



**HAL**  
open science

# Caractérisation fonctionnelle de longs ARNs non codants induits par l'hypoxie et impliqués dans l'agressivité des cancers pulmonaires non à petites cellules

Marine Gautier-Isola

## ► To cite this version:

Marine Gautier-Isola. Caractérisation fonctionnelle de longs ARNs non codants induits par l'hypoxie et impliqués dans l'agressivité des cancers pulmonaires non à petites cellules. Cancer. Université Côte d'Azur, 2020. Français. NNT : 2020COAZ6026 . tel-03157405

**HAL Id: tel-03157405**

**<https://theses.hal.science/tel-03157405v1>**

Submitted on 3 Mar 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THÈSE DE DOCTORAT

Caractérisation fonctionnelle de longs  
ARNs non codants induits par l'hypoxie et  
impliqués dans l'agressivité des cancers  
pulmonaires non à petites cellules

**Marine GAUTIER-ISOLA**

Institut de Pharmacologie Moléculaire et Cellulaire (IPMC) – UMR 7275

**Présentée en vue de l'obtention  
du grade de docteur en Sciences de la Vie et  
de la Santé** *Option Interactions moléculaires et  
cellulaires*  
d'Université Côte d'Azur  
**Dirigée par** : Dr Roger Rezzonico  
**Soutenue le** : 13 Novembre 2020

**Devant le jury, composé de :**  
Dr Bernard Mari, DR, IPMC  
Pr Agnès Noël, Pr, ULiège  
Dr David Gilot, MCU, IGDR  
Dr Roger Rezzonico, CR, IPMC



**Caractérisation fonctionnelle de longs ARNs non codants  
induits par l'hypoxie et impliqués dans l'agressivité  
des cancers pulmonaires non à petites cellules**

**Jury :**

**Président du jury:**

Dr Bernard Mari, Directeur de recherche (DR2), IPMC-CNRS (UMR 7275)

**Rapporteurs:**

Dr Agnès Noël, Professeur, Université de Liège/ GIGA Cancer

Dr David Gilot, Maître de conférences + HDR, Université de Rennes 1/IGDR

**Invité:**

Dr Roger Rezzonico, Chargé de recherche (CR1 + HDR), IPMC-CNRS (UMR 7275)



## **Caractérisation fonctionnelle de longs ARNs non codants induits par l'hypoxie et impliqués dans l'agressivité des cancers pulmonaires non à petites cellules**

Les cancers broncho-pulmonaires non à petites cellules, et plus particulièrement les adénocarcinomes (ADC), constituent la première cause de mortalité due au cancer dans les pays développés. Malgré des prises en charge précoces, leur fort taux de récurrence, nécessite l'identification de nouveaux marqueurs pronostics et de nouvelles cibles thérapeutiques. Dans cette optique, nous nous intéressons à l'hypoxie, un facteur d'agressivité tumoral, et à la famille émergente des longs ARNs non codants qui régulent l'expression génique par le biais de complexes ribonucléoprotéiques.

Des analyses transcriptomiques de cohortes de patients d'ADC pulmonaires ont permis l'identification d'une quarantaine de longs ARNs non codants (lncARN) corrélés au statut hypoxique des tumeurs et/ou à un mauvais pronostic vital. Nous nous sommes focalisés sur deux candidats, NLUCAT1 et LINC01116, induits par l'hypoxie et associés à une diminution de la survie des patients. NLUCAT1 est un transcrite nucléaire de 9800 nt dont l'inactivation par le système CRISPR/Cas9 réduit les propriétés prolifératives et invasives des cellules et augmente la production de RLO et l'apoptose induite par le cisplatine ou un stress oxydatif. L'analyse comparative des transcriptomes de cellules invalidées ou non pour NLUCAT1 a révélé son implication dans la régulation du stress oxydatif via un rétrocontrôle positif sur des gènes de la réponse anti-oxydante.

Par ailleurs, LINC01116 est cytosolique et exprimé conjointement dans les cellules cancéreuses et dans les cellules endothéliales du microenvironnement tumoral. J'ai montré que des siARNs réduisent l'adhésion et augmentent la perméabilité trans-endothéliale suggérant des défauts au niveau des contacts intercellulaires. J'ai ensuite entrepris l'étude du mode d'action de LINC01116 par l'identification de ses partenaires protéiques. Des expériences de "RNA pulldown" couplées à de la spectrométrie de masse ont permis d'identifier les protéines ILF3 et PABPC1. Ces interactions ont été validées par des expériences inverses d'immunoprécipitation d'ARN (RIP). L'implication de ces protéines dans le mode d'action de LINC01116 nécessite des études complémentaires.

L'ensemble des résultats collectés durant ma thèse ont mis en évidence l'action pro-tumorale du transcrite NLUCAT1 dans les ADC et l'implication de LINC01116 dans la modification du microenvironnement vasculaire tumoral. Ces transcrits, associés à un mauvais pronostic des ADC, pourraient représenter des cibles thérapeutiques potentielles pour la prise en charge des patients.

**Mots-clés: Hypoxie, long ARN non codant, cancer broncho-pulmonaires**

---

### **Functional characterization of long non coding RNA induced by hypoxia and associated with non small cell lung carcinoma aggressiveness**

Lung cancers, and notably Lung Adenocarcinomas (LUAD) are the leading cause of cancer death worldwide. Their high rate of recurrence despite early, requires new prognostic markers and new therapeutic targets. The combined study of local cohort (CHU of Nice) and large scale (TCGA) transcriptomes of LUAD allowed the identification of a shortlist of 28 long non-coding RNAs (lncRNA) correlated with hypoxia, a factor of tumor aggressiveness, and a poor prognosis. lncRNAs are transcripts that modulate gene expression through the recruitment of proteins and/or nucleic acids and represent an interesting source of new therapeutic targets. Two lncRNAs candidate were selected and molecular characterization was undertaken by sequencing, RT-PCR and smRNA FISH and concern the nuclear lncRNA NLUCAT1 of 9,8kb and the cytosolic lncRNA LINC01116 of 1,2kb. Experiments with loss of function via CRISPR/Cas9 systems, interference RNA and gain of function allowed to characterize the function of these transcripts. Inactivation of NLUCAT1 by CRISPR/Cas9 reduces proliferation, migration, invasion and increases cisplatin sensitivity and ROS production. Bioinformatic analysis of transcriptomes from cells invalidated or not for NLUCAT1 has demonstrated its involvement in the mechanisms of regulation of oxidative stress via a positive feedback from the NRF2 antioxidant pathway. On the other hand, lncRNA LINC01116 is mainly expressed in endothelial cells of the tumor microenvironment and its inhibition by interfering RNA reduces the adhesion capacities and increases the permeability of the endothelium. Mechanistic characterization was performed for LINC01116 via RNA pulldown-MS co-precipitation experiments and identified a list of potential partner proteins. The proteins of RNA metabolism and stability, ILF3 and PABPC1 were identified and their interactions with LINC01116 were validated by RNA immunoprecipitation (RIP).

Overall, during my thesis, I determine the pro-tumoral action of the NLUCAT1 in LUAD and the involvement of LINC01116 in the modification of the tumor microenvironment. These two transcripts could represent potential therapeutic targets in the management of lung cancer.

**Keywords: Hypoxia, Long non coding RNA, Lung cancer.**

## Remerciements

En premier lieu, je tiens à remercier les membres de mon Jury de thèse d'avoir accepté d'évaluer mes travaux de recherches. Merci à vous, **Prof. Agnès Noël** et **Dr. David Gilot** pour le temps que vous avez accordé à la lecture de mon manuscrit.

Je remercie particulièrement mon directeur de thèse, **Roger Rezzonico**. Merci pour tout le savoir que tu as partagé avec moi et la confiance que tu m'as accordé durant ces 4 années ½ à l'IPMC. Merci pour cette bonne entente que nous avons eu, les discussions enrichissantes et l'humour qui ont rythmé ces dernières années. Merci surtout pour ta patience, notamment lors des phases de rédaction et des préparations aux conférences. J'espère que ta rigueur scientifique me suivra tout au long de ma carrière et me permettra de devenir une chercheuse accomplie dans les années à venir !

Merci aussi à mon chef d'équipe, **Bernard Mari** ou « Papy Bernie » que j'ai tant taquiné durant ces quelques années ! Toujours de bons conseils, surtout en matière de spiritueux, tu sais développer nos ambitions scientifiques et personnelles, alors merci d'avoir toujours cru en moi durant ces quelques années et de m'avoir épaulé à Nice, à Sophia-Antipolis et aussi à Vancouver !

Je remercie également tous les membres de l'équipe de Bernard Mari, pour leur joie de vivre, nos rires partagés, l'entraide et la bienveillance. Je voudrais remercier **Laura** avec qui j'ai travaillé au début de ma thèse, ta bonne humeur a rendu mon intégration au sein de l'équipe très agréable. Un remerciement tout spécial aussi à **Caroline** avec qui nous avons partagé nos passions des chats et du tricot ! Ce fut également une véritable joie de travailler avec le couple italien le plus glamour et le plus gentil de l'IPMC, **Serena** et **Alberto**. Merci à **Georges** également pour son écoute et ses conseils qui me serviront pour la suite mais aussi pour son humour et ses références anglo-saxonnes pas toujours tout public ;) ! Merci également à **Marin** pour sa disponibilité et sa collaboration plus que bienvenu pour les aspects bio-informatiques, Issa Nissa ;) Surtout j'espère pouvoir vous dire à tous, à bientôt !

Merci également à l'équipe de Pascal Barbry, chez qui j'ai effectué mon stage de Master 2 et le début de ma thèse avant notre séparation, heureusement cela n'a pas stoppé nos bonnes relations et ce fut un plaisir de partager les repas, les cafés et mêmes quelques bières avec les membres de cette super équipe ! Merci à vous **Nicolas P., Nicolas N., Karine R., Agnès P., Kevin, Sandra, Marie D., Agnès F., Marie C., Cédric, Christophe, Laure-E, Margot, Magali** et **Gilles**. Merci surtout à mon amie **Carole**, tu es une battante et surtout une femme extrêmement bienveillante et une belle rencontre pour moi !

Merci également aux membres de l'IPMC et aux membres des services de plateforme pour leur aide sur les projets, leurs adaptabilités, leurs disponibilités et leurs offres de formations ! Un merci tout spécial à **Sophie**, un vrai rayon de soleil !

Merci aussi aux équipes pédagogiques de l'Université de Nice Sophia-Antipolis pour leur accueil dans leurs équipes et la formation dispensée dans l'enseignement qui fut un vrai bonheur pour moi. Un remerciement spécial à **Patricia Lebrun, Anny Cupo, Julie Dunys, Laurent Counillon, Delphine Ciais** et **Dominique Colinet** pour votre confiance !

Ensuite, je tiens à remercier mes proches, ma famille et mes amis. Plus particulièrement, je veux remercier la personne qui me soutient depuis de nombreuses années, d'ailleurs on peut le dire depuis le début de mes aventures universitaires, et qui me pousse à aller de l'avant en m'offrant le réconfort de sa présence et de sa confiance. Merci de ne pas trop râler quand je rentre tard, de prendre le relais les jours de grande fatigue, de prendre soin de notre famille et de sa composante féline, d'être mon ancre, merci à toi **Yohann**, mon mari. Toi et notre petite famille, vous restez ma priorité. Je voudrai également remercier notre gentil bébé qui a laissé sa mère écrire sa thèse et qui fait déjà notre bonheur, nous t'attendons impatiemment (mais après le 13 Novembre s'il te plaît :)).

Ensuite, je veux remercier mes parents, **Annie et Fabrice**, car d'une part, je ne serai pas là sans eux, c'est une évidence ;) ! Mais d'autre part, ils m'ont toujours soutenue dans mes projets me laissant l'autonomie et la confiance pour me dépasser donc merci de m'avoir amené jusqu'ici avec toute votre bienveillance et votre générosité.

Je voudrais également remercier mon frère, **Julien**, qui est l'un des soleils qui réchauffe ma vie. Je suis tellement fière de toi, du jeune homme que tu deviens et je voudrais te remercier de n'avoir jamais pris offense de mes absences, je serai toujours là pour toi mon Juju.

Ensuite, je remercie ma sœur de cœur, **Mélody**, nous avons grandi ensemble, nous fondons nos familles côte à côte et je suis si fière de te compter parmi les membres de ma famille. Je suis fière également d'être la marraine de ton merveilleux **Nathan**, merci à toi et à **Nicolas** pour votre confiance ! :)

Je remercie aussi mon ange gardien, je sais que tu n'es jamais très loin de moi... Tu es l'une des raisons qui me pousse en avant, merci à toi, **Clémentine**, de m'avoir tant donné. Je pense également à ma grand-mère, **Madeleine**, qui m'a tant appris et en premier lieu, l'envie de s'améliorer et de ne jamais se contenter de ce qui est facile à obtenir.

Je voudrai également remercier mes grands-pères, **Roger et Gérard**, qui sont des exemples d'hommes accomplis et pour lesquels j'éprouve beaucoup de fierté. Merci d'être un soutien dans ma vie.

Je veux remercier aussi toute ma nombreuse famille, mes oncles et mes tantes, mes cousins et mes cousines, tous ceux qui sont si proches de moi et dont le regard bienveillant m'a permis de trouver ma place dans ce monde, merci de votre soutien indéfectible.

Je veux également remercier ma belle-famille pour leur chaleur, merci à vous **Patricia, Jean-Noël, Sylvie, Thérèse, Julia, Béchir, Léna, Issam et Alexandre**.

On dit aussi que les amis sont la famille que l'on se choisit, et cela ne pourrait pas être plus vrai lorsque je pense à la « Boolteam » : **Michaël V., Michaël J, Silvia, Jérémy, Giulia, Lola et Raphaël**. Votre présence dans nos vies et en réel réconfort et bien que nous serons séparés pendant quelques temps, je sais que nos liens n'en seront que plus fort à notre retour !

Cette famille ne serait pas non plus complète si vous n'étiez pas là, mes « Copines de la Fac » qui sont devenues depuis mes plus grandes amies. Je le dis souvent, mais clairement je n'aurai pas pu aller jusque-là sans vous, dans les moments de démotivation, vous avez su me rendre mon énergie et dans les moments de doutes ou de joies vous avez toujours été présente pour moi alors merci à vous trois **Anasthasia, Cécilia et Mimi** !

Merci également à **Diane, Lucas, Julien T. Stef', Charlie, Raphaël F., Axel, Laetitia, Dorian, Chloé, Fanny, Lucile, Pierre** d'être dans nos vies et de partager avec nous votre bienveillance !

## TABLE DES MATIERES

<b>I. LISTE DES ABREVIATIONS.....</b>	<b>9</b>
<b>II. LISTE DES ILLUSTRATIONS .....</b>	<b>16</b>
<b>III. INTRODUCTION .....</b>	<b>18</b>
<b>Chapitre 1 : Les cancers broncho-pulmonaires.....</b>	<b>19</b>
<b>A. Les poumons .....</b>	<b>19</b>
<b>B. Epidémiologie du cancer du poumon .....</b>	<b>19</b>
<b>C. Les sous-types histologiques des cancers broncho-pulmonaires .....</b>	<b>22</b>
1. Les cancers à petites cellules.....	22
2. Les cancers non à petites cellules.....	22
3. Stades anatomo-cliniques .....	25
<b>D. Les facteurs de risques .....</b>	<b>27</b>
1. Classification des molécules cancérogènes .....	27
2. Le tabagisme.....	28
3. Les expositions environnementales et professionnelles .....	28
4. Les facteurs génétiques .....	29
5. Les pathologies pulmonaires chroniques.....	30
6. Les agents infectieux.....	30
<b>Chapitre 2 : Les principales caractéristiques des cancers broncho-pulmonaires .....</b>	<b>31</b>
<b>A. Classification moléculaire des cancers broncho-pulmonaires .....</b>	<b>31</b>
1. Hétérogénéité moléculaire des cancers broncho-pulmonaires .....	31
2. Les mutations conductrices majoritaires dans les CBNPC .....	33
3. Les autres mutations conductrices .....	37
<b>B. Le microenvironnement tumoral .....</b>	<b>39</b>
1. Les cellules immunitaires .....	39
2. Les cellules endothéliales .....	43
3. Les fibroblastes, architectes de la matrice extracellulaire (MEC).....	45
4. Les échanges gazeux et l'hypoxie.....	47
<b>C. Traitements des cancers broncho-pulmonaires non à petites cellules .....</b>	<b>53</b>
1. Diagnostic et Médecine personnalisée .....	53
2. Exérèse curative .....	53
3. Chimiothérapie .....	53
4. Radiothérapie.....	54
5. Immunothérapie .....	54
<b>Chapitre 3 : Classification des ARNs non codants.....</b>	<b>57</b>
<b>A. De « l'ADN poubelle » aux ARNs non codants .....</b>	<b>57</b>
<b>B. Les ARNs impliqués dans la traduction.....</b>	<b>59</b>
1. Les ARNs ribosomiques.....	59
2. Les snoARNs .....	59
3. Les ARNs de transfert.....	59
<b>C. Les petits ARNs régulateurs .....</b>	<b>61</b>
1. L'ARN interférence et les microARNs.....	61
2. Les piARNs et la régulation des éléments transposables .....	63
<b>D. Les Longs ARNs non codants .....</b>	<b>64</b>

1. Les LncARNs antisens .....	66
2. Les LncARNs divergents.....	66
3. Les LncARNs intergéniques.....	66
4. Les LncARNs introniques.....	66
5. Les IncARNs circulaires .....	67
<b>Chapitre 4 : Synthèse et mode d'action des longs ARNs non codants.....</b>	<b>69</b>
<b>A. Synthèse et maturation des IncARNs .....</b>	<b>69</b>
1. Transcription des IncARNs.....	69
2. Transport des IncARNs, localisation subcellulaire et potentiel codant .....	77
3. Stabilité et dégradation des IncARNs.....	79
<b>B. Modes d'action des IncARNs .....</b>	<b>81</b>
1. Approches technologiques spécifiques à l'étude du mode d'action des IncARNs.....	81
2. Modes d'action des IncARNs nucléaires.....	84
3. Modes d'action des IncARNs cytosoliques .....	93
<b>Chapitre 5 : Fonctions des longs ARNs non codants.....</b>	<b>103</b>
<b>A. Rôles physiologiques des IncARNs .....</b>	<b>103</b>
1. Rôles des IncARNs au cours du développement embryonnaire .....	103
2. Rôles des IncARNs dans des processus de différenciation cellulaire .....	104
<b>B. Rôles pathologiques des IncARNs.....</b>	<b>109</b>
1. LncARNs impliqués dans des pathologies non tumorales .....	109
2. Les IncARNs impliqués dans les processus de tumorigenèse .....	110
3. Les IncARNs et le microenvironnement tumoral.....	114
4. Les IncARNs impliqués dans les cancers broncho-pulmonaires.....	118
<b>IV. RESULTATS .....</b>	<b>123</b>
<b>Avant-propos.....</b>	<b>124</b>
<b>Partie 1 : NLUCAT1, un hypoxa-lnc nucléaire impliqué dans la regulation du stress oxydatif des adénocarcinomes pulmonaires.....</b>	<b>126</b>
The nuclear hypoxia-regulated NLUCAT1 long non-coding RNA contributes to an aggressive phenotype in lung adenocarcinoma through regulation of oxidative stress.....	127
Discussion et perspectives.....	156
<b>Partie 2 : LINC01116, un hypoxa-lnc cytosolique corrélé à l'agressivité des adénocarcinomes pulmonaires et impliqué dans la régulation du stroma tumoral .....</b>	<b>163</b>
Introduction.....	164
Matériels et Méthodes.....	165
Résultats.....	170
Discussion et perspectives.....	194
<b>V. CONCLUSION.....</b>	<b>204</b>
<b>VI. ANNEXE.....</b>	<b>209</b>
One step CRISPRi efficiently represses gene expression in mammalian cells .....	210
<b>VII. REFERENCES BIBLIOGRAPHIQUES .....</b>	<b>224</b>

