory non-coding RNAs. *Biochimie* **2011**, *93*, 1987–1992. [[CrossRef](#)] [[PubMed](#)]
8. Hon, C.C.; Ramiloski, J.A.; Harshbarger, J.; Bertin, N.; Rackham, O.J.L.; Gough, J.; Denisenko, E.; Schmeier, S.; Poulsen, T.M.; Severin, J.; et al. An Atlas of human long non-coding RNAs with accurate 5' ends. *Nature* **2017**, *543*, 199–204. [[CrossRef](#)] [[PubMed](#)]
9. Wu, H.; Yang, L.; Chen, L.L. The diversity of long non-coding RNAs and their generation. *Trends Genet.* **2017**, *33*, 540–552. [[CrossRef](#)] [[PubMed](#)]
10. Quinn, J.J.; Chang, H.Y. Unique features of long non-coding RNA biogenesis and function. *Nat. Rev. Genet.* **2015**, *17*, 47–62. [[CrossRef](#)] [[PubMed](#)]
11. Luo, Q.; Chen, Y. Long noncoding RNAs and Alzheimer's disease. *Clin. Interv. Aging* **2016**, *11*, 867–872. [[CrossRef](#)] [[PubMed](#)]
12. Huang, X.; Luo, Y.L.; Mao, Y.S.; Ji, J.J. The link between long noncoding RNAs and depression. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **2017**, *73*, 73–78. [[CrossRef](#)] [[PubMed](#)]
13. Mirza, A.H.; Kaur, S.; Pociot, F. Long non-coding RNAs as novel players in  $\beta$  cell function and type 1 diabetes. *Hum. Genom.* **2017**, *11*, 17. [[CrossRef](#)] [[PubMed](#)]
14. Bhan, A.; Soleimani, M.; Mandal, S.S. Long Noncoding RNA and cancer: A new paradigm. *Cancer Res.* **2017**, *77*, 3965–3981. [[CrossRef](#)] [[PubMed](#)]
15. Angrand, P.O.; Vennin, C.; Le Bourhis, X.; Adriaenssens, E. The role of long non-coding RNAs in genome formatting and expression. *Front. Genet.* **2015**, *6*. [[CrossRef](#)] [[PubMed](#)]
16. Brannan, C.I.; Dees, E.C.; Ingram, R.S.; Tilghman, S.M. The product of the *H19* gene may function as an RNA. *Mol. Cell. Biol.* **1990**, *10*, 28–36. [[CrossRef](#)] [[PubMed](#)]
17. Bartolomei, M.S.; Zemel, S.; Tilghman, S.M. Parental imprinting of the mouse *H19* gene. *Nature* **1991**, *351*, 153–155. [[CrossRef](#)] [[PubMed](#)]
18. Vennin, C.; Dahmani, F.; Spruyt, N.; Adriaenssens, E. Role of long non-coding RNA in cells: Example of the *H19/IGF2* locus. *Adv. Biosci. Biotechnol.* **2013**, *4*, 34–44. [[CrossRef](#)]
19. Delaval, K.; Wagschal, A.; Feil, R. Epigenetic deregulation of imprinting in congenital diseases of aberrant growth. *BioEssays* **2006**, *28*, 453–459. [[CrossRef](#)] [[PubMed](#)]
20. Rainier, S.; Johnson, L.A.; Dobry, C.J.; Ping, A.J.; Grundy, P.E.; Feinberg, A.P. Relaxation of imprinted genes in human cancer. *Nature* **1993**, *362*, 747–749. [[CrossRef](#)] [[PubMed](#)]
21. Raveh, E.; Matouk, I.J.; Gilon, M.; Hochberg, A. The *H19* long non-coding RNA in cancer initiation, progression and metastasis – A proposed unifying theory. *Mol. Cancer* **2015**, *14*. [[CrossRef](#)] [[PubMed](#)]
22. Cai, X.; Cullen, B.R. The Imprinted *H19* noncoding RNA is a primary microRNA precursor. *RNA* **2007**, *13*, 313–316. [[CrossRef](#)] [[PubMed](#)]