## I. LISTE DES ABREVIATIONS

AEC : Alveolar Epithelial Cell  
MEC : Matrice Extracellulaire  
CIRC : Centre International de Recherche sur le Cancer  
OMS : l'Organisation Mondiale de la Santé  
PAH : polycycliques aromatiques hydrocarbonées  
BPCO: Bronchopneumathie Obstructive Chronique  
VIH : Virus de l'Immunodéficience humaine  
ADN : Acide desoxyribonucléique  
CBPC : Cancer Broncho-pulmonaire à petite Cellule  
CBPNP : Cancer Broncho-pulmonaire Non à Petite Cellule  
ADC: Adénocarcinome  
LUAD: Lung Adenocarcinoma  
SCC: Squamous Cancer Cell  
LCC: Large-cell Cancer  
TNM: Tumour-Node-Metastasis  
EGFR: Epidermal Growth Factor Receptor  
PI3K : Phosphoinositide 3 Kinase  
AKT/PKB : Proteine Kinase B  
mTOR : mechanistic target of rapamycin  
RAS : Rat Sarcoma  
RAF : **R**apidly **A**ccelerated **F**ibrosarcoma  
MAPK/MEK : Mitogen-Activated protein kinase  
COSMIC : Catalog of somatic mutations in cancer  
GEF : Guanine nucleotide exchange factor  
GDP : Guanine Di-Phosphate  
GAP : GTPase Activating Protein  
MDM2 : Murin Double Minute 2  
Bax : Bcl-2 associated X  
SV40 (simian virus 40)  
p53RE :p53 Response Element  
SREBPs : Sterol Regulatory element binding protein  
TCGA : The Cancer Genome Atlas  
NK : Natural Killer  
DC : Cellules dendritiques  
LT : Lymphocyte T  
CPAg : Cellule présentatrice d'antigène  
MCP1 : Monocyte chemoattractant protein 1  
TAMs : Tumor-associated macrophage  
IL-10 : Interleukin 10  
PD-L1 : Programmed Death Ligand 1  
TGF $\beta$  : Transforming Growth Factor  $\beta$   
TNF $\alpha$  : Tumor Necrosis Factor  $\alpha$   
LEC : Lymphatic endothelial cell  
VEGF : Vascular Endothelial Growth Factor  
PDGF : Platelet derived growth factor  
CAF : Cancer associated fibroblast  
RLO : Radicaux Libres oxygénés  
IFN $\gamma$  : Interféron  $\gamma$

MMP : Métalloprotéases  
Ang2 : Angiopoietin 2  
FGF : Fibroblast Growth Factor  
PAI-1 : Plasminogen activator inhibitor-1  
NO : Nitric Oxyde  
COX-2 : cyclo-oxygenase-2  
TSP-2 : thrombospondin-2  
NICD : Notch Intracellular Domain  
DLL1-4 : Delta-Like Canonical Notch 1 Ligand 1-4  
ICAM1 : Intercellular adhesion molecule 1  
VCAM1 : Vascular cell adhesion molecule 1  
 $\alpha$ -SMA :  $\alpha$ -smooth muscle actin  
MCP1 : Monocyte chemotactic protein 1.  
HIF : Hypoxia Induced Factor  
EPAS1 : Endothelial PAS Domain Protein 1  
HRE : Hypoxia Response Element  
PHD : Prolyl hydroxylase  
p-VHL : Von Hippel-Lindau  
CA9 : Carbonic Anhydrase 9  
HAF : Hypoxia Associated Factor  
NF- $\kappa$ B : Nuclear Factor kappa B  
I $\kappa$ B : Inhibitor of Nuclear Factor kappa B  
IKK : I $\kappa$ B Kinase  
KEAP1 : Kelch Like ECH Associated Protein 1  
GLUT1 : Glucose Transporter 1  
LDHA : Lactate Deshydrogenase A  
PDK1 : Pyruvate déshydrogénase 1  
PKM2 : Pyruvate kinase isozyme M1/M2  
VEGFR2 : Récepteur au VEGF 2  
NOS : Nitric acid synthetase  
TEM : Transition épithélio-mésenchymateuse  
LOX : Lysyl oxydase  
ABCA2 : ATP Binding Cassette Subfamily A Member 2  
MDR1 : Multidrug resistance gene 1  
IASLC: International Association for the Study of Lung Cancer  
ATS: American Thoracic Society  
ERS: European Respiratory Society  
CTLA4: Cytotoxic T-Lymphocyte antigen 4  
TAA : Tumor Associated Antigen  
ALK: Anaplastic Lymphoma Kinase  
BRAF1: B-Raf Proto-Oncogene, Serine/Threonine Kinase  
CAR-T : Chimeric Antigen Receptor T cell  
CMH : Complexe Majeur d'Histocompatibilité  
ARN : acide ribonucléique  
ENCODE : ENCyclopedia Of DNA Element  
FANTOM : Fonctionnal Annotation of The Mammalian genome  
miARN : micro-ARN  
snoARN : small nucleolar ARN  
piARN : piwi-ARN  
ARNr : ARN ribosomique  
ARNt : ARN de transfert  
ARNm : ARN messenger

snoRNP : small nucleolar RiboNucleoProtein  
 LncARN : Long Non Coding ARN  
 CHS : chalcone synthetase  
 DGCR8 : DiGeorge Syndrome Critical Region Gene 8  
 miARNdb : micro-ARN double brin  
 RISC: RNA-induced Silencing Complex  
 EXP5: Exportin 5  
 siRNA: small interfering RNA  
 shRNA: short-hairpin-RNA  
 piARN: Piwi-ARN  
 ARNpol: ARN polymérase  
 RBP: RNA Binding Protein  
 COOLAIR: Cold Induced Antisense Transcription of Flowering Locus C  
 GENCODE: Encyclopedia of genes and gene variants  
 circRNA: circular RNA  
 LincRNA: Long Intergenic Non Coding RNA  
 NAT: Natural Antisens Transcript  
 SINEUPs: SINEB2 sequence to UP-regulate translation  
 TLR: Toll-like recptor  
 MYMLR: Myc Modulating LncRNA  
 EVX1: Even-Skipped Homeobox 1  
 NORAD: Non coding RNA activated by DNA damage  
 Kcnq1ot1: KCNQ Opposite Strand/ Antisens Transcript 1  
 KCNQ1: Potassium Voltage-gated channel subfamily Q member 1  
 PCR2: Polycomb repressor complex 2  
 TFIID: Transcription factor IID  
 TBP: TATA-box binding protein  
 TAF: TBP associated factor  
 CDK7 : Cyclin dependant kinase 7  
 CTD: C-terminal Domain of RNA pol II  
 NELF: Negative Elongation Factor  
 CE: Capping Enzyme  
 CDK9: Cyclin dependant kinase 9  
 p-TEF: Positive transcription elongation factor  
 H3K9me: Méthylation de la lysine 9 de l'histone 3  
 H3K36me: Méthylation de la lysine 36 de l'histone 3  
 H3K27Ac: Acétylation de la lysine 27 de l'histone 3  
 H3K4me3: Tri-méthylation de la lysine 4 de l'histone 3  
 HMT: Histone Méthyltransférase  
 HDM: histone déméthylase  
 HAT: histone acétyltransférase  
 HDAC: histone désacétylase  
 FACT: Facilitate Chromatin Transport  
 CAF-1: Chromatin Assembly Factor 1  
 PSPC1: Paraspeckles component 1  
 PSF: Protein splicing factor  
 FCP1/CTDP1: CTD phosphatase 1  
 CPSF: Cleavage and polyadenylation specificity factor  
 CstF: Clivage stimulation Factor  
 PAS: Polyadénylation signal  
 XRN2: 5'-3' Exoribonuclease 2  
 MALAT1: Metastasis associated Lung adenocarcinoma Transcript 1



NEAT1: Nuclear Paraspeckles Assembly Transcript 1  
 Nrd1 : Nardilysin 1  
 meRNA: multiexonic RNA  
 sORF: small Open Read Frame  
 AU-rich: motifs riche en séquence AU  
 ARE-BP:AU-Rich Element Binding Protein  
 PABP1: Poly(A) Binding Protein 1  
 HuR/ELAVL1: Human antigen R  
 Tia-1: T-Cell-Restricted Intracellular Antigen-1  
 TTP: Tristetraproline  
 AUF1/ HNRNPD : AU-Rich Element RNA-Binding Protein 1  
 ILF3: Interleukin enhanced binding factor 3  
 NMD: Nonsens mediated decay  
 NSD: Non-stop mRNA decay  
 NGD: Non-go mRNA decay  
 Dcp1/Dcp2: Decapping protein 1/2  
 Caf1/CCR4/Not: Carbon Catabolite Repressor Protein (CCR4)-Associative Factor 1  
 DIS3: 3'-5' Exoribonuclease  
 ChIRP: Chromatine Isolation by RNA purification  
 CHART: Capture Hybridization Analysis of RNA Targets  
 RIP: RNA Immunoprecipitation  
 CLIP: Crosslinked Immunoprecipitation  
 RAP: RNA Antisens Purification  
 LIGR-seq: LIGation of interacting RNA followed by high-throughput sequencing  
 GRID-seq: Global RNA Interactions with DNA by Deep sequencing  
 Xist: X inactive specific transcript  
 SAF: Scaffold Attachment Factor  
 SHARP: SMART/HDAC1-Associated Repressor Protein  
 HNRNPK: Heterogeneous Nuclear Ribonucleoprotein K  
 DNMT: ADN methyltransferase  
 HOTAIR: HOX Transcript Antisens RNA  
 CCND1: Cyclin D1  
 FUS/TLS: Translocated in sarcoma  
 THRIL: TNF and HNRNPL related immunoregulatory lncRNA  
 lncRNA LET : lncRNA Low Expression in Tumor  
 PANDA : P21 Associated lncRNA DNA damage Activated  
 LEENE: Enhancer-associated lncRNA that enhances eNOS expression  
 KLF2/4: Kruppel Like Factor 2/4  
 Med1: Mediator Complex Subunit 1  
 SAR: Scaffold-Attached Region  
 eRNA: enhancer RNA  
 Lockd: lncRNA downstream of Cdkn1b  
 CDKN1: Cyclin Dependent Kinase Inhibitor 1  
 lincRNA-Cox2: Long intergenic RNA cyclooxygenase 2  
 CCAT2: Colon Cancer Associated Transcript 2  
 GLS: Glutaminase  
 SNP: Single Nucleotide Polymorphism  
 MBL: muscleblind  
 arcARN: ARN architecturaux  
 SR: Serine-Arginine-Rich splicing factors  
 PRP6: Pre-mRNA Processing Factor 6  
 Fos: FBJ Murine Osteosarcoma

Irs2: Insulin Receptor Substrate 2  
 ceARN: Competing endogenous ARN  
 PTENP1: PTEN pseudogene 1  
 PTEN: Phosphatase And Tensin Homolog  
 TINCR: Differentiation induced lncRNA  
 STAU-1: Staufen1  
 UPF1/UPF2: Up-Frameshift Suppressor 1 Homolog  
 SMD: STAU1-mediated RNA decay  
 SERPINE1: Serine (Or Cysteine) Proteinase Inhibitor  
 IMP2:IGF2 mRNA-binding protein  
 Uch11: Ubiquitin C-Terminal Hydrolase L1  
 IRES: Internal Ribosome Entry Sequence  
 SENCR: Smooth muscle and endothelial cell enriched  
 CKAP4: Cytoskeleton Associated Protein 4  
 MetaLnc9/Lnc00963 : Long Intergenic Non-Protein Coding RNA 963  
 PGK1: Phosphoglycerate Kinase 1  
 Lnc-DC: LncRNA of dendritic cells  
 SHP1/PTPN6 : Protein Tyrosine Phosphatase Non-Receptor Type 6  
 STAT3: Signal Transducer And Activator Of Transcription 3  
 NKILA : NF-KappaB Interacting LncRNA  
 PB : p-bodies ou processing bodies  
 GS : granules de stress  
 LSM14A/LSM14B : LSM Family Member 14A/14B  
 IGF2BP1 : Insulin Like Growth Factor 2 mRNA Binding Protein 1  
 G3BP : GTPase Activating Protein (SH3 Domain) Binding Protein 1  
 Gas5 : Growth Arrest Specific 5  
 BCYRN1 : Brain Cytoplasmic RNA 1  
 RMRP : RNA Component Of Mitochondrial RNA Processing Endoribonuclease  
 CSF : Colony Stimulated Factor  
 LHDB : Lactate déshydrogénase B  
 Neo : neomycine  
 lnc-Brn1a/PANTR1: POU3F3 Adjacent Non-Coding Transcript 1  
 lnc-Brn1b/LINC01159 : Long Intergenic Non-Protein Coding RNA 1159  
 FabI/ELDR : EGFR Long Non-Coding Downstream RNA  
 linc-Enc1/LINC01157 : Long Intergenic Non-Protein Coding RNA 1157  
 Fendrr : FOXF1 Adjacent Non-Coding Developmental Regulatory RNA  
 Haunt : HOXA upstream non-coding transcript  
 Mdgt/HAGLR : HOXD Antisense Growth-Associated Long Non-Coding RNA  
 Celr : cerebellum expressed regulatory RNA  
 Crnde : Colorectal Neoplasia Differentially Expressed  
 Spasm/KANTR : KDM5C Adjacent Non-Coding Transcript  
 linc-Pint : Long Intergenic Non-Protein Coding RNA, P53 Induced Transcript  
 linc-Ppara : LincRNA of Peroxisome Proliferator Activated Receptor Alpha  
 mESc : mouse embryonic stem cell  
 Oct4 : Octamer-Binding Protein 4  
 GFP : Green Fluorescent protein  
 TUNA/Megamind : Tcl1 Upstream Neuron-Associated lincRNA  
 PTBP1 : Polypyrimidine Tract Binding Protein 1  
 Sox2 : SRY-Box Transcription Factor 2  
 RMST : Rhabdomyosarcoma 2-associated transcript  
 ANCR : anti-differentiation ncRNA  
 KRT1, KRT10 et GHRL3

DEANR1 : Definitive Endoderm-Associated lncRNA 1  
 FOXA2 : forkhead box protein A2  
 MESP1 : Mesoderm Posterior BHLH Transcription Factor 1  
 MIXL1 : Mix Paired-Like Homeobox  
 Rroid : RNA-demarcated Regulatory region of *Id2*  
 Id2 : Inhibitor Of DNA Binding 2  
 NANCI : Nkx2.1-Associated Noncoding Intergenic RNA  
 Bvht : BRAVEHEART  
 Lnc-RAP-n : Long non coding RNA Regulated in Adipogenesis  
 ORLNC1 : osteoporosis-related lncRNA 1  
 BACE1 : B amyloid cleavage enhanced 1  
 CCL5/RANTES : Regulated Upon Activation, Normally T-Expressed, And Presumably Secreted  
 TERRA : TElomeric Repeat-containing RNA  
 CTCF : CCCTC-Binding Factor  
 TCF7 : Transcription Factor 7  
 Wnt : Wingless-Type  
 LncRNA THOR : Testis Associated Oncogenic lncRNA  
 LUNAR1 : Leukemia-Associated Non-Coding IGF1R Activator RNA 1  
 IGF1R : Insulin Growth Factor Receptor 1  
 PCGEM1 : Prostate-Specific Transcript 1 (Non-Protein Coding)  
 ANRIL/CDKN2B-as : Antisense Noncoding RNA In The INK4 Locus  
 p15/CDKN2B : Cyclin-dependent kinase inhibitor 2B  
 UCA1 : Urothelial Cancer Associated 1  
 HOTTIP : HOXA Distal Transcript Antisense RNA  
 HOXA13 : Homeobox A13  
 TUG1 : Taurine Up-Regulated 1  
 LIMK2b : LIM Domain Kinase 2  
 BRCA1 : Breast Cancer Associated 1  
 LncRNA MEG3 : lncRNA Maternally Expressed 3  
 lncRNA-HIT : HOXA transcript induced by TGF $\beta$   
 LncRNA-ATB : lncRNA activated by TGF- $\beta$   
 ZEB1 : Zinc finger E-box-binding homeobox 1  
 TGF $\beta$ -AS1 : TGF $\beta$  antisens 1  
 BANCR : BRAF-Activated Non-Protein Coding RNA  
 LINK-A : long intergenic non-coding RNA for kinase activation  
 Lnc-ROR : Long Intergenic Non-Protein Coding RNA, Regulator Of Reprogramming  
 LncRNA H19 : H19 Imprinted Maternally Expressed Transcript  
 HINCUT1 : Hypoxia Induced Non coding Ultra conserved Transcript  
 EFNA3-as : Ephrin 3 antisens RNA  
 LncRNA DILC : lncRNA downregulated in liver cancer stem cells  
 TSLNC8 : Tumor Suppressor lncRNA On Chromosome 8p12  
 FOXO1 : Forkhead Box O1  
 NKX2-1-AS1 : NKX2 Antisens RNA 1  
 LCT : lymphocytes cytotoxiques T  
 Tregs lymphocytes T régulateurs  
 IDO1 : Indoleamine 2,3-dioxygenase 1  
 INSR : récepteur à l'insuline  
 HULC : Highly Up-regulated in Liver Cancer (non-protein coding)  
 STEEL : spliced-transcript endothelial-enriched lncRNA  
 EGFL7OS : Epidermal Growth Factor-Like Protein 7 Opposite Strand RNA  
 Tie1-AS : Tyrosine Kinase Like And EGF Like Domains 1 antisens RNA  
 PVT1 : Plasmacytoma Variant Translocation 1

LUADT1 : Lung Adenocarcinoma Associated Transcript 1  
CASC9.5 : Cancer Susceptibility 9  
Lnc-Sox2ot : SOX2 Overlapping Transcript  
EPEL : E2F-mediated cell proliferation enhancing lncRNA  
FEN1 : Flap Structure-Specific Endonuclease 1  
LKB1 : Liver Kinase B1  
NONO : Non-POU Domain Containing Octamer Binding  
CRTC : CREB Regulated Transcription Coactivator  
LCAT1 : Lung Cancer Associated Transcript 1  
VELUCT : Viability Enhancing Lung Cancer Transcript  
PILAR1 : Prognostic in Lung Adenocarcinoma lncRNA-1  
ALAL-1 : Amplified lncRNA Associated with Lung cancer -1  
SART3 : Squamous cell carcinoma antigen recognized by T cells 3  
USP4 : Ubiquitin specific protease 4  
ELIT-1/LINC00842 : EMT associated lncRNA induced by TGFb1  
NSCLCAT1 : non-small cell lung cancer—associated transcript-1  
LIN28B-AS1 : Lin-28 Homolog B Antisens 1  
MUC5B-AS1 : Mucin 5B, Oligomeric Mucus/Gel-Forming Antisens 1  
RGMB-AS1 : Repulsive Guidance Molecule BMP Co-Receptor B antisens 1  
NBR2 : Neighbor Of BRCA1 lncRNA 2  
PICART1 : P53 Inducible Cancer Associated RNA Transcript 1  
TARID : TCF21 Antisense RNA Inducing Promoter Demethylation  
GADD45A : Growth Arrest And DNA Damage Inducible Alpha  
TCF21 : Transcription Factor 21  
SPRY4-IT1 : Sprouty RTK Signaling Antagonist 4 Intronic Transcript 1  
P53RRA : P53 Related lncRNA  
YBX1 : Y-Box Binding Protein 1  
GIAT4RA : lncRNA between GINS4 and GPAT4  
LSH : Lymphoid-Specific Helicase  
Uchl3 : Ubiquitin C-Terminal Hydrolase L3  
LINP1 : lncRNA In Non-Homologous End Joining Pathway 1  
SCAL1 : Smoke and Cancer Associated lncRNA 1  
ILF3 : Inducible Interleukin Factor  
PABPC1 : Poly(A) Binding Protein Cytoplasmic 1

## II. LISTE DES ILLUSTRATIONS

<b>Figure 1</b> : Représentation schématique des poumons et de l'épithélium respiratoire broncho-pulmonaire. ....	20
<b>Figure 2</b> : Incidence des cancers dans la population française par genre et tissu.....	21
<b>Figure 3</b> : Les sous types histologiques des cancers broncho-pulmonaires.....	24
<b>Figure 4</b> : Classification internationale TNM des cancers pulmonaires en fonction de leur extension anatomique	26
<b>Figure 5</b> : Les caractéristiques des cancers. Adapté de <sup>27</sup> .....	32
<b>Figure 6</b> : Voies de signalisation de l'EGFR et mutations impliquées dans les cancers broncho-pulmonaires.....	34
<b>Figure 7</b> : Rôles suppresseur de tumeur et oncogénique de p53. ....	36
<b>Figure 8</b> : Fréquences des mutations oncogéniques dans les adénocarcinomes pulmonaires. Adapté de <sup>57</sup> .....	38
<b>Figure 9</b> : Organisation et diversité du microenvironnement tumoral.....	40
<b>Figure 10</b> : La surveillance des cancers broncho-pulmonaires par le système immunitaire.....	42
<b>Figure 11</b> : L'activation des fibroblastes associés au cancer (CAF) favorise l'agressivité tumorale.....	46
<b>Figure 12</b> : Les différentes voies de signalisation de l'hypoxie (NF-κB, HIF1, NRF2) et leur mécanisme de régulation en normoxie. ....	49
<b>Figure 13</b> : Conséquences des voies de signalisation de l'hypoxie.....	51
<b>Figure 14</b> : La réponse immune anti-tumorale et les stratégies de l'immunothérapie. ....	56
<b>Figure 15</b> : Classification des acides ribonucléiques codants et non codants chez les mammifères.....	58
<b>Figure 16</b> : ARN ribosomique, ARN de transfert et snoARN sont les ARN régulateurs de la traduction des ARNm en protéines. ....	60
<b>Figure 17</b> : Les différentes voies de synthèses des microARNs, siARNs et piwiARNs. ....	62
<b>Figure 18</b> : Classification des longs ARNs non codants en fonction de leur localisation génique par rapport aux gènes codants. ....	65
<b>Figure 19</b> : Voies de synthèse des ARNs circulaires par épissage circulaire ou "back-splicing".....	68
<b>Figure 20</b> : Mécanismes d'initiation, de pause et d'élongation transcriptionnelle de l'ARN Pol II. ....	70
<b>Figure 21</b> : Différences dans les mécanismes de transcription des lncARNs et des ARNm. ....	72
<b>Figure 22</b> : Différents mécanismes de terminaison de la transcription et de maturation des ARNs.....	76
<b>Figure 23</b> : Mécanismes de dégradation des ARNs au cours des différentes étapes de leur synthèse et en fonction des compartiments subcellulaires. Adapté de <sup>197</sup> .....	78
<b>Figure 24</b> : Méthodologies appliquées à l'étude du mode d'action des lncARNs.....	83
<b>Figure 25</b> : Mécanismes moléculaires d'action des lncARNs en fonction de leur localisation subcellulaire.....	84
<b>Figure 26</b> : Modes d'actions des lncARNs nucléaires Xist, THRIL, LEENE et Lockd sur la transcription des gènes codants. ....	87
<b>Figure 27</b> : Modes d'actions des lncARNs CCAT2 et circMbl sur la maturation co-transcriptionnelle des ARNm. ....	90
<b>Figure 28</b> : Rôles structuraux nucléaires de MALAT1 et NEAT1 dans la formation des granules d'épissage (nuclear speckles) et des paraspeckles. ....	92
<b>Figure 29</b> : Mode d'actions de lncARNs interagissant avec d'autres transcrits cytosoliques. ....	94
<b>Figure 30</b> : Régulation de la traduction par les lncARNs. ....	96

<b>Figure 31:</b> Modes d'action des lncARNs SENCER, NORAD, MetaLnc9, Lnc-DC et NKILA sur les modifications post-traductionnelles des protéines.....	98
<b>Figure 32:</b> Interactions des lncARNs et de leurs partenaires protéiques au niveau des granules d'ARNs. ....	100
<b>Figure 33:</b> Trafic intracellulaire et sécrétions des lncARNs dans le compartiment extracellulaire.....	102
<b>Figure 34:</b> Rôle de lncARNs dans le développement et la différenciation cellulaire.....	106
<b>Figure 35:</b> lncARNs impliqués dans les processus cellulaires de la tumorigenèse. ....	117
<b>Figure 36:</b> Les lncARNs dans les cancers broncho-pulmonaires. ....	122
<b>Figure 37:</b> NLUCAT1 favorise la croissance de sphéroïdes tumoraux. ....	159
<b>Figure 38:</b> Recherche des partenaires moléculaires du transcrit NLUCAT1 par CHART-DNAseq.....	162
<b>Figure 39:</b> Recherche d'interactants protéiques de LINC01116 dans les cellules HUVEC. ....	202
<b>Figure 40:</b> Recherche d'interactions potentielles entre la protéine ILF3, LINC01116 et l'ARNm de VEGFA .....	203
<b>Figure 41:</b> NLUCAT1, un lncARN nucléaire protumoral qui régule la réponse anti-oxydante au sein des CBNPC. ....	206
<b>Figure 42:</b> LINC01116, un hypoxa-lnc cytosolique impliqué dans l'adhésion et les jonctions intercellulaires.....	208

### **III. INTRODUCTION**

## **Chapitre 1 : Les cancers broncho-pulmonaires**

### **A. Les poumons**

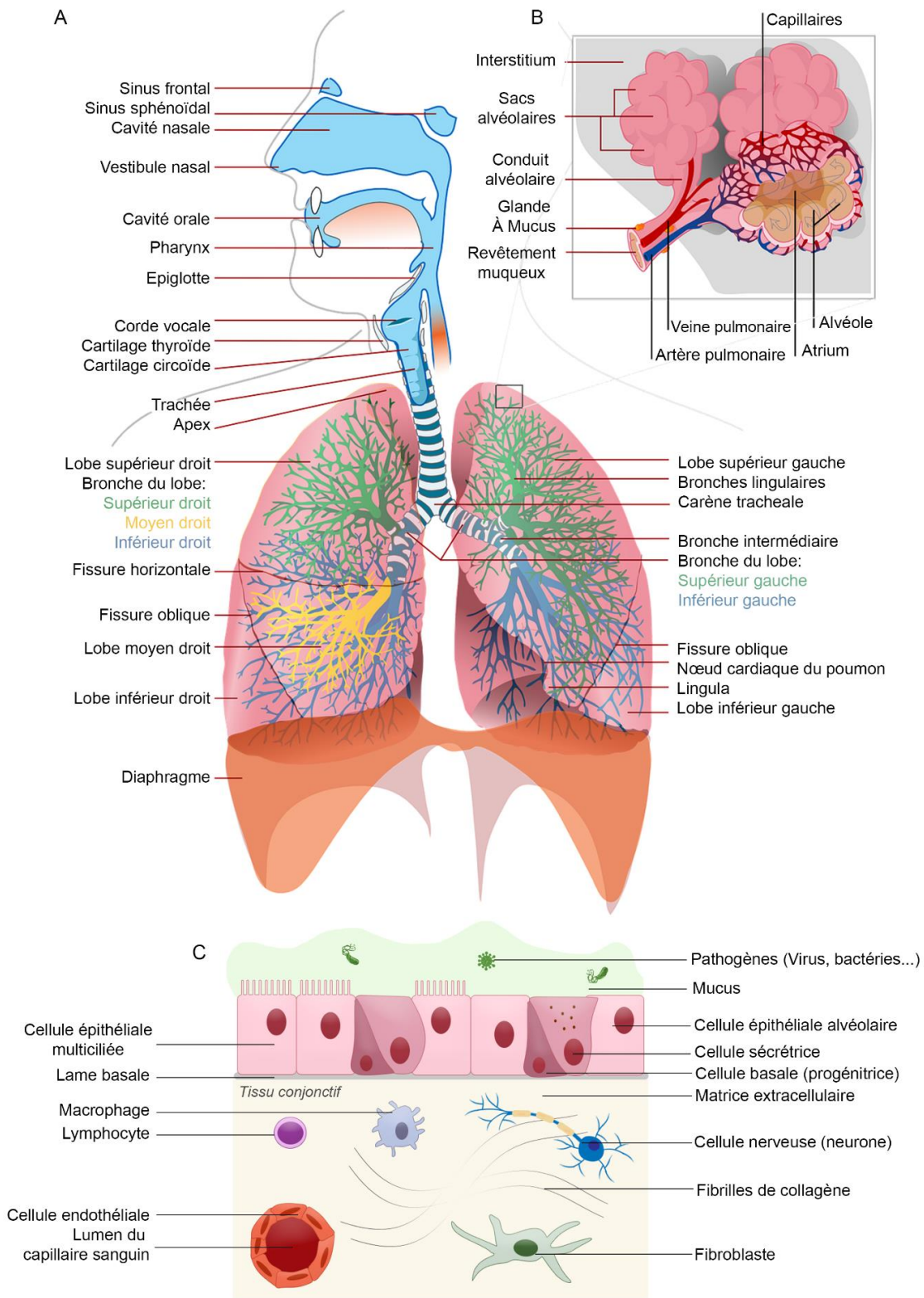
Les poumons sont des organes vitaux au même titre que le cerveau, le cœur, le foie, le pancréas et les reins. Situés dans le thorax de part et d'autre du cœur, ils sont constitués de deux lobes du côté gauche et de trois lobes du côté droit. Ils sont reliés à la trachée par deux bronches souches centrales qui se séparent dans chacun des lobes en bronches puis en bronchioles (Figure 1A). Les bronchioles se terminent par des alvéoles pulmonaires dans lesquelles des capillaires permettent les échanges gazeux d'oxygène (O<sub>2</sub>) et de dioxyde de carbone (CO<sub>2</sub>) entre l'air et le sang (Figure 1B). Le tractus respiratoire est constitué d'un tissu conjonctif tapissé par un épithélium complexe. Ils sont séparés par une lame basale riche en collagène de type IV. L'épithélium pulmonaire est composé de cellules basales progénitrices, de cellules sécrétrices de mucus, de cellules épithéliales alvéolaires (AEC) et de cellules épithéliales multiciliées qui par le mouvement de leurs cils vibratiles déplacent le mucus et permettent la clairance mucociliaire qui évacue les impuretés et les pathogènes. Le tissu conjonctif est majoritairement composé de cellules mésenchymateuse (fibroblastes) qui produisent les constituants de la matrice extracellulaire (MEC), de cellules nerveuses et de cellules endothéliales formant les vaisseaux sanguins et les vaisseaux lymphatiques (Figure 1C) <sup>1,2</sup>.

Il existe de nombreuses pathologies pulmonaires qui peuvent être d'origines infectieuses (tuberculose, coqueluche, grippe, coronavirus), vasculaires (maladie veineuse trombo-embolique, hypertension pulmonaire), chroniques (asthme, broncho-pneumopathie chronique obstructive, fibrose pulmonaire idiopathique) ou cancéreuses.

### **B. Epidémiologie du cancer du poumon**

Les cancers broncho-pulmonaires sont la première cause mondiale de décès par cancer. En France, on dénombre 28 000 décès par an et 45 000 nouveaux cas par an <sup>3</sup>. Le taux de survie à 5 ans n'est que de 17% malgré une prise en charge thérapeutique précoce par exérèse curative associée à différentes chimio-, radio- et immunothérapies. En termes d'incidence, les cancers pulmonaires arrivent en seconde position des cancers les plus fréquents après les cancers de la prostate chez l'homme et ceux du sein chez la femme (Figure 2). Sur la période 2005-2012, la propension à développer des cancers pulmonaires a été stable chez les hommes (-0,3%), alors qu'elle a augmenté de 5,4% chez les femmes en raison de l'accroissement de leur comportement tabagique considéré comme le premier facteur de risque de ces cancers.





**Figure 1** : Représentation schématique des poumons et de l'épithélium respiratoire broncho-pulmonaire. Représentation schématique du tractus respiratoire comportant les voies respiratoires, les poumons et le diaphragme. **B**. Représentation schématique d'une alvéole pulmonaire. **C**. Représentation schématique de l'épithélium respiratoire.



Prostate	240 890	29%	Sein	230 480	30%
<b>Poumon</b>	<b>115 060</b>	<b>14%</b>	<b>Poumon</b>	<b>115 060</b>	<b>14%</b>
Colon & Rectum	71 850	9%	Colon & Rectum	69 360	9%
Vessie	52 020	6%	Utérus	46 470	6%
Mélanome de la peau	40 010	5%	Thyroïde	36 550	5%
Rein	37 120	5%	Lymphome	30 300	4%
Lymphome	36 060	4%	Mélanome de la peau	30 220	4%
Oral et pharynx	27 710	3%	Rein	23 800	3%
Leucémie	25 320	3%	Ovaire	21 990	3%
Pancréas	22 050	3%	Pancréas	21 980	3%
<i>Tous cancers</i>	<i>822 300</i>	<i>100%</i>	<i>Tous cancers</i>	<i>774 370</i>	<i>100%</i>
<b>A. Estimation du nombre de cas par an</b>					
<b>Poumon</b>	<b>85 600</b>	<b>28%</b>	<b>Poumon</b>	<b>71 340</b>	<b>26%</b>
Prostate	33 720	11%	Sein	39 520	15%
Colon & Rectum	25 250	8%	Colon & Rectum	24 130	9%
Pancréas	19 360	6%	Pancréas	18 300	7%
Foie	13 260	4%	Ovaire	15 460	6%
Leucémie	13 740	4%	Lymphome	9 570	4%
Œsophage	11 910	4%	Leucémie	9 040	3%
Vessie	10 670	4%	Utérus	8 120	3%
Lymphome	9 750	3%	Foie	6 330	2%
Rein	8 270	3%	Cerveau	5 670	2%
<i>Tous cancers</i>	<i>300 430</i>	<i>100%</i>	<i>Tous cancers</i>	<i>271 520</i>	<i>100%</i>
<b>B. Estimation du nombre de décès par an</b>					

**Figure 2:** Incidence des cancers dans la population française par genre et tissus.

*Données de l'Organisation Mondiale de la Santé (OMS), 2012.*

## C. Les sous-types histologiques des cancers broncho-pulmonaires

Les cancers broncho-pulmonaires se divisent en différents sous-types ayant des origines tissulaires et des propriétés invasives distinctes.

### 1. Les cancers à petites cellules

Les cancers broncho-pulmonaires à petites cellules (CBPC) représentent 15% des cancers broncho-pulmonaires et sont liés à 95% au tabagisme. Le développement tumoral est proximal au niveau des muqueuses du tronc bronchique et a une origine neuroendocrine. Les CBPC ont une activité mitotique élevée avec un temps de doublement de 30 jours. Ainsi bien qu'ils soient plus sensibles aux drogues cytotoxiques, il s'agit du sous-type le plus agressif. Histologiquement, ces tumeurs se distinguent des autres sous-types par un cytoplasme étroit, une chromatine finement granuleuse et des vésicules neuroendocrines intra-cytoplasmiques.

### 2. Les cancers non à petites cellules

Les cancers broncho-pulmonaires non à petites cellules (CBNPC) sont les plus fréquents (85%) et regroupent les cancers à grandes cellules (10%), les cancers épidermoïdes (40%) et les adénocarcinomes (50%) (Figure 3A).

#### i. Les adénocarcinomes

Les adénocarcinomes constituent 40 à 45% des cancers broncho-pulmonaires et sont, pour 70% d'entre eux, corrélés au tabagisme. Leur dénomination provient de « adeno » qui veut dire « glande » en grec. En effet, la plupart de ces tumeurs sont issues de cellules productrices de mucus et ont majoritairement une origine distale dans le poumon. Il existe 4 sous types histologiques d'adénocarcinomes: les adénocarcinomes acineux qui sont des petits amas de cellules avec des cavités, les adénocarcinomes papillaires qui ont une forme caractéristique en « doigt de gant », les adénocarcinomes solides à sécrétions mucineuses produisant du mucus et sans arrangements particuliers et les adénocarcinomes bronchio-alvéolaires qui respectent l'agencement du tissu pulmonaire mais ont une hypersécrétion de mucus (Figure 3B).