23. Berteaux, N.; Aptel, N.; Cathala, G.; Genton, C.; Coll, J.; Daccache, A.; Spruyt, N.; Hondermarck, H.; Dugimont, T.; Curgy, J.J.; et al. A novel *H19* antisense RNA overexpressed in breast cancer contributes to paternal *IGF2* expression. *Mol. Cell. Biol.* **2008**, *28*, 6731–6745. [[CrossRef](#)] [[PubMed](#)]
24. Xia, W.K.; Lin, Q.F.; Shen, D.; Liu, Z.L.; Su, J.; Mao, W.D. Clinical implication of long noncoding RNA 91H expression profile in osteosarcoma patients. *OncoTargets Ther.* **2016**, *9*, 4645–4652. [[CrossRef](#)]
25. Zhu, M.; Chen, Q.; Liu, X.; Sun, Q.; Zhao, X.; Deng, R.; Wang, Y.; Huang, J.; Xu, M.; Yan, J.; et al. lncRNA *H19*/miR-675 axis represses prostate cancer metastasis by targeting *TGFBI*. *FEBS J.* **2014**, *281*, 3766–3775. [[CrossRef](#)] [[PubMed](#)]
26. Liu, C.; Chen, Z.; Fang, J.; Xu, A.; Zhang, W.; Wang, Z. *H19*-derived miR-675 contributes to bladder cancer cell proliferation by regulating *P53* activation. *Tumor Biol.* **2016**, *37*, 263–270. [[CrossRef](#)] [[PubMed](#)]
27. Lottin, S.; Adriaenssens, E.; Dupressoir, T.; Berteaux, N.; Montpellier, C.; Coll, J.; Dugimont, T.; Curgy, J.J. Overexpression of an ectopic *H19* gene enhances the tumorigenic properties of breast cancer cells. *Carcinogenesis* **2002**, *23*, 1885–1895. [[CrossRef](#)] [[PubMed](#)]

28. Adriaenssens, E.; Dumont, L.; Lottin, S.; Bolle, D.; Leprêtre, A.; Delobelle, A.; Bouali, F.; Dugimont, T.; Coll, J.; Curgy, J.J. 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. [[CrossRef](#)]
29. Berteaux, N.; Lottin, S.; Adriaenssens, E.; Coppenolle, F.W.; Leroy, X.; Coll, J.; Dugimont, T.; Curgy, J.J. Hormonal regulation of H19 gene expression in prostate epithelial cells. *J. Endocrinol.* **2004**, *183*, 69–78. [[CrossRef](#)] [[PubMed](#)]
30. Basak, P.; Chatterjee, S.; Weger, S.; Bruce, M.C.; Murphy, L.C.; Raouf, A. Estrogen regulates luminal progenitor cell differentiation through H19 gene expression. *Endocr. Relat. Cancer* **2015**, *22*, 505–517. [[CrossRef](#)] [[PubMed](#)]
31. Sun, H.; Wang, G.; Peng, Y.; Zeng, Y.; Zhu, Q.N.; Li, T.L.; Cai, J.Q.; Zhou, H.H.; Zhu, Y.S. H19 lncRNA mediates 17 $\beta$ -estradiol-induced cell proliferation in MCF-7 breast cancer cells. *Oncol. Rep.* **2015**, *33*, 3045–3052. [[CrossRef](#)] [[PubMed](#)]
32. Tsang, W.P.; Enders, K.O.N.; Simon, S.M.N.; Jin, H.; Yu, J.; Sung, J.J.Y.; Kwok, T.T. Oncofetal H19-derived miR-675 regulates tumor suppressor RB in human colorectal cancer. *Carcinogenesis* **2010**, *31*, 350–358. [[CrossRef](#)] [[PubMed](#)]
33. Vennin, C.; Spruyt, N.; Dahmani, F.; Julien, S.; Bertucci, F.; Finetti, P.; Chassat, T.; Bourette, R.P.; Le Bourhis, X.; Adriaenssens, E. H19 non-coding RNA-derived miR-675 enhances tumorigenesis and metastasis of breast cancer cells by downregulating c-Cbl and Cbl-b. *Oncotarget* **2015**, *6*. [[CrossRef](#)] [[PubMed](#)]
34. Zhai, L.L.; Wang, P.; Zhou, L.Y.; Yin, J.Y.; Tang, Q.; Zhang, T.J.; Wang, Y.X.; Qin, Y.; Lin, J.; Deng, Z.Q. Overexpression of miR-675 in formalin-fixed paraffin-embedded (ffpe) tissues of breast cancer patients. *Int. J. Clin. Exp. Med.* **2015**, *8*, 11195–11201. [[PubMed](#)]
35. Cordero, F.; Ferrero, G.; Polidoro, S.; Fiorito, G.; Campanella, G.; Sacerdote, C.; Mattiello, A.; Masala, G.; Agnoli, C.; Frasca, G.; et al. Differentially methylated microRNAs in prediagnostic samples of subjects who developed breast cancer in the european prospective investigation into nutrition and cancer (epic-italy) cohort. *Carcinogenesis* **2015**, *36*, 1144–1153. [[CrossRef](#)] [[PubMed](#)]
36. Huang, Y.; Zheng, Y.; Jin, C.; Li, W. Long non-coding RNA H19 inhibits adipocyte differentiation of bone marrow mesenchymal stem cells through epigenetic modulation of histone deacetylases. *Sci. Rep.* **2016**. [[CrossRef](#)] [[PubMed](#)]
37. Liu, L.; An, X.; Li, Z.; Song, Y.; Li, L.; Zuo, S.; Liu, N.; Yang, G.; Wang, H.; Cheng, X.; et al. The H19 long noncoding RNA is a novel negative regulator of cardiomyocyte hypertrophy. *Cardiovasc. Res.* **2016**, *111*, 56–65. [[CrossRef](#)] [[PubMed](#)]
38. Cai, B.; Ma, W.; Bi, C.; Yang, F.; Zhang, L.; Han, Z.; Huang, Q.; Ding, F.; Li, Y.; Yan, G.; et al. Long noncoding RNA H19 mediates melatonin inhibition of premature senescence of C-kit<sup>+</sup> cardiac progenitor cells by promoting miR-675. *J. Pineal Res.* **2016**, *61*, 82–95. [[CrossRef](#)] [[PubMed](#)]
39. Costa, V.; Lo Dico, A.; Rizzo, A.; Rajata, F.; Tripodi, M.; Alessandro, R.; Conigliaro, A. MiR-675-5p supports hypoxia induced epithelial to mesenchymal transition in colon cancer cells. *Oncotarget* **2017**, *8*. [[CrossRef](#)] [[PubMed](#)]
40. Chen, S.; Dingfang, B.; Yuanyuan, M.; Zhu, J.; Chen, G.; Sun, L.; Li, T.; Pan, Y.; Wang, X.; et al. H19 Overexpression induces resistance to 1,25(OH)2D3 by targeting VDR through miR-675-5p in colon cancer cells. *Neoplasia* **2017**, *19*, 226–236. [[CrossRef](#)] [[PubMed](#)]
41. Li, X.; Hao, W.; Yao, B.; Xu, W.; Chen, J.; Zhou, X. lncRNA H19/miR-675 axis regulates cardiomyocyte apoptosis by targeting VDAC1 in diabetic cardiomyopathy. *Sci. Rep.* **2016**, *6*. [[CrossRef](#)] [[PubMed](#)]
42. Zhou, Y.W.; Zhang, H.; Duan, C.J.; Duan, C.J.; Gao, Y.; Cheng, Y.D.; He, D.; Li, R.; Zhang, C.F. miR-675-5p enhances tumorigenesis and metastasis of esophageal squamous cell carcinoma by targeting REPS2. *Oncotarget* **2016**, *7*. [[CrossRef](#)] [[PubMed](#)]
43. 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*. [[CrossRef](#)] [[PubMed](#)]
44. Liu, G.; Xiang, T.; Wu, Q.F.; Wang, W.X. Long noncoding RNA H19-derived miR-675 enhances proliferation and invasion via RUNX1 in gastric cancer cells. *Oncol. Res.* **2016**, *23*. [[CrossRef](#)] [[PubMed](#)]
45. 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. [[CrossRef](#)] [[PubMed](#)]