#### ii. Les cancers épidermoïdes

Les carcinomes épidermoïdes ou "Squamous Cell Carcinoma" (SCC) en anglais, sont très corrélés au tabagisme (99% des cas). Ils proviennent généralement d'une métaplasie malpighienne ou d'une lésion de l'épithélium respiratoire. A terme, les cellules cancéreuses vont se différencier à partir de cette lésion et produire de la kératine. Ces tumeurs cancéreuses se développent généralement aux

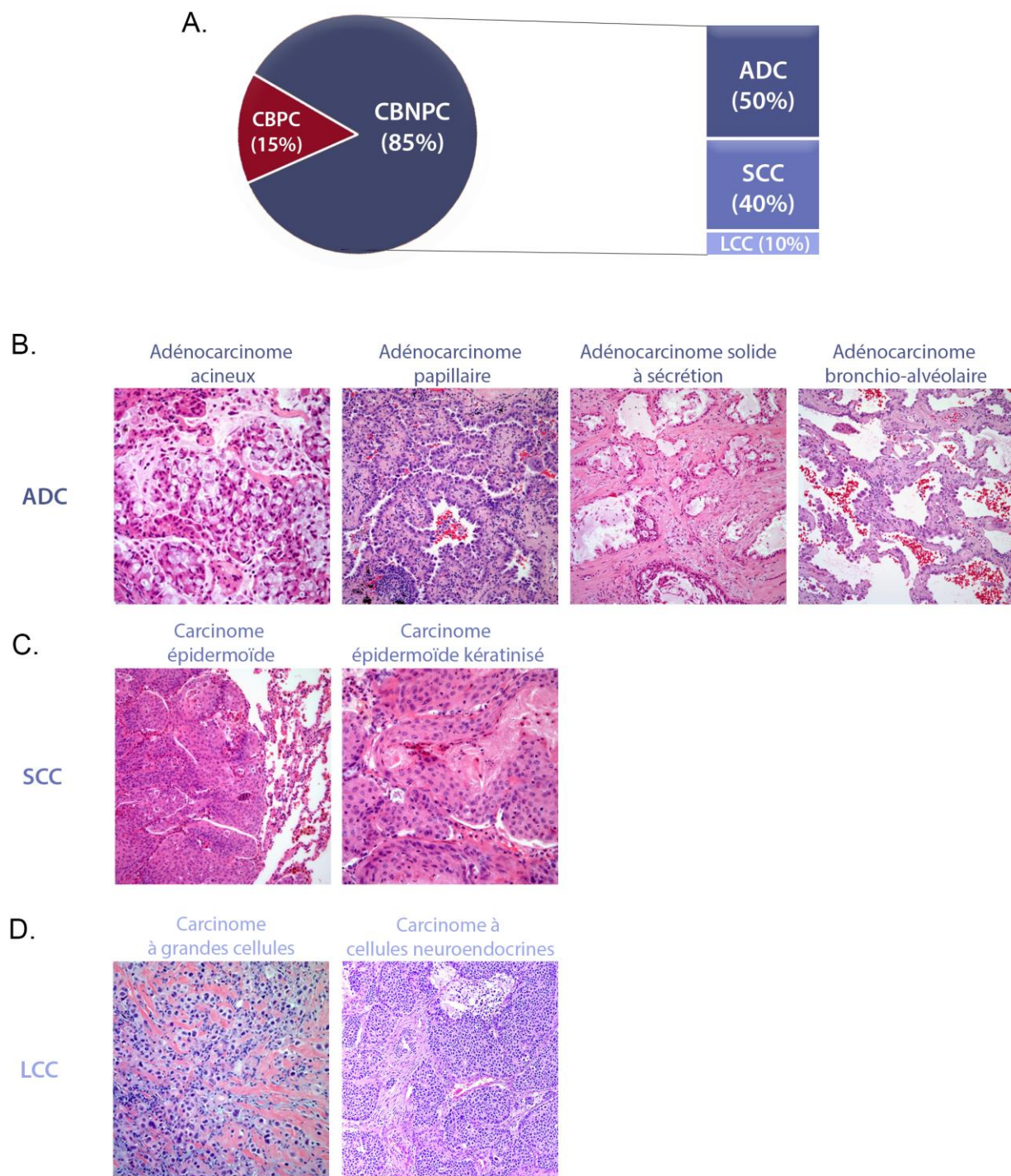
seins des bronches, on parle de bourgeon endobronchique qui peut être excavé par un centre nécrotique (Figure 3C).

### iii. Les cancers à grandes cellules

Les cancers à grandes cellules ou "Large-Cell Carcinoma" (LCC) en anglais sont dûs à 90% au tabagisme. Ces cancers sont également considérés comme « indifférenciés » car ce sont des cellules de grandes tailles ne sécrétant pas de mucus ou de kératine. Elles ont un temps de doublement long, une absence de jonctions cellulaires et une chromatine claire. On distingue dans cette catégorie deux sous-types histologiques : les cancers à grandes cellules et les cancers neuroendocrines à grandes cellules qui ont des vésicules endocrines (Figure 3D).

### iv. Les sous-types histologiques rares

Les tumeurs cancéreuses adéno-squameuses, sarcomatoïdes, muco-épidermoïdes, adénoïdes kystiques et les sarcomes des tissus mous sont des tumeurs rares du tractus pulmonaires qui peuvent s'apparenter à une mixité des sous-types histologiques cités précédemment.



**Figure 3:** Les sous types histologiques des cancers broncho-pulmonaires.

**A.** Les cancers broncho-pulmonaires sont différenciés en deux sous-types majoritaires, les cancers broncho-pulmonaires à petites cellules (CBPC, 15%) et les cancers broncho-pulmonaires non à petites cellules (CBNPC, 85%). Les CBNPC sont divisés en trois sous-types histologiques, les adénocarcinomes (ADC, 50%) (**B**), les cancers épidermoïdes (SCC « Squamous Cell Carcinoma », 40%) (**C**) et les cancers à grandes cellules (LCC « Large-Cell Carcinoma », 10%) (**D**). Coupes de biopsies colorées à l'hématoxyline-éosine pour visualiser les différences histologiques. (Source photographies <sup>4</sup>)

### 3. Stades anatomo-cliniques

Les cancers broncho-pulmonaires peuvent être classés selon la classification internationale TNM qui permet d'évaluer le stade d'avancement des tumeurs. Cette classification prend en considération l'état de la tumeur primitive (T), l'état d'invasion des ganglions lymphatiques régionaux (N pour node en anglais) et la présence de métastases distantes (M).

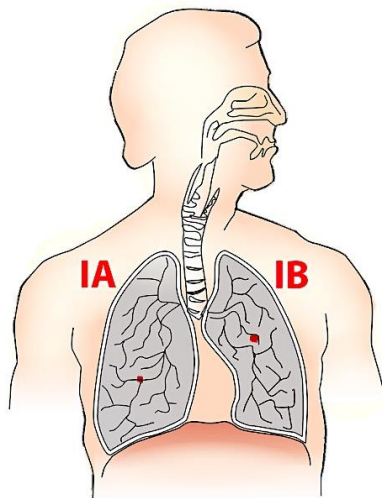
Ainsi, T0 désigne une tumeur initiale non localisée, T1 une tumeur inférieure à 3 cm de diamètre et entourée par du tissu pulmonaire, T2 une tumeur de plus de 3 cm ou qui envahit les bronches principales, T3 une tumeur qui envahit la plèvre ou la paroi thoracique et T4 une tumeur qui envahit le cœur, la trachée, l'œsophage et/ou les vertèbres. Concernant les adénopathies régionales, N0 caractérise l'absence d'invasion ganglionnaire, N1 la présence de ganglions métastatiques péri-bronchiques homolatéraux, N2 la présence de ganglions métastatiques médiastinaux et N3 la présence de ganglions métastatiques contralatéraux, sous-carènes ou sous claviculaires. La présence de métastases est classifiée en fonction de l'absence (M0) ou de la présence de métastases viscérales (M1).

On distingue ainsi les différents stades d'avancement de la maladie (Figure 4) :

- Stade IA: tumeur inférieure à 3 cm sans métastases viscérales ou dans les ganglions périphériques « T1N0M0 ».
- Stade IB: tumeur supérieure à 3 cm sans métastases viscérales ou dans les ganglions périphériques « T2N0M0 ».
- Stade IIA: tumeur inférieure à 5 cm avec des ganglions métastatiques péri-bronchiques « T1N1M0 ».
- Stade IIB: tumeur supérieure à 5 cm ou qui infiltre les bronches supérieures et qui peut présenter des ganglions métastatiques péri-bronchiques « T2N1M0 » ou « T3N0M0 »
- Stade IIIA : tumeur envahissant les bronches principales avec des ganglions métastatiques péri-bronchiques « T3N1M0 » ou « TnN2M0 » existence de ganglions métastatiques médiastinaux.
- Stade IIIB : tumeur envahissant le cœur, la paroi thoracique ou la plèvre « T4NnM0 » ou « TnN3M0 » présence de ganglions métastatiques médiastinaux contralatéraux
- Stade IV : Présence de métastases dans d'autres tissus « TnNnM1 ».

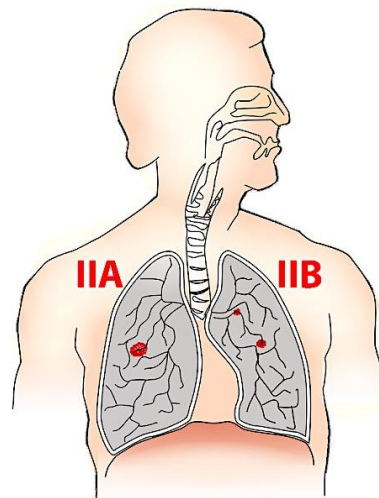


## STADE I



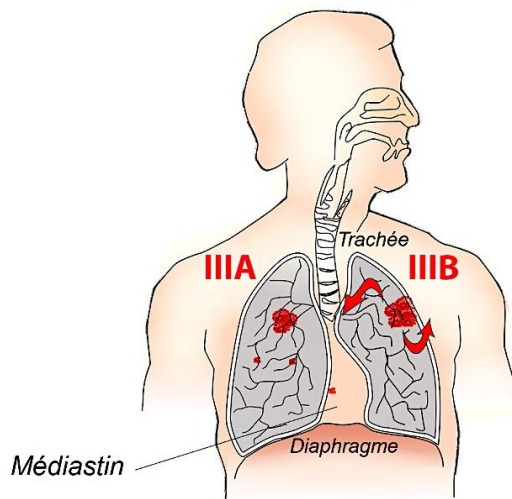
Pas de métastases (M0)  
Pas de nodules (N0)  
**Stade IA:** Tumeur <3cm  
(T1a: <2cm; T1b: 2<>3cm)  
**Stade IB:** Tumeur T2a: 3<>5cm

## STADE II



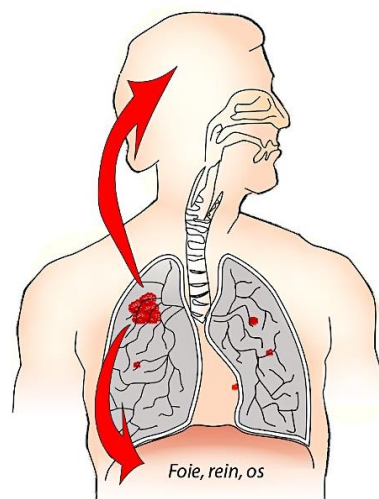
Pas de métastases (M0)  
**Stade IIA:** Tumeur >5cm  
(T2b: 5<>7cm; T3:>7cm)  
et sans nodules (N0)  
**Stade IIB:** Tumeur <5cm  
avec nodule homolatéral (N1)

## STADE III



Pas de métastases (M0)  
**Stade IIIA:** Tumeur envahissant les bronches (T3)  
et présence de nodules péri-bronchiques (N1)  
ou tumeur (T1a-T2b) avec nodules médiastinaux  
homolatéraux (N2)  
**Stade IIIB:** Tumeur envahissant les tissus adjacents (T4)  
(Coeur, plèvre, diaphragme, trachée...) ou présence  
de nodules médiastinaux controlatéraux (N3).

## STADE IV



Présence de métastases  
dans d'autres tissus (M1)  
(Foie, rein, cerveau, os)

**Figure 4:** Classification internationale TNM des cancers pulmonaires en fonction de leur extension anatomique

## D. Les facteurs de risques

Les facteurs de risques correspondent à l'ensemble des éléments favorisant le développement du cancer. Dans le cadre des cancers broncho-pulmonaires, on prend en compte différentes molécules cancérigènes présentes notamment dans le tabac ou dans l'environnement, des facteurs génétiques et des infections aiguës ou chroniques.

### 1. Classification des molécules cancérigènes

Les substances cancérigènes sont référencées par le Centre International de Recherche sur le Cancer (CIRC) de l'Organisation Mondiale de la Santé (OMS) basé à Lyon en France afin d'identifier la dangerosité de ces molécules. Plus de 1000 molécules ont été référencées depuis 1971, et 500 d'entre elles sont annotées comme cancérigènes ou probablement cancérigènes selon la classification suivante:

- Éléments de classe I : agents cancérigènes avérés pour l'homme (n=120).
- Éléments de classe 2A : agents probablement cancérigènes pour lesquels des informations suffisantes de susceptibilité existent chez l'animal mais pas chez l'homme (n=81).
- Éléments de classe 2B : agents potentiellement cancérigènes pour lesquels des informations obtenues chez l'animal sont insuffisantes mais présomptives (n=299).
- Éléments de classe 3 : agents inclassables quant à leur cancérigénicité (n=502).
- Élément de classe 4: agents probablement pas cancérigènes (n=1, le caprolactame).

On considère que le tabac et notamment les cigarettes conventionnelles sont le premier facteur de risque des cancers broncho-pulmonaires. En effet, la cigarette contient 70 éléments chimiques ayant une action carcinogène démontrée chez l'homme ou l'animal dont 16 molécules qui sont des éléments de classe I. Parmi ces substances de classe I, on distingue les polycycliques aromatiques hydrocarbonés (PAH) et les N-nitrosamines qui sont émis lors de la combustion du tabac.



## 2. Le tabagisme

En 1964, le rapport accablant de *The Surgeon General of United States Public Health Service* indique que le tabac est le premier facteur de risque du cancer du poumon avec une augmentation de 81% de l'incidence de ces cancers <sup>5</sup>. En 1995, ce même service a démontré que les non-fumeurs exposés au tabagisme passif étaient soumis qualitativement aux mêmes substances cancérigènes ce qui augmenterait leur risque de développer un cancer broncho-pulmonaire de 24% <sup>6</sup>. Les 70 substances cancérigènes du tabac ont plusieurs actions sur la cancérogénèse. Certaines vont activer des voies de signalisations métaboliques induisant la prolifération anarchique des cellules tandis que d'autres molécules vont créer des modifications sur l'ADN (C/G>T) et induire entre 1000 et 10 000 mutations oncogéniques par cellule (environ 9 mutations/mégabases) <sup>7,8</sup>.

Aujourd'hui, l'utilisation de la cigarette électronique (e-cigarette) est en plein essor dans le cadre du sevrage tabagique. Cependant, aucune étude sur les risques de son utilisation à long terme n'a été effectuée. Toutefois, les premiers travaux montrent que les quantités de N-nitrosamines inhalées dans la vapeur sont négligeables. Une solubilisation de formaldéhyde, un agent fixateur cancérigène, a été mise en évidence mais à des niveaux proches de ceux utilisés dans la pharmacologie actuelle <sup>9</sup>. Toutefois, la chronicité d'inhalation et la puissance électrique des e-cigarettes constituent un danger pour la santé par la création d'un stress oxydant chronique génotoxique au niveau du tractus respiratoire sollicitant les systèmes de réparation de l'ADN pouvant, à long terme, entraîner des mutations délétères <sup>10</sup>.

## 3. Les expositions environnementales et professionnelles

Les expositions environnementales et professionnelles constituent le deuxième facteur de risque. Par ordre d'incidence, elles sont liées à l'exposition au radon, aux particules fines, à l'amiante, aux pesticides et aux radiations ionisantes.

### i. Le radon

Le radium a été mis en évidence par Pierre et Marie Curie en 1899. L'année suivante, le physicien Ernst Frederick Dorn a découvert le radon, un gaz inerte incolore, inodore et radioactif d'origine naturelle provenant de la transformation du radium. Dans les années 1950, après la forte incidence des cancers broncho-pulmonaires chez les mineurs, le radon a été répertorié comme un facteur de risque des cancers broncho-pulmonaires qui agit en synergie avec la consommation de tabac <sup>11</sup>. Au contact de l'air et de l'eau, le radon est dégradé en polonium qui émet des rayonnements alpha causant des dommages à l'ADN et des mutations susceptibles d'entraîner la formation de tumeurs cancéreuses <sup>12</sup>.

## ii. Les particules fines

Les particules fines émises par les moteurs diesel ne sont pas clairement répertoriées comme facteurs de risque à cause de la grande variabilité d'exposition existant au sein des cohortes étudiées. Toutefois, une étude récente imputerait 3% des cancers broncho-pulmonaires survenus en France en 2015, à l'exposition aux particules fines <sup>13</sup>.

## iii. L'amiante

L'amiante provient de fibres minérales silicatées et est grandement utilisée dans l'industrie. Elle augmente de 2 à 6 fois le risque de développer un cancer du poumon. Les effets de l'amiante sur le cancer agissent en synergie avec le tabagisme et touchent principalement les ouvriers de l'industrie. En France, 2,3% des cancers broncho-pulmonaires chez l'homme seraient liés à l'exposition à l'amiante <sup>14</sup>.

## iv. Les pesticides

Les pesticides font également parti des facteurs environnementaux impactant l'incidence des cancers broncho-pulmonaires. Le CIRC a classifié près d'une soixantaine de pesticides incluant l'arsenic (carcinogène de classe I), le captafol et le dibromure d'éthylène (classe 2A) et 19 autres molécules (classe 2B). Aujourd'hui, les études épidémiologiques relient la proximité du lieu d'habitation de champs traités aux pesticides au risque de développer différents cancers (seins, prostate, poumon et cerveau) <sup>15</sup>.

## v. Les radiations ionisantes

Les derniers facteurs environnementaux ayant un impact sur le risque de contracter un cancer du poumon sont les radiations ionisantes. En effet, l'étude épidémiologique des survivants des bombes atomiques d'Hiroshima et de Nagasaki a révélé un effet dose dépendant de l'exposition aux radiations ionisantes sur l'incidence des cancers broncho-pulmonaires. On retrouve cette même corrélation chez les patients régulièrement exposés aux examens tomographiques et rayons X <sup>16</sup>.

## 4. Les facteurs génétiques

Les différentes études de ségrégation des gènes dans les cancers broncho-pulmonaires ne permettent pas de lier directement une mutation ou un polymorphisme à cette pathologie. Toutefois, certaines mutations peuvent aggraver l'impact de l'environnement sur le développement tumoral pulmonaire. Ainsi, certaines familles ont une apparition plus précoce de la maladie au vu de leur consommation tabagique avec une augmentation de 50% de l'incidence des cancers broncho-pulmonaires <sup>17,18</sup>. Plus spécifiquement, la délétion allélique 6q23-25 semble présente dans environ 50%

des tumeurs broncho-pulmonaires dans 4 cohortes différentes <sup>19,20</sup>. D'autres liens ont été démontré entre l'incidence des cancers broncho-pulmonaires et les mutations dans les loci chromosomiques 12q24, 5q31-q33, 14q11, 16q24 <sup>21,22</sup>

#### 5. Les pathologies pulmonaires chroniques

La bronchopneumopathie obstructive chronique (BPCO) est la pathologie chronique pulmonaire la plus répandue avec 1,7 millions de personnes concernées en France en 2015. La BPCO est due à un contexte inflammatoire important entraînant une sécrétion anormale de mucus, un épaissement des parois musculaires et une destruction des alvéoles pulmonaires appelée emphysème. Les symptômes sont une toux persistante, une difficulté à respirer (dyspnée) et des expectorations fréquentes. Elle augmente de 2 à 3 fois le risque de développer un cancer broncho-pulmonaire même lorsqu'elle n'est pas associée à une consommation tabagique (20% des cas). D'autres pathologies chroniques, créant un contexte inflammatoire pulmonaire comme l'asthme ou la tuberculose, augmentent également de 2,4 et 3,5 fois l'incidence d'apparition des cancers broncho-pulmonaires <sup>23</sup>.

#### 6. Les agents infectieux

Les derniers facteurs de risques connus des cancers broncho-pulmonaires sont les infections virales. De nombreuses études ont démontré les liens entre les infections aiguës de pneumonie et l'augmentation de l'incidence des cancers broncho-pulmonaires. Les pneumonies sont liées à des infections virales diverses créant un contexte inflammatoire pulmonaire qui favorise le développement de cancers broncho-pulmonaires <sup>24</sup>.

Par leur statut immunodéprimé, les patients infectés par le virus de l'immunodéficience humaine (VIH) sont plus sensibles aux infections pulmonaires conduisant à un contexte inflammatoire chronique favorable au développement tumoral <sup>25</sup>. Enfin, les papillomavirus sont des virus oncogéniques infectant les muqueuses dont l'ADN viral a été mis en évidence dans le tractus respiratoire <sup>26</sup>. Ces virus produisent une protéine E6 capable de dégrader la protéine suppresseur de tumeur TP53, gardienne de l'intégrité du génome et de la prolifération cellulaire maligne dans les poumons.

## Chapitre 2 : Les principales caractéristiques des cancers broncho-pulmonaires

Les tumeurs ont des caractéristiques particulières (ou « Hallmarks » en anglais) qui leur confèrent leur agressivité (Figure 5) <sup>27,28</sup>. Parmi ces caractéristiques on distingue : i) les altérations génétiques et épigénétiques des cellules cancéreuses modifiant leur métabolisme et leurs capacités prolifératives et/ou invasives, et ii) les dérégulations d'expression génique affectant le microenvironnement tumoral qui participent au développement de la pathologie.

Parmi les propriétés intrinsèques des cellules cancéreuses, on note un pouvoir de multiplication élevé, l'inhibition des points de contrôle du cycle cellulaire, l'immortalité due à la réplication des télomères, la résistance à la mort cellulaire, la modification du métabolisme et l'acquisition de propriétés invasives conduisant à la formation de métastases. Ces modifications sont dues à l'instabilité du génome et au fort taux de mutation au sein de ces cellules entraînant la dérégulation de l'expression de gènes codants ou non codants contrôlant ces processus cellulaires <sup>29</sup>.

D'autres caractéristiques des cancers sont liées au microenvironnement tumoral et impliquent les cellules stromales environnantes et la matrice extracellulaire. Ces caractéristiques participent à l'invasion des métastases dans la matrice extracellulaire, à la formation de nouveaux vaisseaux sanguins (angiogenèse), à l'échappement des cellules cancéreuses à l'immunosurveillance ainsi qu'à la formation d'un contexte inflammatoire.

### A. Classification moléculaire des cancers broncho-pulmonaires

#### 1. Hétérogénéité moléculaire des cancers broncho-pulmonaires

Une des particularités des CBNPC concerne leur important statut mutationnel induit par le tabagisme qui entraîne des réparations et des réarrangements aberrants de l'ADN <sup>7,8,30</sup>. Au cours de la progression tumorale, les nombreux réarrangements chromosomiques et mutations s'accumulent au sein des cellules créant une grande hétérogénéité moléculaire mise en évidence par les techniques actuelles de séquençage d'ARN à l'échelle de la cellule unique <sup>31-33</sup>.

Cette instabilité génomique se traduit par des réarrangements de segments chromosomiques, une modification du nombre de copies d'un gène (appelée "Mutant allele specific imbalance" ou "MASI" en anglais) ou des mutations somatiques de perte ou de gain de fonction <sup>34</sup>. Ces modifications génomiques peuvent déréguler l'expression des gènes codants ou non codants soit directement par des mutations/réarrangements au sein des gènes, soit indirectement en entraînant des modifications épigénétiques <sup>35</sup>. Les mutations et réarrangements conférant aux cellules leurs caractéristiques cancéreuses sont appelés des mutations conductrices ("drivers") tandis que les mutations qui n'impactent pas les cellules sont appelées les mutations passagères ("passengers") <sup>36</sup>.



Figure 5: Les caractéristiques des cancers. Adapté de <sup>27</sup>

## 2. Les mutations conductrices majoritaires dans les CBNPC

Les mutations conductrices peuvent cibler des oncogènes afin d'induire leur activation constitutive ou des gènes suppresseurs de tumeurs afin d'invalider leur fonction. Ces mutations peuvent survenir spontanément ou être induites par des molécules cancérigènes ou d'autres facteurs de risques décrits précédemment. Différentes mutations et réarrangements chromosomiques peuvent co-exister au sein des tumeurs et cibler un même gène ou différents gènes pro-tumoraux. Les oncogènes sont impliqués dans différentes voies de signalisations incluant les récepteurs à activité tyrosine kinase, la signalisation mTOR, la régulation du stress oxydatif, le contrôle du cycle cellulaire ou encore la réparation de l'ADN <sup>37</sup>. Les altérations géniques les plus représentées dans les cancers broncho-pulmonaires concernent les gènes *EGFR*, *KRAS* et *TP53*.

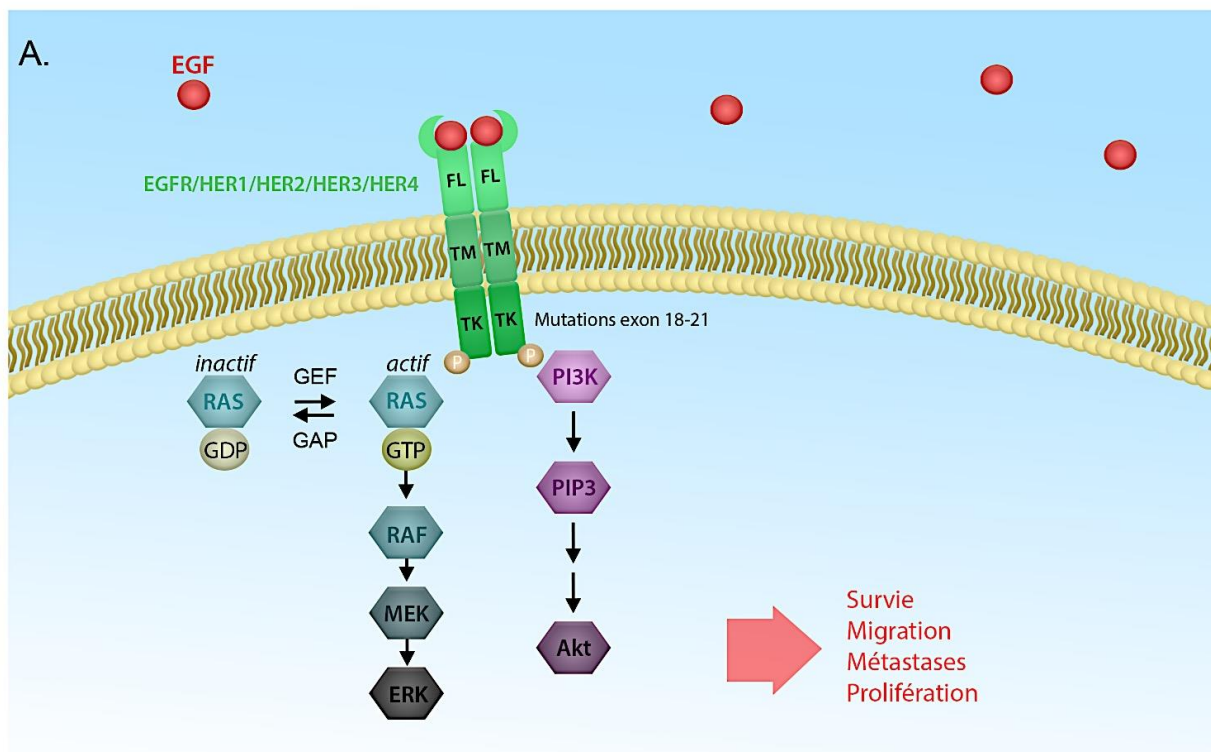
### *i. Les Mutations du récepteur de l'EGF*

Les récepteurs de l'EGF (Epidermal Growth Factor) appartiennent à la famille des récepteurs à activité tyrosine kinase et regroupent quatre membres : EGFR (HER1, ErbB1), HER2 (ErbB2), HER3 (ErbB3) et HER4 (ErbB4). L'EGF se fixe sur un récepteur dimérique (homo/hétérodimère) qui s'active par l'autophosphorylation de résidus de tyrosine dans les domaines cytosoliques. En aval, deux cascades de signalisation peuvent être activées par phosphorylation, les voies PI3K/AKT/mTOR et RAS/RAF/MEK/MAPK qui contrôlent divers processus impliqués dans le développement tumoral comme l'angiogenèse, la prolifération, la survie, la migration et l'invasion cellulaires (Figure 6A).

A l'inverse, l'inhibition de ce récepteur entraîne la surexpression de molécules pro-apoptotiques comme la protéine Bim, activant la voie apoptotique mitochondriale et conduisant à la mort des cellules cancéreuses <sup>38</sup>.

Les mutations de l'EGFR induisent une activation constitutive de l'activité tyrosine kinase du récepteur dans les cellules cancéreuses favorisant la croissance tumorale et la formation de métastases. L'incidence de ces mutations varie en fonction de l'ethnie. En effet, 30% des cancers broncho-pulmonaires portent ce type de mutation en asie contre 15% en Europe <sup>39</sup>. Les trois mutations majoritaires de l'EGFR (85% des cas) concernent le site catalytique du récepteur. Il s'agit d'une délétion dans l'exon 19, et de substitutions dans l'exon 20 (T790M) et dans l'exon 21 (L858R). D'autres mutations existent au niveau du domaine transmembranaire telle que la délétion des exons 2-7 mimant la fixation du ligand <sup>40</sup> (Figure 6B).

Les tumeurs présentant des mutations EGFR auront un phénotype plus agressif ainsi qu'une probabilité accrue de résistance aux inhibiteurs de tyrosine kinases utilisés en chimiothérapie.



**B.**

	(...) 688	728 729	761 762	823 824	875	Séquence protéique				
	Exon 18				Exon 19	Exon 20	Exon 21	Séquence génique		
Mutations (%) <sup>*</sup>	G719A (0,77) G719S (0,47) G719C (0,26) Autres (0,91)				Toutes les délétions (46) E746_A750 (39,4) L747_P753>S (1,4) L747-A750>P (0,81) L747-T751 (0,96) E746_S752>V (0,49) L747_S752 (0,44) E746_751 (0,29)		T790M (4,1) S768I (0,55) Toutes les insertions (1,45) V769-D770insASV (0,14) D770-N771insSVD (0,11)		L858R (37,5) L861Q (1,12)	
					Toutes les insertions (0,2)					

**Figure 6:** Voies de signalisation de l'EGFR et mutations impliquées dans les cancers broncho-pulmonaires. Représentation des voies de signalisation MAPK et PI3K/AKT activées par le récepteur de l'EGF et régulant divers processus cellulaires. (FL : domaine de fixation du ligand, TM : domaine transmembranaire, TK : domaine à activité tyrosine kinase) **B.** Cartographie des différentes mutations du gène du récepteur de l'EGF répertoriées dans les cancers broncho-pulmonaires. Concernant la nomenclature présentée ici, les appellations « G719A » représentent les cas de mutations ponctuelles et notamment l'exemple d'une mutation ponctuelle d'une glycine en alanine en position 719. Les appellations « E746\_A750 » représentent les délétions des acides aminés, avec en exemple la délétion du glutamate en position 746 à l'alanine en position 750 et les appellations « V769-D770insASV » représentent une insertion d'acides aminés avec en exemple l'insertion des acides aminés alanine-serine-valine entre la valine en position 769 et la glutamine en position 770. Le pourcentage de ces mutations reporté par le COSMIC (Catalogue Of Somatic Mutations In Cancer) en 2014 est indiqué entre parenthèses. \* La mutation T790M est souvent présente avec d'autres mutations primaires. Par conséquent, ce résultat peut être sous représentatif par rapport à des études sur des mutations multiples. Adapté de <sup>41</sup>.

## *ii. Les Mutations de KRAS*

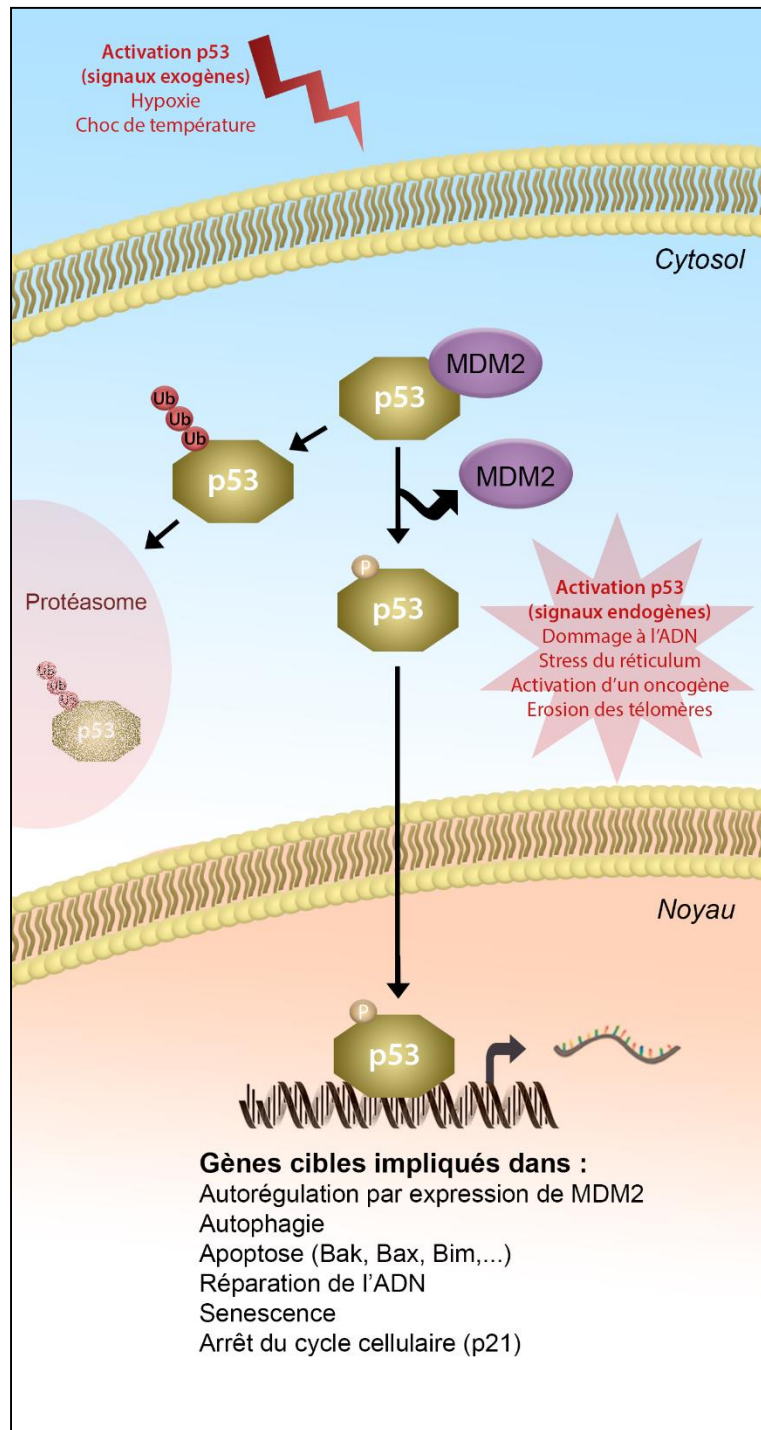
La protéine KRAS pour « Kirsten Rat Sarcoma » appartient à la famille des petites GTPases encore appelées petites protéines G. Elles permettent la transduction des signaux des récepteurs aux facteurs de croissance tels que l'EGFR aux voies PI3K/AKT/mTOR et RAF1/MEK/MAPK. A l'état inactif, elles se situent au niveau membranaire et sont liées avec une molécule de faible énergie, le GDP (Guanine Di-Phosphate). Elles sont activées par un facteur d'échange GEF (Guanine nucleotide Exchange Factor), qui permet de remplacer une molécule de GDP par une molécule de GTP (Guanine Tri-Phosphate) de plus haute énergie <sup>42</sup>. La protéine KRAS-GTP est ensuite hydrolysée par la protéine GAP (GTPase Activating Protein) qui active les effecteurs permettant la transduction du signal <sup>43</sup>. Les protéines RAS ont également une activité GTPase intrinsèque leur permettant d'activer les effecteurs.

Certaines mutations de KRAS entraînent une activation constitutive de l'activité GTPase et une augmentation de la prolifération, de la survie et de la migration cellulaire. La fréquence de ces mutations varie également en fonction des groupes ethniques et est plus fréquente chez les patients présentant un adénocarcinome (32%, Figure 8) et ayant un comportement tabagique <sup>44,45</sup>. Ainsi contrairement aux mutations affectant l'EGFR, 40% des CBPNC ont une mutation KRAS dans la population européenne contre seulement 10% dans la population asiatique. Les mutations sont généralement des substitutions au niveau des acides aminés G12, G13 et Q61. La mutation G12C est la plus fréquente (40%). Dans le cas des mutations qui touchent le gène KRAS, la mutation d'un allèle est suffisante pour augmenter le métabolisme et la prolifération des cellules cancéreuses. Les mutations de *KRAS* ont un effet dose-dépendant c'est à dire que leur impact est plus fort si les deux allèles du gène sont altérés ou si des réarrangements chromosomiques entraînent un gain de copie du gène. D'ailleurs l'activation de cette voie de signalisation par réarrangement chromosomique est plus fréquente que l'apparition d'une mutation somatique par substitution <sup>46,47</sup>.