46. Yan, J.; Zhang, Y.; She, Q.; Li, X.; Peng, L.; Wang, X.; Liu, S.; Shen, X.; Zhang, W.; Dong, Y.; et al. Long noncoding RNA H19/miR-675 axis promotes gastric cancer via FADD/Caspase 8/Caspase 3 signaling pathway. *Cell. Physiol. Biochem.* **2017**, *42*, 2364–2376. [[CrossRef](#)] [[PubMed](#)]
47. Shi, Y.; Wang, Y.; Wenkang, L.; Wang, P.; Tao, T.; Zhang, J.; Qian, J.; Liu, N.; You, Y. Long non-coding RNA H19 promotes glioma cell invasion by deriving miR-675. *PLoS ONE* **2014**. [[CrossRef](#)] [[PubMed](#)]
48. Hernandez, J.M.; Elahi, A.; Clark, C.W.; Wang, J.; Humphries, L.A.; Centeno, B.; Bloom, G.; Fuchs, B.C.; Yeatman, T.; Shibata, D. miR-675 mediates downregulation of Twist1 and Rb in AFP-secreting hepatocellular carcinoma. *Ann. Surg. Oncol.* **2013**, *20*, 625–635. [[CrossRef](#)] [[PubMed](#)]
49. He, D.; Wang, J.; Zhang, C.; Shan, B.; Deng, X.; Li, B.; Zhou, Y.; Chen, W.; Hong, J.; Gao, Y.; et al. Down-regulation of miR-675-5p contributes to tumor progression and development by targeting pro-tumorigenic GPR55 in non-small cell lung cancer. *Mol. Cancer* **2015**, *14*. [[CrossRef](#)] [[PubMed](#)]
50. Huang, Y.; Zheng, Y.; Jia, L.; Li, W. Long Noncoding RNA H19 Promotes Osteoblast Differentiation via TGF- $\beta$ 1/Smad3/HDAC Signaling Pathway by Deriving miR-675: H19/miR-675 Promotes Osteogenesis. *Stem Cells* **2015**, *33*, 3481–3492. [[CrossRef](#)] [[PubMed](#)]
51. Gao, W.L.; Liu, M.; Yang, Y.; Xu, Y.; Li, R.; Deng, Q.; Sun, H.; Wang, S. The Imprinted H19 gene regulates human placental trophoblast cell proliferation via encoding miR-675 that targets nodal modulator 1 (NOMO1). *RNA Biol.* **2012**, *9*, 1002–1010. [[CrossRef](#)] [[PubMed](#)]
52. Sun, T.; Leung, F.; Lu, W. miR-9-5p, miR-675-5p and miR-138-5p damages the strontium and LRP5-mediated skeletal cell proliferation, differentiation, and adhesion. *Int. J. Mol. Sci.* **2016**, *17*. [[CrossRef](#)] [[PubMed](#)]
53. Dey, B.K.; Pfeifer, K.; Dutta, A. 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.* **2014**, *28*, 491–501. [[CrossRef](#)] [[PubMed](#)]
54. Chen, S.W.; Wang, P.Y.; Liu, Y.C.; Sun, L.; Zhu, J.; Zuo, S.; Ma, J.; Li, T.Y.; Zhang, J.L.; Chang, G.W.; et al. Effect of long noncoding RNA H19 overexpression on intestinal barrier function and its potential role in the pathogenesis of ulcerative colitis. *Inflamm. Bowel Dis.* **2016**, *22*, 2582–2592. [[CrossRef](#)] [[PubMed](#)]
55. Kim, N.H.; Choi, S.H.; Lee, T.R.; Lee, C.H.; Lee, A.Y. Cadherin 11, a miR-675 target, induces N-cadherin expression and epithelial-mesenchymal transition in melasma. *J. Investig. Dermatol.* **2014**, *134*, 2967–2976. [[CrossRef](#)] [[PubMed](#)]
56. Kim, N.H.; Choi, S.H.; Kim, C.H.; Lee, C.H.; Lee, T.R.; Lee, A.Y. Reduced MiR-675 in exosome in H19 RNA-related melanogenesis via MITF as a direct target. *J. Investig. Dermatol.* **2014**, *134*, 1075–1082. [[CrossRef](#)] [[PubMed](#)]