## *iii. Les Mutations de TP53*

La protéine p53 est un facteur de transcription codé par le gène *TP53*. Elle a été découverte suite à son induction dans des cellules infectées par le virus oncogénique SV40 (simian virus 40) <sup>48</sup>. La protéine p53 induit l'arrêt du cycle cellulaire et l'apoptose en réponse à un stress oxydatif, à un stress génotoxique ou à une infection virale oncogénique. Dans un contexte physiologique, elle est faiblement exprimée car elle est dégradée par l'ubiquitine ligase E3 MDM2, elle-même induite par p53 dans un mécanisme d'autorégulation (Figure 7) <sup>49</sup>. La protéine p53 est un suppresseur de tumeurs et est appelée « gardienne du génome ».





**Figure 7:** Rôles suppresseur de tumeur et oncogénique de p53.

La protéine p53 (TP53) est induite en réponse à divers types de stress (génotoxique, hypoxie, thermique, activation d'oncogène, érosion des télomères, stress au réticulum), et entraîne l'arrêt de la prolifération et l'apoptose des cellules concernées afin d'empêcher l'apparition de cellules cancéreuses. Lors de mutations de la protéine p53, il y a une inhibition de l'apoptose et des réparations de l'ADN.

En effet, lors d'un stress à l'ADN, p53 est phosphorylée sur des résidus sérine présents dans le domaine de transactivation N-terminal ce qui induit son import nucléaire, sa tétramérisation et sa liaison au niveau de séquences ADN spécifiques pour activer la transcription de gènes cibles tels que la protéine p21 qui inhibe le cycle cellulaire ou la protéine pro-apoptotique Bax<sup>50</sup>.

Les mutations impactant la voie p53 peuvent cibler directement le gène *TP53* bloquant la synthèse de la protéine ou provoquant sa perte de fonction ou son incapacité à se lier à l'ADN. Elles peuvent également impacter les partenaires de p53 comme l'ubiquitine ligase MDM2 (8%, Figure 8<sup>37</sup>), les séquences génomiques de fixation du facteur de transcription p53 (p53RE, p53 Response Element) ou les facteurs de transcription p63, p73, SP1, SREBPs, le récepteur de la vitamine D ou NRF2 qui contrôle la réponse anti-oxydante<sup>51</sup>. Il s'agit des principales mutations dans les cancers à mauvais pronostic tels que les cancers à petites cellules, les cancers épidermoïdes, les cancers du sein triple-négatif, les cancers ovariens de haut-grade et les cancers de l'œsophage. Plus précisément, l'analyse du statut mutagène de cohortes de patients du TCGA (The Cancer Genome Atlas) indique que ces mutations touchent 46% des patients présentant un adénocarcinome pulmonaire et 80% des patients présentant un carcinome épidermoïde ou à petites cellules. Ces mutations sont souvent combinées à des mutations de l'EGFR ou de KRAS<sup>37,47,52</sup>.

### 3. Les autres mutations conductrices

Comme dit précédemment, différentes mutations et réarrangements chromosomiques sont concomittants au sein des cellules cancéreuses conférant une grande hétérogénéité moléculaire au sein des tumeurs. Cette hétérogénéité n'est pas toujours synonyme de mauvais pronostic puisque les conséquences de ces mutations peuvent être ciblées par différents traitements chimiothérapeutiques.

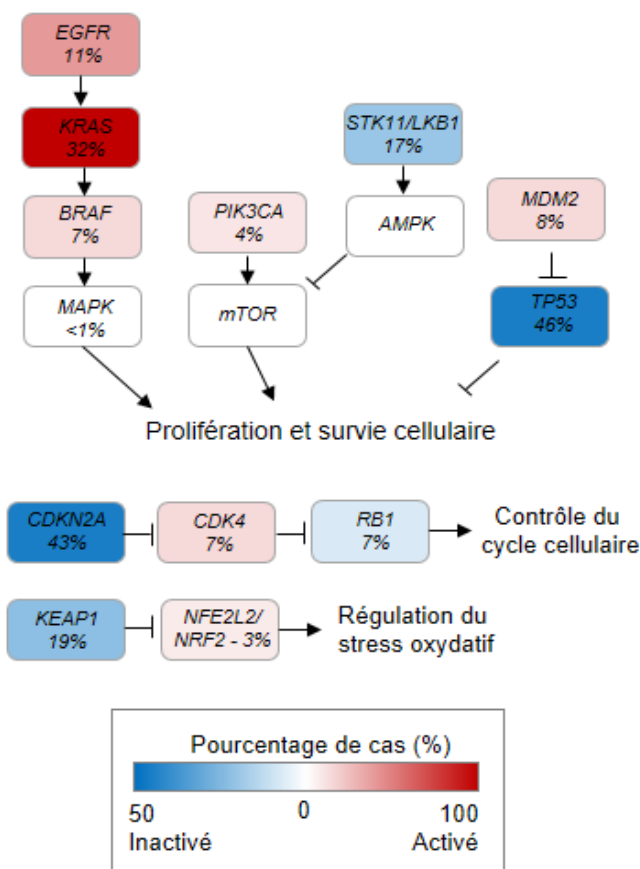
Parmi les autres gènes "drivers" fréquemment mutés dans les cancers<sup>53,54</sup>, certains affectent la prolifération des cellules cancéreuses. C'est le cas du gène *CDKN2A* qui contrôle le cycle cellulaire. Ce gène est fréquemment inhibé dans les cancers (43%, Figure 8) soit par hyperméthylation de son promoteur sur les deux allèles du gène soit par MASI. L'inactivation de *CDKN2A* conduit à un arrêt du contrôle du cycle cellulaire et à une croissance anarchique des cellules<sup>37</sup>.

Le gène *BRAF* (b-Raf proto-oncogène sérine/thréonine kinase) est impliqué dans la cascade de signalisation MAPK du récepteur de l'EGF. Dans les CBNPC, la mutation somatique gain de fonction la plus fréquemment identifiée concerne la substitution V600E (7%, Figure 8) qui induit une activation des voies MAPK<sup>37,55</sup>. De même, la fusion du gène *ALK* ("Anaplastic Lymphoma Receptor Tyrosine Kinase") avec le gène de son partenaire *EML4* ("echinoderm microtubule-associated protein like 4") est un événement apparaissant dans 1% des cas de CBNPC et induisant une activation constitutive des voies

MAPK (Figure 8) <sup>54,56</sup>. Les mutations somatiques impactant le gène PI3KCA (Phosphatidylinositol 3-kinase) induisent son expression dans 4% des CBNPC et activent ainsi la voie de signalisation mTOR également impliquée dans la prolifération et la survie cellulaire <sup>54,57</sup>.

D'autre part, le gène LKB1/STK11 (Sérine/Thréonine kinase 11) est ciblé par des mutations somatiques inhibant son expression dans les CBNPC invasifs (17% des cas, Figure 8). LKB1 a une action suppresseur de tumeurs en activant la voie p53, le gène CDKN2A et la voie AMPK (AMP activated protein kinase) qui inhibe la voie mTOR <sup>58,59</sup>.

La voie NRF2/NFE2L2 qui permet de diminuer le stress oxydatif des cellules cancéreuses est également la cible de mutations conductrices. Le stress oxydatif induit la production de radicaux libres oxygénés (RLO) qui induisent la mort apoptotique des cellules cancéreuses. Ainsi, dans les cancers, la voie de signalisation NRF2 est constitutivement activée par des mutations induisant la perte de fonction du gène *KEAP1*, inhibiteur de NRF2 (19%) <sup>57,60</sup>. Dans ce contexte, la recherche de nouveaux gènes porteurs de mutations "drivers" est au centre de nombreuses études visant à identifier de nouveaux marqueurs pronostics et de nouvelles cibles thérapeutiques anticancéreuses.



**Figure 8** : Fréquences des mutations oncogéniques dans les adénocarcinomes pulmonaires. Adapté de <sup>57</sup>

## B. Le microenvironnement tumoral

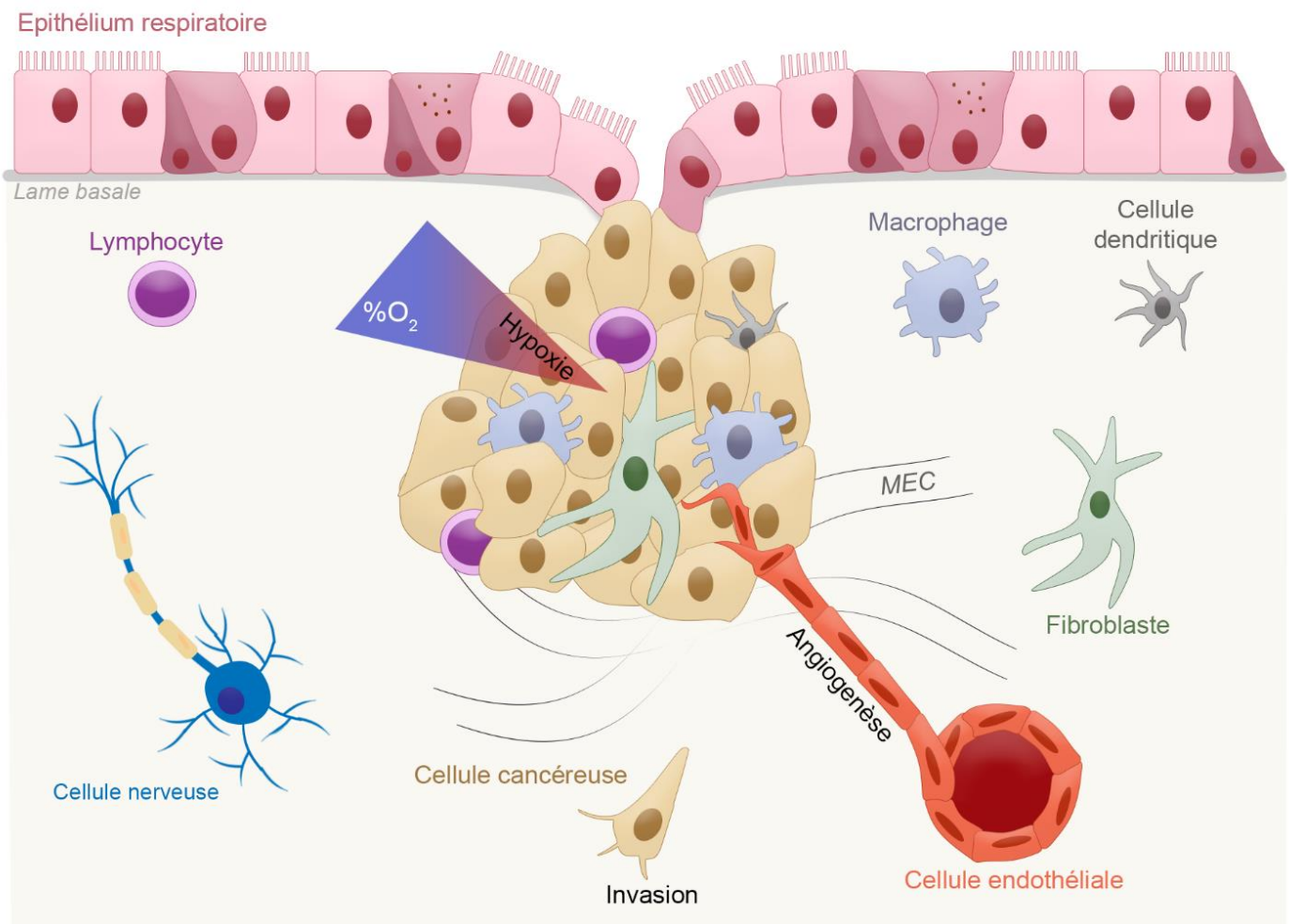
Le microenvironnement tumoral est constitué par la matrice extracellulaire (MEC) et par l'ensemble des cellules non tumorales appelées cellules stromales à proximité des cellules cancéreuses. Ces cellules regroupent les cellules immunitaires infiltrant la tumeur, les cellules constituant les vaisseaux sanguins et lymphatiques (cellules endothéliales, péricytes...), les cellules mésenchymateuses du tissu conjonctif (fibroblastes), les cellules nerveuses et les cellules épithéliales du tissu sain péri-tumoral (Figure 9).

Ce stroma tumoral et des paramètres physico-chimiques comme le pH, la rigidité ("stiffness" en anglais) matricielle ou le taux d'oxygène, influencent l'expression génique et les capacités prolifératives et invasives des cellules cancéreuses qui doivent s'adapter à cet environnement <sup>61</sup>. Le microenvironnement tumoral est donc un acteur dans la mise en place des caractéristiques des cancers <sup>62</sup>.

Parmi les cellules stromales, trois types cellulaires majoritaires vont moduler le potentiel agressif des tumeurs. Ce sont : i) les cellules immunitaires par leur sécrétion de nombreux facteurs de croissance et cytokines ; ii) les cellules endothéliales par leur capacité angiogénique (création de nouveaux vaisseaux sanguins) et ii) les fibroblastes par leur sécrétion de constituants de la matrice extracellulaire modifiant la rigidité du microenvironnement tumoral.

### 1. Les cellules immunitaires

Le système immunitaire se subdivise en systèmes immunitaires inné et acquis. Le système immunitaire inné est composé de cellules « Natural Killer » (NK), de cellules phagocytaires telles que les polynucléaires neutrophiles, les macrophages (de type M1 et M2) et les cellules dendritiques (DC), et de granulocytes (éosinophiles, basophiles, neutrophiles). Parmi ces cellules, les DC et les macrophages sont résidents dans le tissu pulmonaire afin d'assurer l'immunosurveillance et d'éliminer par phagocytose les cellules anormales ou infectées par un pathogène. Ces cellules sont également appelées cellules présentatrices d'antigènes (CPAg) car une fois le pathogène digéré, les CPAg présenteront à leur surface des peptides antigéniques afin d'activer le système immunitaire acquis composé des lymphocytes B, producteurs d'anticorps, et des lymphocytes T. Les CPAg se lient avec les lymphocytes T par le biais de différents récepteurs membranaires créant une synapse immunologique générant des signaux d'activation et d'expansion clonale des lymphocytes T. Cette réponse immunitaire prévient l'apparition de nombreuses pathologies dont les cancers <sup>63</sup>.



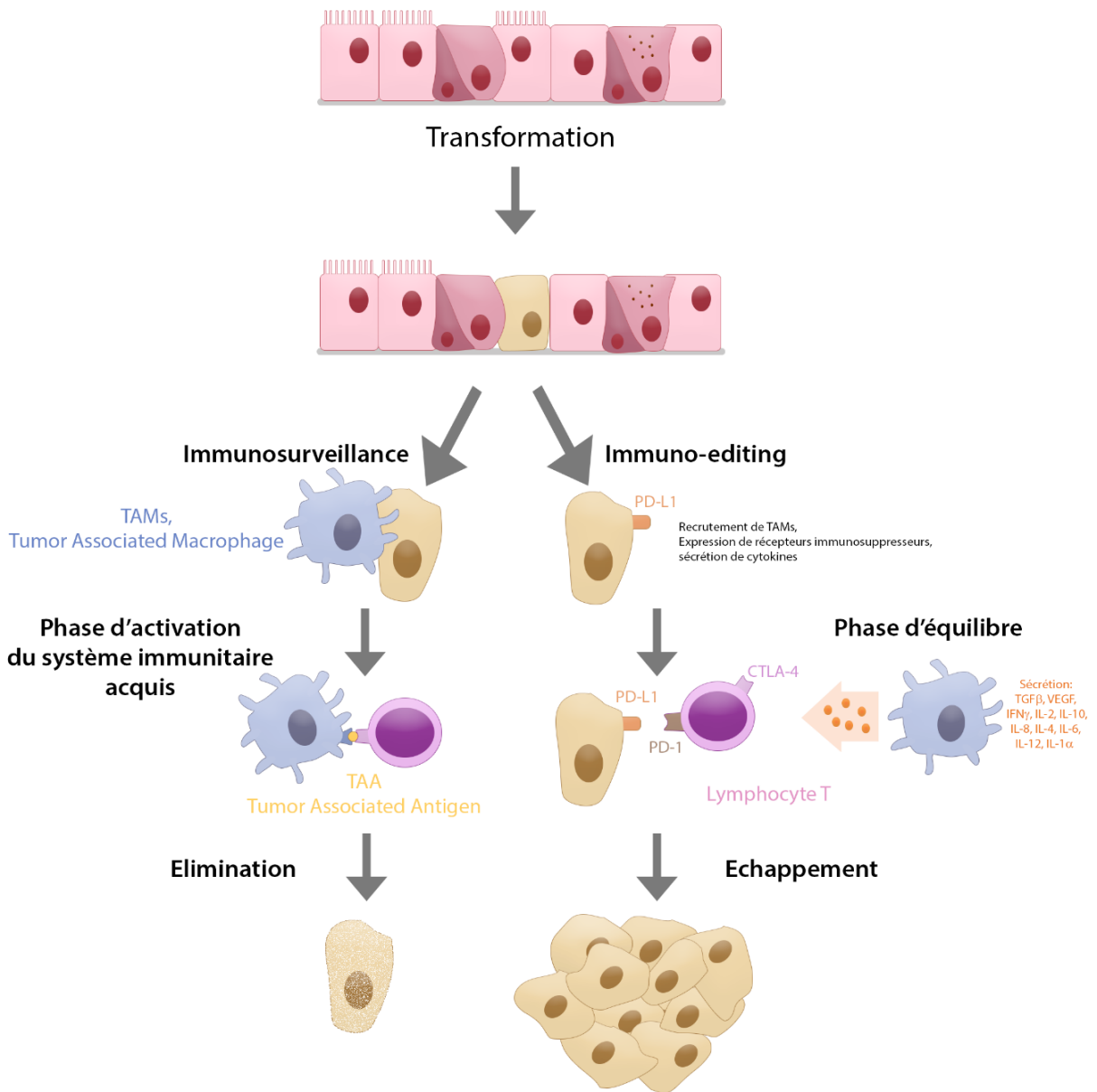
**Figure 9:** Organisation et diversité du microenvironnement tumoral.

Les cellules cancéreuses formant la tumeur primaire désorganisent la structure du tissu créant une niche tumorale possédant son propre microenvironnement. Au sein du microenvironnement tumoral se trouvent les cellules tumorales, la matrice extracellulaire et les cellules stromales présentes dans le tissu sain qui infiltrent ce microenvironnement telles que les cellules endothéliales, les cellules immunitaires et les fibroblastes. Cette niche tumorale peut modifier les apports en oxygène et en nutriments créant un contexte hypoxique. Adapté de <sup>64</sup>.

Toutefois, il y a 150 ans, le Dr Virchow a émis le postulat que les cellules immunitaires peuvent également favoriser la tumorigenèse car on observe un infiltrat de cellules immunitaires au sein de nombreuses tumeurs solides. De plus, comme nous l'avons vu précédemment, les pathologies inflammatoires chroniques entraînent des altérations et un remodelage permanent de l'épithélium respiratoire favorable à l'émergence de mutations et à la transformation tumorale (I-C-6) <sup>24,63,65-67</sup>.