57. Keniry, A.; Oxley, D.; Monnier, P.; Kyba, M.; Dandolo, L.; Smits, G.; Reik, W. The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. *Nat. Cell Biol.* **2012**, *14*, 659–665. [[CrossRef](#)] [[PubMed](#)]
58. Adriaenssens, E.; Lottin, S.; Berteaux, N.; Hornez, L.; Fauquette, W.; Fafeur, V.; Peyrat, J.P.; Le Bourhis, X.; Hondermarck, H.; Coll, J.; et al. Cross-talk between mesenchyme and epithelium increases H19 gene expression during scattering and morphogenesis of epithelial cells. *Exp. Cell Res.* **2002**, *275*, 215–229. [[CrossRef](#)] [[PubMed](#)]
59. Tay, Y.; Kats, L.; Salmena, L.; Weiss, D.; Tan, S.M.; Ala, U.; Karreth, F.; Poliseno, L.; Provero, P.; Di Cunto, F.; et al. Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. *Cell* **2011**, *147*, 344–357. [[CrossRef](#)] [[PubMed](#)]
60. Peng, F.; Li, T.T.; Wang, K.L.; Xiao, G.Q.; Wang, J.H.; Zhao, H.D.; Kang, Z.J.; Fan, W.J.; Zhu, L.L.; Li, M.; et al. H19/let-7/LIN28 reciprocal negative regulatory circuit promotes breast cancer stem cell maintenance. *Cell Death Dis.* **2017**, *8*. [[CrossRef](#)] [[PubMed](#)]
61. Li, Z.; Li, Y.; Li, Y.; Ren, K.; Li, X.; Han, X.; Wang, J. Long non-coding RNA H19 promotes the proliferation and invasion of breast cancer through upregulating DNMT1 expression by sponging miR-152. *J. Biochem. Mol. Toxicol.* **2017**. [[CrossRef](#)] [[PubMed](#)]
62. Zhou, W.; Ye, X.L.; Xu, J.; Cao, M.G.; Fang, Z.Y.; Li, L.Y.; Guan, G.H.; Liu, G.; Qian, Y.H.; Xie, D. The lncRNA H19 mediates breast cancer cell plasticity during EMT and MET plasticity by differentially sponging miR-200b/c and let-7b. *Sci. Signal.* **2017**, *10*. [[CrossRef](#)] [[PubMed](#)]
63. Zhao, T.F.; Jia, H.Z.; Zhang, Z.Z.; Zhao, X.S.; Zou, Y.F.; Zhang, J.; Chen, X.F. LncRNA H19 regulates ID2 expression through competitive binding to hsa-miR-19a/b in acute myelocytic leukemia. *Mol. Med. Rep.* **2017**, *16*, 3687–3693. [[CrossRef](#)] [[PubMed](#)]

64. Lv, M.; Zhong, Z.; Huang, M.; Tian, Q.; Jiang, R.; Chen, J. lncRNA *H19* regulates epithelial-mesenchymal transition and metastasis of bladder cancer by miR-29b-3p as competing endogenous RNA. *Biochim. Biophys. Acta* **2017**, *1864*, 1887–1899. [[CrossRef](#)] [[PubMed](#)]
65. Huang, Z.W.; Tian, L.H.; Yang, B.; Guo, R.M. Long noncoding RNA *H19* acts as a competing endogenous RNA to mediate CTGF expression by sponging miR-455 in cardiac fibrosis. *DNA Cell Biol.* **2017**. [[CrossRef](#)] [[PubMed](#)]
66. Ghazal, S.; McKinnon, B.; Zhou, J.; Mueller, M.; Men, Y.; Yuang, L.; Mueller, M.; Flannery, C.; Huang, Y.; Taylor, H.S. *H19* lncRNA alters stromal cell growth via IGF signaling in the endometrium of women with endometriosis. *EMBO Mol. Med.* **2015**, *7*, 996–1003. [[CrossRef](#)] [[PubMed](#)]
67. Wu, W.; Hu, Q.; Nie, E.; Yu, T.; Wu, Y.; Zhi, T.; Jiang, K.; Shen, F.; Wang, Y.; Zhang, J.; et al. Hypoxia induces *H19* expression through direct and indirect Hif-1 $\alpha$  activity, promoting oncogenic effects in glioblastoma. *Sci. Rep.* **2017**, *7*. [[CrossRef](#)] [[PubMed](#)]
68. 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. The imprinted *H19* lncRNA antagonizes Let-7 microRNAs. *Mol. Cell* **2013**, *52*, 101–112. [[CrossRef](#)] [[PubMed](#)]
69. Imig, J.; Brunschweiler, A.; Brümmer, A.; Guennewig, B.; Mittal, N.; Kishore, S.; Tsirikla, P.; Gerber, A.P.; Zavolan, M.; Hall, J. miR-CLIP Capture of a miRNA targetome uncovers a lincRNA *H19*–miR-106a interaction. *Nat. Chem. Biol.* **2014**, *11*, 107–114. [[CrossRef](#)] [[PubMed](#)]
70. Gao, Y.; Wu, F.; Zhou, J.; Yan, L.; Jurczak, J.; Lee, H.Y.; Yang, L.; Mueller, M.; Zhou, X.B.; Dandolo, L.; et al. The *H19*/let-7 double-negative feedback loop contributes to glucose metabolism in muscle cells. *Nucleic Acids Res.* **2014**, *42*, 13799–13811. [[CrossRef](#)] [[PubMed](#)]
71. Liang, W.C.; Fu, W.M.; Wang, F.B.; Sun, Y.X.; Xu, L.L.; Wong, C.W.; Chan, K.M.; Li, G.; Waye, M.M.Y.; Zhang, J.F. *H19* Activates Wnt signaling and promotes osteoblast differentiation by functioning as a competing endogenous RNA. *Sci. Rep.* **2016**, *6*. [[CrossRef](#)] [[PubMed](#)]
72. Li, M.; Chen, H.; Zhao, Y.; Shuming, G.; Cai, C. *H19* functions as a ceRNA in promoting metastasis through decreasing miR-200s activity in osteosarcoma. *DNA Cell Biol.* **2016**, *35*, 235–240. [[CrossRef](#)] [[PubMed](#)]
73. Liu, L.; Yang, J.; Zhu, X.; Li, D.; Lv, Z.; Zhang, X. Long noncoding RNA *H19* competitively binds miR-17-5p to regulate YES1 expression in thyroid cancer. *FEBS J.* **2016**, *283*, 2326–2339. [[CrossRef](#)] [[PubMed](#)]
74. Luo, M.; Li, Z.; Wang, W.; Zeng, Y.; Liu, Z.; Qiu, J. Long non-coding RNA *H19* increases bladder cancer metastasis by associating with EHZ2 and inhibiting E-cadherin expression. *Cancer Lett.* **2013**, *333*, 213–221. [[CrossRef](#)] [[PubMed](#)]
75. Zhuo, C.; Jiang, R.; Lin, X.; Shao, M. lncRNA *H19* inhibits autophagy by epigenetically silencing of DIRAS3 in diabetic cardiomyopathy. *Oncotarget* **2017**, *8*. [[CrossRef](#)] [[PubMed](#)]
76. Monnier, P.; Martinet, C.; Pontis, J.; Stancheva, I.; Ait-Si-Ali, S.; Dandolo, L. *H19* lncRNA controls gene expression of the imprinted gene network by recruiting MBD1. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 20693–20698. [[CrossRef](#)] [[PubMed](#)]
77. Zhou, J.; Yang, L.; Zhong, T.; Mueller, M.; Men, Y.; Zhang, N.; Xie, J.; Jiang, K.; Chung, H.; Sun, X.; et al. *H19* lncRNA alters DNA methylation genome wide by regulating S-adenosylhomocysteine hydrolase. *Nat. Commun.* **2015**, *6*. [[CrossRef](#)] [[PubMed](#)]
78. Deng, Q.; He, B.; Gao, T.; Pan, Y.; Sun, H.; Xu, Y.; Li, R.; Ying, H.; Wang, F.; Liu, X.; Chen, J.; et al. Up-regulation of *91H* promotes tumor metastasis and predicts poor prognosis for patients with colorectal cancer. *PLoS ONE* **2014**, *9*. [[CrossRef](#)] [[PubMed](#)]