Ainsi, bien que les cellules cancéreuses soient la plupart du temps éliminées par le système immunitaire, certaines peuvent échapper à cette surveillance par des mécanismes « d'Immuno-editing » (Figure 10). Les cellules cancéreuses expriment à leur surface des récepteurs tels que PD-L1 (Programmed Death Ligand 1) qui inhibent la réponse immunitaire au niveau de la synapse immunologique <sup>65</sup>. Ces récepteurs sont ciblés par les approches d'immunothérapies anticancéreuses.

Cette phase dite « d'équilibre » est également caractérisée par le recrutement de cellules immunitaires via la sécrétion de chimiokines comme l'IL-8 ou le MCP1/CCL2 (Monocyte Chemoattractant Protein 1). Les macrophages infiltrants les tumeurs (TAMs, Tumor-Associated Macrophages) dérivent des monocytes et sécrètent au cœur de la tumeur des facteurs de croissance tels que le TGF $\beta$  (Transforming Growth Factor  $\beta$ ), le VEGF (Vascular Endothelial Growth Factor), l'interféron (IFN $\gamma$ ), des cytokines anti-tumorales (IL-1 $\alpha$ , IL-6, IL-12) et pro-tumorales (IL-10, IL-4) et des protéases matricielles (Metalloprotéases, MMP1-9) <sup>66,67</sup>. L'ensemble de ces cytokines et facteurs de croissance produits par les cellules immunitaires et par d'autres cellules stromales vont induire un contexte inflammatoire au sein des tumeurs et promouvoir l'échappement c'est à dire la prolifération puis l'invasion des cellules cancéreuses et l'angiogenèse des cellules endothéliales au sein de la tumeur solide.



**Figure 10:** La surveillance des cancers broncho-pulmonaires par le système immunitaire.

La majorité des cellules épithéliales se transformant en cellules cancéreuses sont rapidement éliminées par le système immunitaire. Toutefois, dans certains cas, la cellule cancéreuse échappe à l'immunosurveillance et atteint un équilibre avec l'infiltration du système immunitaire favorisant sa prolifération. Cet état permet l'apparition d'une dernière phase dite d'échappement qui utilise le système immunitaire à des fins pro-tumorales.

## 2. Les cellules endothéliales

Les cellules endothéliales constituent les vaisseaux sanguins et lymphatiques <sup>68,69</sup>. Les vaisseaux sanguins permettent de distribuer l'oxygène, les nutriments et les hormones aux cellules des tissus via la composante artérielle mais également de collecter le dioxyde de carbone et de recycler 90% de la fraction liquide du sang via la composante veineuse du système circulatoire. Le système lymphatique a pour fonction de réguler l'homeostasie des fluides en captant les 10% de la fraction liquide extravasée dans les tissus et de transporter les cellules immunitaires <sup>70,71</sup>.

Les vaisseaux sanguins sont composés de cellules endothéliales et de péricytes qui maintiennent la contractibilité des vaisseaux. Les cellules endothéliales des vaisseaux sanguins ont des jonctions intercellulaires complexes (desmosomes, jonctions adhérentes, serrées et communicantes) et expriment des molécules spécifiques (VEGFR-1, IL8, IL6, MCP-1 et NRP-1) permettant le recrutement tissulaire de cellules immunitaires provenant de la circulation systémique.

Les vaisseaux lymphatiques sont composés de cellules endothéliales lymphatiques (« LEC » pour Lymphatic Endothelial Cell en anglais), de cellules de muscles lisses et d'un système interne de valves afin d'éviter les reflux de liquide. Les cellules endothéliales lymphatiques (LEC) expriment d'autres marqueurs moléculaires spécifiques (VEGFR-2, VEGFR-3, NRP-2, Prox-1 et LYVE-1). Elles ont des jonctions intercellulaires plus perméantes qui favorisent la recaptation des fluides <sup>71,72</sup>.

La néo-vascularisation tumorale par recrutement de cellules endothéliales permet d'approvisionner la tumeur en nutriments et en oxygène via la création de nouveaux vaisseaux sanguins (angiogenèse), tout en favorisant la dissémination des métastases par création de nouveaux vaisseaux lymphatiques (lymphangiogenèse) <sup>73-75</sup>. En effet, les vaisseaux lymphatiques infiltrant la tumeur permettent aux cellules métastatiques d'afférées jusqu'aux ganglions sentinelles du tissu. Ces métastases migrent ensuite dans la circulation systémique pour atteindre d'autres organes tels que le foie, le cerveau et d'autres zones du poumon <sup>75-77</sup>. La néo-vascularisation peut impliquer trois mécanismes différents, i) le recrutement de précurseurs des cellules endothéliales provenant de la moelle osseuse, les angioblastes, ii) l'expansion de la vascularisation préexistante, ou iii) la contrainte du tissu tumoral qui crée des invaginations au sein du réseau vasculaire (intussusception) <sup>78</sup>. Ces différentes cellules endothéliales sont recrutées par divers stimuli tels que les stress hypoxique et/ou métabolique, mécanique (rigidité de la MEC), l'inflammation ou encore des mutations de gènes pro- ou anti-angiogéniques <sup>78</sup>. Ces mécanismes sont liés à la sécrétion de molécules favorisant la prolifération, la migration et les interactions cellule-cellule, telles que le VEGF-A, le VEGF-C, l'angiopoïétine 2 (Ang2) et le bFGF (basic Fibroblast Growth Factor) <sup>79,80</sup>. Ces facteurs de croissance sont produits par les cellules immunitaires, les fibroblastes, les cellules tumorales et les cellules endothéliales du



microenvironnement tumoral. D'autres molécules pro-angiogéniques sont produites spécifiquement dans le contexte tumoral, telles que l'intégrine  $\alpha 5\beta 3$ , l'inhibiteur de l'activateur du plasminogène PAI-1, le monoxyde d'azote (NO), la cyclo-oxygénase-2 (COX-2) responsable de la synthèse de prostaglandine PGE2 et la thrombospondine-2 (TSP-2) <sup>69,78</sup>. La voie Notch est également activée au cours de la néo-vascularisation, par l'interaction du récepteur Notch1 avec différents ligands membranaires tels que Jagged 1 ou les Delta-Like Canonical Notch 1 Ligand 1-4 (DLL1-4) <sup>81</sup>. Cette interaction induit la coupure du domaine intracellulaire NICD (Notch Intracellular domain) et sa translocation dans le noyau où il active les gènes cibles Hey1, Hey2, Myc, p21 et d'autres gènes impliqués dans la migration, la différenciation et la prolifération des cellules endothéliales <sup>81-84</sup>.

L'angiogenèse et la lymphangiogenèse sont donc directement liées à la dissémination des métastases et sont corrélées à un mauvais pronostic vital des patients <sup>85</sup>. Aujourd'hui, la recherche de nouvelles stratégies thérapeutiques se base sur le ciblage de ces processus cellulaires. Dans ce but, des études cliniques ont été initiées pour déterminer l'impact de l'utilisation d'anticorps ciblant les LECs au niveau de leurs récepteurs (VEGFR-3 ou VEGFR-2) dans le traitement des cancers <sup>86(p1),87</sup>. Certains traitements ont été validés par les agences de santé publique malgré des effets secondaires multiples (hypertension, nausées, diarrhées, fatigue et vomissements) et un effet anti-tumoral minime. Parmi ces traitements, le ramucirumab cible uniquement le récepteur VEGFR-2 et le bevacizumab cible concomitamment les récepteurs VEGFR-2 et VEGFR-3 <sup>88,89</sup>. Actuellement d'autres traitements ciblant d'autres molécules telles que le VEGF-C (VGX-100 utilisé en combinaison avec le bevacizumab), le VEGF-A et l'angiopoïétine 2 (vanucizumab) sont en essais cliniques de phases I ou II <sup>90,91</sup>. Ainsi, la compréhension des mécanismes de la lymphangiogenèse est au cœur de la recherche de nouvelles cibles thérapeutiques contre le cancer.

### 3. Les fibroblastes, architectes de la matrice extracellulaire (MEC)

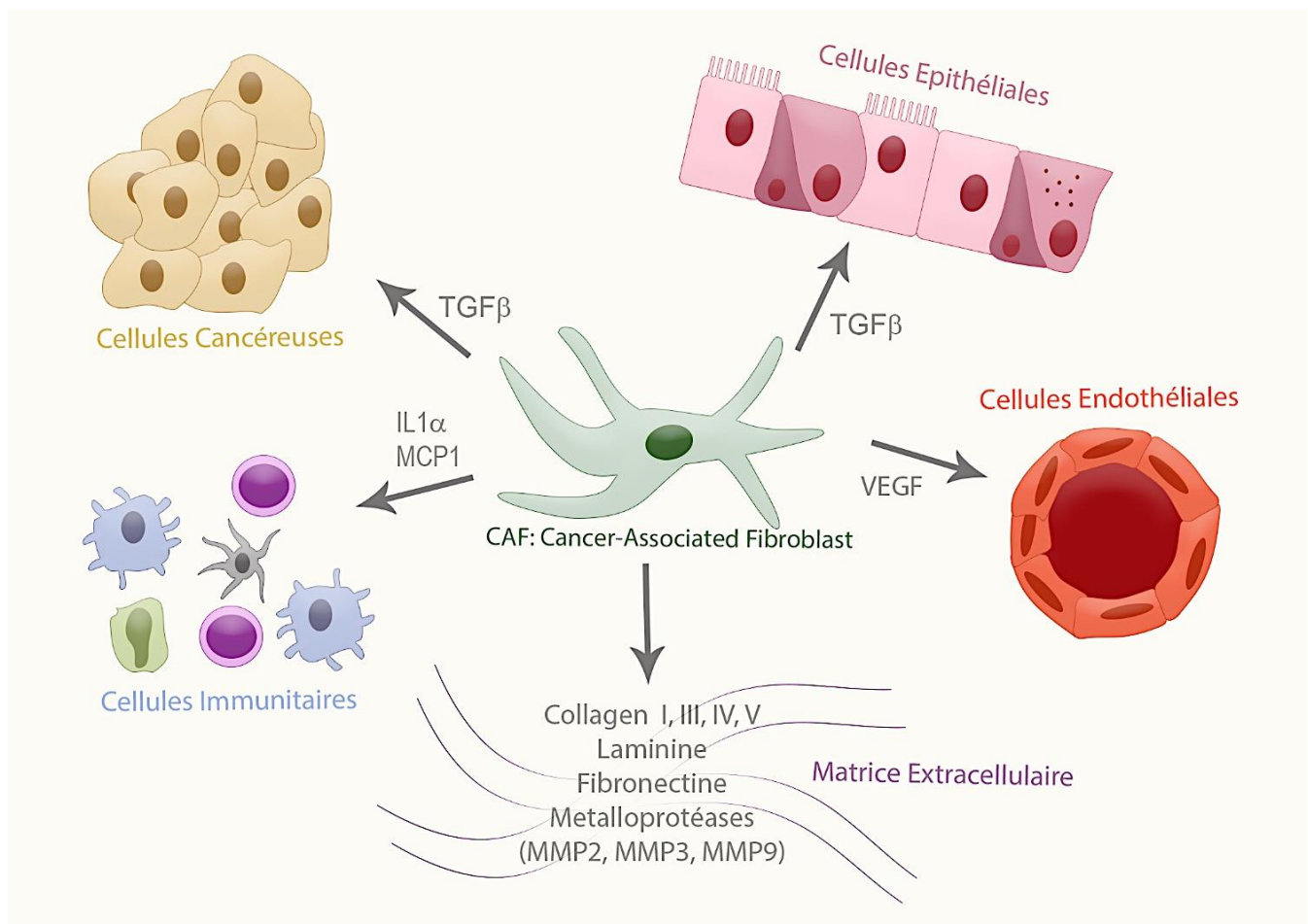
Les fibroblastes sont la composante cellulaire majoritaire du tissu conjonctif séparé de l'épithélium par la lame basale. Ils ont pour fonction de produire les composants de la MEC (Fibronectine, laminine, collagènes fibrillaires de types I, III, V et protéases matricielles) et de réguler les mécanismes de réparation et de différenciation de l'épithélium <sup>92</sup>. Une activation anormale des fibroblastes ou une modification de leur profil de sécrétion peut favoriser l'apparition de pathologies telles que la fibrose pulmonaire ou les cancers <sup>92-95</sup>.

Les fibroblastes peuvent être activés par différents constituants de la matrice extracellulaire dont différents facteurs de croissance (EGF, PDGF, FGF et TGF $\beta$ ), les intégrines ( $\alpha 6\beta 1$ ,  $\alpha \nu\beta 5$ ) mais également par des interactions intercellulaires au niveau des récepteurs ICAM1 (Intercellular adhesion molecule 1) ou VCAM1 (Vascular cell adhesion molecule 1) ou encore par le stress oxydatif via les radicaux libres oxygénés (RLO) <sup>96</sup>. On nomme les fibroblastes activés péri-tumoraux ou infiltrant la tumeur, les fibroblastes associés au cancer ou "Cancer-Associated Fibroblasts" (CAF) en anglais (Figure 11). Les CAFs ont différentes origines, ils proviennent des fibroblastes résidents, des cellules mésenchymateuses intra-tumorales, des adipocytes ou encore des cellules endothéliales <sup>92,97,98</sup>. Les dernières expériences de séquençage à l'échelle de la cellule unique ont mis en évidence les différents marqueurs d'expressions des CAFs en fonction de leurs origines histologiques <sup>99</sup>. Les CAFs expriment soit des marqueurs communs à toutes les sous-populations tels que l' $\alpha$ -SMA ( $\alpha$ -smooth muscle actin), TGF $\beta$ , SPARC (Secreted Protein Acidic and Cysteine Rich) et TAGLN (Transgelin), soit des marqueurs spécifiques de leurs origines. Par exemple, les CAFs ayant une origine vasculaire expriment les protéines EpCAM, PECAM1 et CD34 tandis que les CAFs provenant des fibroblastes résidents expriment différentes protéines de la matrice (COL14a1, Fibulin2, Lox).

Ils stimulent la progression tumorale en participant à chaque étape de la tumorigenèse depuis l'initiation jusqu'au développement métastatiques au sein d'organes secondaires. Les CAFs accroissent les capacités invasives des cellules tumorales et développent la résistance aux traitements des cancers <sup>100,101</sup>. En effet, ils vont sécréter de grandes quantités de métalloprotéases (MMP2, MMP3 et MMP9) ce qui diminue la rigidité de la MEC permettant la création d'un chemin migratoire <sup>102</sup>. Ils guident également les cellules tumorales qui migrent collectivement en suivant une cellule dite « leader » qui peut être une cellule cancéreuse ou un CAF sécrétant du TGF $\beta$  <sup>103</sup>. Les CAFs sécrètent également du VEGF permettant l'attraction des cellules endothéliales <sup>92,94</sup>.

Enfin, ils sécrètent également de nombreux facteurs de croissance, cytokines (IL-1) et chimiokines (MCP1, Monocyte chemotactic protein 1) permettant l'attraction des cellules immunitaires et la création d'un contexte immunosuppresseur <sup>104</sup>. Ainsi, depuis quelques années, les CAFs sont les

cibles de différentes thérapies anti-cancéreuses. Différentes études ont démontrés que la déplétion des CAFs à l'aide d'inhibiteurs pharmacologiques ou d'anticorps réduit la taille des tumeurs de cancer du sein, du poumon, du colon et du pancréas <sup>105,106</sup>. Toutefois, la déplétion des CAFs par un inhibiteur pharmacologique ciblant la protéine  $\alpha$ -SMA peut également induire un phénotype plus agressif avec une tumeur plus invasive et une infiltration de lymphocytes immunosupresseurs (Tregs) au sein de la tumeur <sup>107</sup>. L'ambivalence de ces résultats est liée notamment à l'hétérogénéité des sous-populations de CAFs et à leurs profils moléculaires qui font actuellement l'objet de nombreuses recherches <sup>104</sup>.



**Figure 11:** L'activation des fibroblastes associés au cancer (CAF) favorise l'agressivité tumorale. En sécrétant différents facteurs de croissance, cytokines, chimiokines et protéases, les CAFs promeuvent l'invasion des cellules cancéreuses, la transition épithélio-mésenchymateuse de l'épithélium respiratoire, l'infiltration de cellules inflammatoires et l'angiogenèse des cellules endothéliales. Adapté de <sup>92</sup>.

#### 4. Les échanges gazeux et l'hypoxie

Dans les tumeurs solides, la prolifération anarchique des cellules tumorales crée des zones éloignées des vaisseaux sanguins caractérisées par un apport réduit en oxygène créant un environnement hypoxique. L'hypoxie est un facteur important de l'agressivité des tumeurs. En effet, les cellules tumorales vont s'adapter à leur environnement appauvri en oxygène en activant différentes voies de signalisation qui va notamment modifier le métabolisme des cellules, leurs capacités invasives et les rendre plus résistantes à différents traitements anticancéreux <sup>108</sup>.

##### *i. Voies de signalisation activées par l'hypoxie*

Les facteurs de transcription HIFs (Hypoxia Inducible Factors) sont les régulateurs clés de la réponse hypoxique (Figure 12). Ce sont des hétérodimères, constitués de sous-unités alpha, HIF-1 $\alpha$  ou HIF-2 $\alpha$  (EPAS1), associées à une sous-unité bêta, HIF-1 $\beta$  (ANRT), qui après formation dans le cytoplasme transloquent dans le noyau et se fixent sur des séquences régulatrices HRE (Hypoxia Response Element) pour activer l'expression de gènes cibles.