79. Vennin, C.; Spruyt, N.; Robin, Y.M.; Chassat, T.; Le Bourhis, X.; Adriaenssens, E. The long non-coding RNA *91H* increases aggressive phenotype of breast cancer cells and up-regulates *H19*/*IGF2* expression through epigenetic modifications. *Cancer Lett.* **2017**, *385*, 198–206. [[CrossRef](#)] [[PubMed](#)]
80. Berteaux, N.; Lottin, S.; Monté, D.; Pinte, S.; Quatannens, B.; Coll, J.; Hondermarck, H.; Curgy, J.J.; 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. [[CrossRef](#)] [[PubMed](#)]
81. Barsyte-Lovejoy, D.; Lau, S.K.; Boutros, P.C.; Khosravi, F.; Jurisica, I.; Andrulis, I.L.; Tsao, M.S.; Penn, L.Z. The c-Myc oncogene directly induces the *H19* noncoding RNA by allele-specific binding to potentiate tumorigenesis. *Cancer Res.* **2006**, *66*, 5330–5337. [[CrossRef](#)] [[PubMed](#)]

82. Dugimont, T.; Montpellier, C.; Adriaenssens, E.; Lottin, S.; Dumont, L.; Iotsova, V.; Lagrou, C.; Stéhelin, D.; Coll, J.; Curgy, J.J. The *H19* TATA-less promoter is efficiently repressed by wild-type tumor suppressor gene product P53. *Oncogene* **1998**, *16*, 2395–2401. [[CrossRef](#)] [[PubMed](#)]
83. Yang, F.; Bi, J.; Xue, X.; Zheng, L.; Zhi, K.; Hua, J.; Fang, G. Up-regulated long non-coding RNA *H19* contributes to proliferation of gastric cancer cells. *FEBS J.* **2012**, *279*, 3159–3165. [[CrossRef](#)] [[PubMed](#)]
84. Zhang, K.; Luo, Z.; Zhang, Y.; Li, Z.; Lichun, W.; Lian, L.; Yang, J.; Song, X.; Liu, J. Circulating lncRNA *H19* in plasma as a novel biomarker for breast cancer. *Cancer Biomark.* **2016**, *17*, 187–194. [[CrossRef](#)] [[PubMed](#)]
85. Zhou, X.; Yin, C.; Dang, Y.; Ye, F.; Zhang, G. Identification of the long non-coding RNA *H19* in plasma as a novel biomarker for diagnosis of gastric cancer. *Sci. Rep.* **2015**, *5*. [[CrossRef](#)] [[PubMed](#)]
86. Chen, J.S.; Wang, Y.F.; Zhang, X.Q.; Lv, J.M.; Li, Y.; Liu, X.X.; Xu, T.P. *H19* serves as a diagnostic biomarker and up-regulation of *H19* expression contributes to poor prognosis in patients with gastric cancer. *Neoplasma* **2016**, *63*, 223–230. [[CrossRef](#)] [[PubMed](#)]
87. Gofrit, O.N.; Benjamin, S.; Halachmi, S.; Leibovitch, I.; Dotan, Z.; Lamm, D.L.; Ehrlich, N.; Yutkin, V.; Ben-Am, M.; Hocherg, A. DNA based therapy with diphtheria toxin-A BC-819: A phase 2b marker lesion trial in patients with intermediate risk nonmuscle invasive bladder cancer. *J. Urol.* **2014**, *191*, 1697–1702. [[CrossRef](#)] [[PubMed](#)]
88. Lavie, O.; Edelman, D.; Levy, T.; Fishman, A.; Hubert, A.; Segev, Y.; Raveh, E.; Gilon, M.; Hichberg, A. A phase 1/2a, dose-escalation, safety, pharmacokinetic, and preliminary efficacy study of intraperitoneal administration of BC-819 (*H19*-DTA) in subjects with recurrent ovarian/peritoneal cancer. *Arch. Gynecol. Obstet.* **2017**, *295*, 751–761. [[CrossRef](#)] [[PubMed](#)]

© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).







## Résumé

Le gène *H19* est soumis à l’empreinte génomique parentale et ne code aucune protéine. Le produit de ce gène, le long ARN non codant (lncRNA) *H19*, agit en tant qu’ARN et est impliqué dans le développement embryonnaire ainsi que dans la tumorigenèse. Le lncRNA *H19* est le précurseur du miR-675. Mes travaux de thèse ont permis d’identifier de nouveaux mécanismes impliqués dans la régulation de la tumorigenèse mammaire par le gène *H19*. Nous avons mis en évidence que le lncRNA *H19* régule négativement la protéine p53 dans les cellules cancéreuses mammaires. Mes travaux ont démontré que le lncRNA *H19* interagit physiquement avec la protéine p53 et MDM2 afin d’induire sa dégradation et empêcher sa translocation dans le noyau. Ce nouveau mode d’action d’*H19* dans les cancers du sein pourrait expliquer le manque de pertinence clinique de l’étude du statut mutationnel de p53 par immunohistochimie dans ce cancer. J’ai également mis en évidence que non seulement le lncRNA *H19* est impliqué dans la régulation des cellules souches cancéreuses, mais également le miR-675-5p. En effet, les tumeurs exprimant des signatures géniques associées aux marqueurs de cellules souches cancéreuses sont des tumeurs qui surexpriment le gène *H19*. De plus, la modulation de l’expression du lncRNA *H19* et de son microARN régule les capacités fonctionnelles associées aux cellules souches cancéreuses mammaires. Pour finir, j’ai initié un projet permettant l’identification, sans *a priori*, des gènes cibles du lncRNA *H19* et de son microARN dans les cellules cancéreuses mammaires.

Mots clefs : *H19*, cancer du sein, long ARN non codant, miR-675, p53, cellules souches cancéreuses.

## Abstract

The *H19* gene is subject to genomic imprinting and does not encode protein. The product of this gene, the long non coding RNA (lncRNA) *H19*, act as an RNA and is involved in development and the tumorigenesis. The *H19* RNA is the precursor of miR-675. My thesis work identified new mechanism involved in the regulation of breast tumorigenesis by *H19*.

We have demonstrated that the lncRNA *H19* negatively regulates the p53 protein in breast cancer cell lines. My work revealed that *H19* interacts with p53 and MDM2 to induce the degradation of p53 and impedes its nuclear localization. This new mechanism of *H19* in breast cancer could explain the lack of clinical relevance of the p53 mutational state measured by immunohistochemistry in breast cancer. My work also revealed that not only the lncRNA *H19* is involved in the regulation of breast cancer stem cells but also the miR-675-5p. Indeed, we have shown a correlation between overexpression of *H19* and expression of a cancer stem cell phenotype in patient tumors. Furthermore, the modulation of *H19* or miR-675 expression regulates the functional capacities associated with breast cancer stem cells. I also initiated a project that will allow the identification of *H19* and miR-675 target genes in breast cancer cell lines. To conclude, I highlighted the implication of the lncRNA *H19* and miR-675 in different process involved in breast cancer tumorigenesis.

Keywords: *H19*, breast cancer, long non coding RNA, miR-675, p53, cancer stem cells.