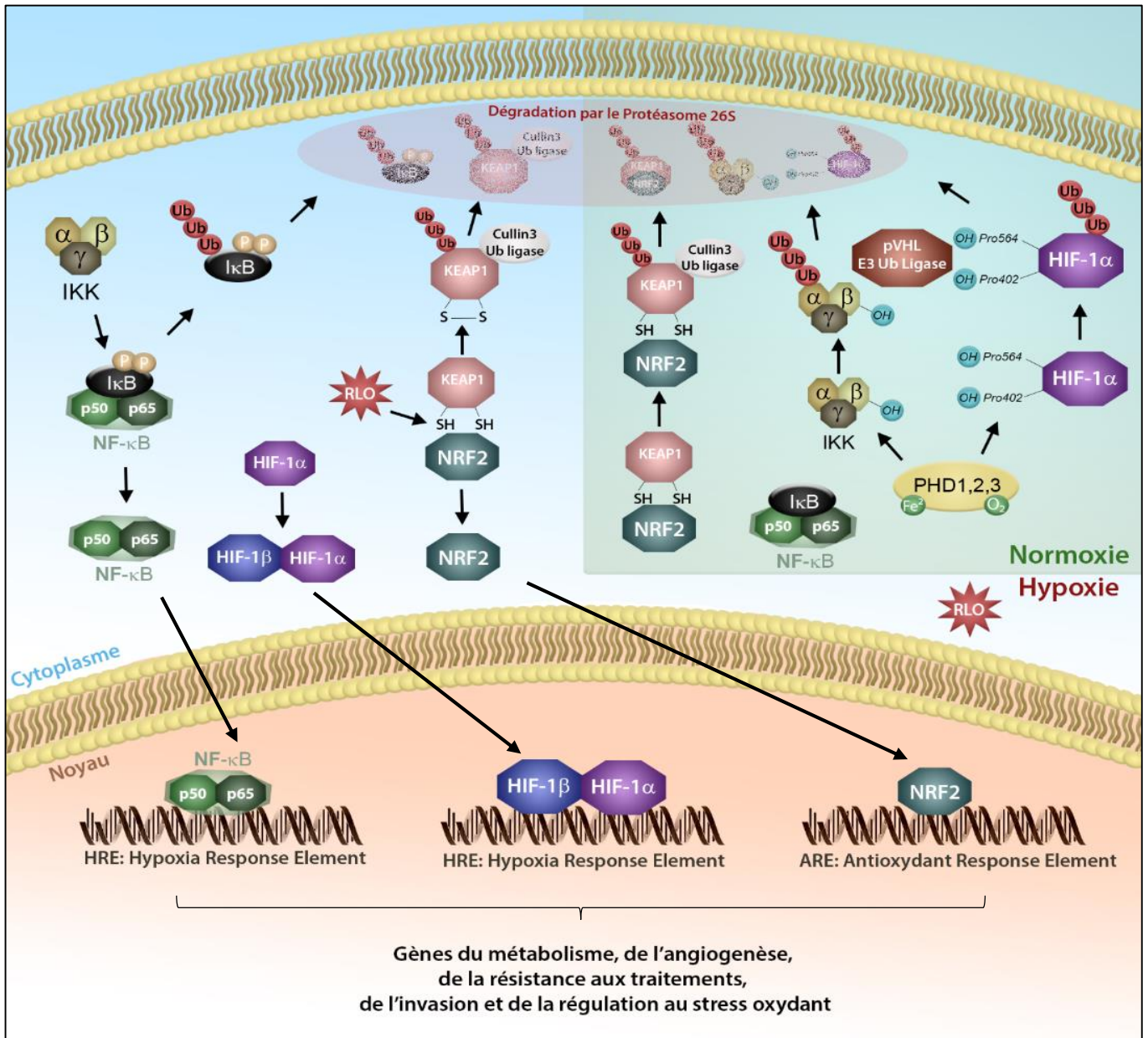
En présence d'oxygène, c'est-à-dire en condition normoxique, les protéines HIFs sont très instables. En effet, des prolyl hydroxylases (PHDs), senseurs de la concentration en oxygène, vont hydroxyler un résidu proline présent dans la séquence protéique de HIF-1 $\alpha$  (Proline 564) ou de HIF-2 $\alpha$  modifiant leur conformation et entraînant leur ubiquitination par l'ubiquitine ligase pVHL (Von Hippel-Lindau) puis leur dégradation par le protéasome <sup>109,110</sup>. En condition hypoxique, les PHDs sont inactives et les protéines HIF-1 $\alpha$  ou HIF-2 $\alpha$  sont stabilisées.

Les isoformes HIF-1 $\alpha$  et HIF-2 $\alpha$  se différencient au niveau de leurs domaines N-terminaux, ce qui leur permet de cibler des séquences HRE différentes et donc de réguler l'expression de gènes spécifiques. Ainsi, le gène *CA9* (Carbonic Anhydrase 9) est induit par HIF-1 $\alpha$  tandis que le facteur de transcription c-Myc est régulé par HIF-2 $\alpha$ . Cependant, ces isoformes peuvent avoir une action synergique sur certains gènes comme pour l'induction du gène *VEGF*. Enfin, ces facteurs peuvent avoir des effets antagonistes comme pour le gène de l'IL-8, cytokine pro-angiogénique, qui est réprimé par HIF-1 $\alpha$  et induit par HIF-2 $\alpha$  <sup>110</sup>.

D'un point de vue cinétique, bien que stabilisés et transloqués concomitamment dans le noyau, ces facteurs de transcription n'interviennent pas au même moment dans la réponse hypoxique. Le facteur HIF1 $\alpha$  intervient dans la réponse précoce et agit à l'hypoxie, puis il est ciblé par l'ubiquitine ligase E3 HAF (Hypoxia Associated Factor) qui induit sa dégradation. En revanche HIF-2 $\alpha$  n'est pas ubiquitiné par HAF et son activation se maintient dans le temps ce qui lui permet de réguler les effets d'une hypoxie chronique <sup>111</sup>.

D'autres facteurs de transcription sont activés dans un contexte hypoxique. C'est le cas de la voie NF- $\kappa$ B classiquement liée à l'intégration de stimuli inflammatoires (tels que le TNF $\alpha$ ) qui induisent la phosphorylation de l'inhibiteur de NF- $\kappa$ B, I $\kappa$ B par un complexe de I $\kappa$ B kinases (IKK $\alpha$ , IKK $\beta$ , IKK $\gamma$  et Nemo). La phosphorylation d'I $\kappa$ B libère le facteur NF $\kappa$ B, induisant sa translocation nucléaire et l'activation de ses gènes cibles. En normoxie, le complexe IKK est inhibé par hydroxylation de IKK $\beta$  par PHD-1. Dans ces conditions, NF- $\kappa$ B est inhibé par I $\kappa$ B. En hypoxie, PHD-1 est inactive et n'induit plus la dégradation des kinases du complexe IKK ce qui permet l'inactivation par phosphorylation de I $\kappa$ B, la libération de NF- $\kappa$ B et l'expression de ses gènes cibles <sup>113,114</sup>.

Enfin, une autre voie de signalisation est activée par l'environnement hypoxique. En effet, l'hypoxie induit un stress oxydatif par la génération de radicaux libres oxygénés (RLO) qui ont un impact délétère sur la cellule en oxydant les lipides, les nucléotides des acides nucléiques (ADN et ARN) ou les cystéines des protéines. La régulation du stress oxydatif est liée à l'activation de la voie anti-oxydante NRF2/NFE2L2. En absence de RLO, le facteur de transcription NRF2 se fixe à la protéine répresseur KEAP1 qui permet son ubiquitination par l'ubiquitine ligase Cullin3 et sa dégradation par le protéasome. A l'inverse, en hypoxie, la génération de RLO oxyde la protéine KEAP1 au niveau de résidus de cystéine impactant sa conformation et libérant le facteur de transcription NRF2 qui transloque dans le noyau afin d'activer ses gènes cibles <sup>115</sup>.



**Figure 12 :** Les différentes voies de signalisation de l'hypoxie (NF-κB, HIF1, NRF2) et leur mécanisme de régulation en normoxie.

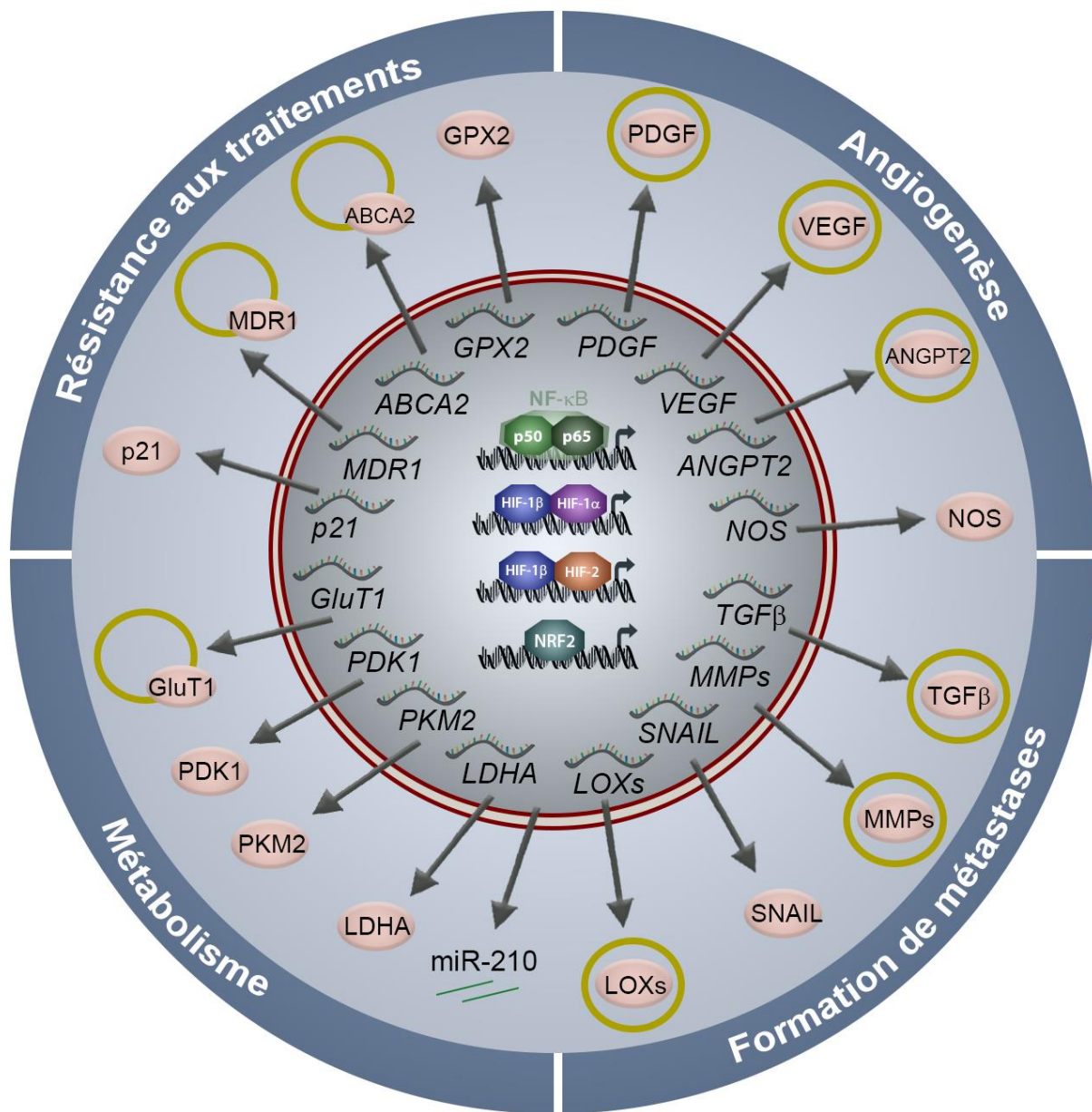
## *ii. Conséquences du microenvironnement tumoral hypoxique*

Les gènes régulés par l'hypoxie permettent aux cellules de survivre dans cet environnement. Ils interviennent dans différents processus biologiques favorisant l'agressivité des cellules cancéreuses notamment la transition d'un métabolisme aérobie à un métabolisme anaérobie, l'angiogenèse, la formation de métastases et la résistance aux traitements <sup>116,117</sup> (Figure 13).

L'un des premiers effets notables concerne le métabolisme des cellules cancéreuses qui subissent un basculement d'un métabolisme aérobie vers un métabolisme anaérobie appelé « effet Warburg » <sup>118</sup>. Les cellules captent le glucose, carburant énergétique des cellules, via le transporteur membranaire GLUT1 (Glucose Transporter 1) et le métabolise par une succession de réactions enzymatiques appelées glycolyse. Dans le cas d'un métabolisme aérobie soit en présence d'oxygène, la glycolyse transforme une molécule de glucose en deux molécules de pyruvate qui est métabolisé en acétyl-Coa par la décarboxylation oxydative. L'acétyl-Coa via le cycle de Krebs permet la production de molécules énergétiques qui seront utilisées au sein des mitochondries par la chaîne de phosphorylation oxydative pour produire de l'ATP, molécule énergétique indispensable à de nombreuses réactions enzymatiques.

En absence d'oxygène, les réactions enzymatiques de décarboxylation oxydative et de phosphorylation oxydative sont inactives. Le pyruvate est transformé en lactate via le processus de fermentation lactique catalysé par la Lactate Déshydrogénase A (LDHA). Le bilan énergétique d'un métabolisme anaérobie est de 2 molécules d'ATP soit bien moins rentable que celui d'un métabolisme aérobie de 36 molécules d'ATP. L'expression de différents gènes est augmentée par le facteur HIF-1 afin que l'apport énergétique soit suffisant et permette la prolifération des cellules cancéreuses. Ainsi, la captation du glucose est augmentée par surexpression du récepteur GLUT1. D'autres activités enzymatiques sont également surexprimées comme la LDHA ou des enzymes de la glycolyse telles que la Pyruvate déshydrogénase 1 (PDK1) et l'isozyme M1/M2 de la Pyruvate kinase (PKM2). Enfin, une autre conséquence de l'effet de Warburg concerne l'acidification du microenvironnement par exportation du lactate qui se trouve en excès au sein des cellules hypoxiques <sup>108,117</sup>.





**Figure 13:** Conséquences des voies de signalisation de l'hypoxie.

Les différents facteurs de transcription induit par l'hypoxie (HIF1, HIF2, NRF2, NF-κB) permettent l'expression de différents gènes codants des protéines impliquées dans le métabolisme énergétique, l'angiogenèse, la formation des métastases ou la résistance aux traitements.



La diminution du taux d'oxygène est également un stimulus d'activation des processus angiogéniques indispensables à la croissance de la tumeur via l'induction de différents facteurs pro-angiogéniques par les facteurs de transcription HIFs. C'est le cas de l'oxyde nitrique synthase (NOS) et des facteurs de croissance VEGF, PDGF (platelet-derived growth factor) et angiopoïétine 2 (Ang2)<sup>78,119,120</sup>. L'induction de ces protéines et l'activation des récepteurs à activité tyrosine kinase, tels que le VEGFR2 pour la création de nouveaux vaisseaux sanguins ou le VEGFR3 pour la création de nouveaux vaisseaux lymphatiques, favorisent la migration, la prolifération et la perméabilité de l'endothélium<sup>66,69,121</sup>.

L'hypoxie favorise également la transition épithélio-mésenchymateuse (TEM) des cellules cancéreuses nécessaire à la formation de métastases. La TEM se traduit par l'augmentation de marqueurs des cellules mésenchymateuses tels que les protéines Snail, Slug, la cadhérine N (CDH2), l' $\alpha$ -SMA, le TGF $\beta$ , et inversement par la diminution de marqueurs épithéliaux tels que la cadhérine E (CDH1) et par l'acquisition de capacités migratoires et invasives des cellules<sup>122</sup>. Les métastases sécrètent notamment des metalloprotéases matricielles (MMP1-9) et des lysyl oxydases (LOX et LOX-Like 1-4) qui modifient la rigidité de la matrice extracellulaire et favorisent la migration cellulaire<sup>123,124</sup>.

Le contexte hypoxique rend également les cellules cancéreuses plus résistantes aux traitements anticancéreux par la sur-expression de différents transporteurs (ABCA2, MDR1 et P-gp) permettant l'efflux des drogues chimiothérapeutiques, par l'inhibition de la voie p53 et par l'activation de la voie anti-oxydante NRF2<sup>125</sup>. En effet, la radiothérapie ou certains agents chimiothérapeutiques tels que le cisplatine produisent des RLO qui sont neutralisés par l'activation de la voie NRF2<sup>115,126</sup>. D'ailleurs, NRF2 est fréquemment activé de façon constitutive dans les cancers pulmonaires via la répression de son inhibiteur KEAP1, expliquant en partie la forte chimio- et radiorésistance de ces tumeurs<sup>127</sup>.

Les effets de l'hypoxie sur ces différents processus cellulaires peuvent également impliquer des gènes non codants. Nous reviendrons plus en détails sur cette problématique dans les chapitres suivants.

## C. Traitements des cancers broncho-pulmonaires non à petites cellules

L'agressivité des cancers broncho-pulmonaires varie en fonction du stade, du sous-type histologique, du statut mutationnel et de la nature du micro-environnement immunitaire<sup>33</sup>. Les traitements doivent être adaptés en fonction de ces différents paramètres. La médecine personnalisée est donc au centre des choix thérapeutiques des cancers broncho-pulmonaires. L'enjeu majeur est de découvrir de nouveaux marqueurs pronostiques et de nouvelles cibles thérapeutiques.

### 1. Diagnostic et Médecine personnalisée

La médecine personnalisée s'appuie sur les caractéristiques des tumeurs, c'est à dire sur la stadification, le statut mutationnel et la nature de l'infiltrat immunitaire et stromal. Dans ce but, une nouvelle classification des tumeurs a été mise en place par différents organismes comme l'International Association for the Study of Lung Cancer (IASLC), l'American Thoracic Society (ATS) et l'European Respiratory Society (ERS). La première étape dans la démarche de la médecine personnalisée consiste à augmenter les outils diagnostiques pour améliorer la prise en charge des patients. Le premier test diagnostique mis en œuvre consiste à caractériser la tumeur primaire par imagerie médicale (radiologie rayons X, scanner et PET-scan). Ensuite, une biopsie est prélevée par chirurgie afin d'identifier le sous-type histologique et d'établir le statut mutationnel afin de proposer le traitement le plus adapté (Chimio- ou Immunothérapie).

### 2. Exérèse curative

L'ablation chirurgicale ou exérèse curative est proposée lorsque le stade d'avancement de la tumeur est inférieur au stade IIIb, c'est à dire lorsque la tumeur se limite aux poumons. Dans les cas précoces (stade 0 et I) et lorsque le patient est en bonne santé, la chirurgie reste la meilleure opportunité de guérison pour le patient. Selon la taille et la colonisation des cellules cancéreuses, les chirurgiens pourront effectuer soit une résection, soit une lobectomie, soit une pneumonectomie (excision de l'ensemble d'un poumon). Dans les cas de tumeurs de stade IIIa, des traitements de chimiothérapies et/ou de radiothérapies sont réalisés avant la chirurgie. Enfin, après exérèse curative, l'analyse des marqueurs génétiques de la tumeur permet de déterminer les facteurs d'agressivité de la tumeur (micro-environnement tumoral et mutations) afin d'adapter éventuellement un traitement adjuvant.

### 3. Chimiothérapie

Le traitement chimiothérapeutique peut être local après chirurgie ou systémique. Le traitement est adapté aux résultats anatomo-cliniques post-chirurgie. La chimiothérapie consiste le plus souvent

en une combinaison de sels de platines (carboplatine, cisplatine ou oxaliplatine) et d'inhibiteurs des voies de signalisation altérées par les mutations (EGFR, ALK, KRAS ou BRAF1) <sup>128</sup>. Les molécules ciblant les voies de signalisation inhibent l'intégration ou la transmission des signaux intracellulaires tandis que les sels de platines sont des agents alkylants qui se fixent sur les bases puriques (A ou G) de l'ADN induisant d'une part des modifications conformationnelles qui bloquent la réplication et la transcription et d'autre part un stress oxydatif par la production de RLO entraînant l'apoptose des cellules cancéreuses <sup>129</sup>.

#### 4. Radiothérapie

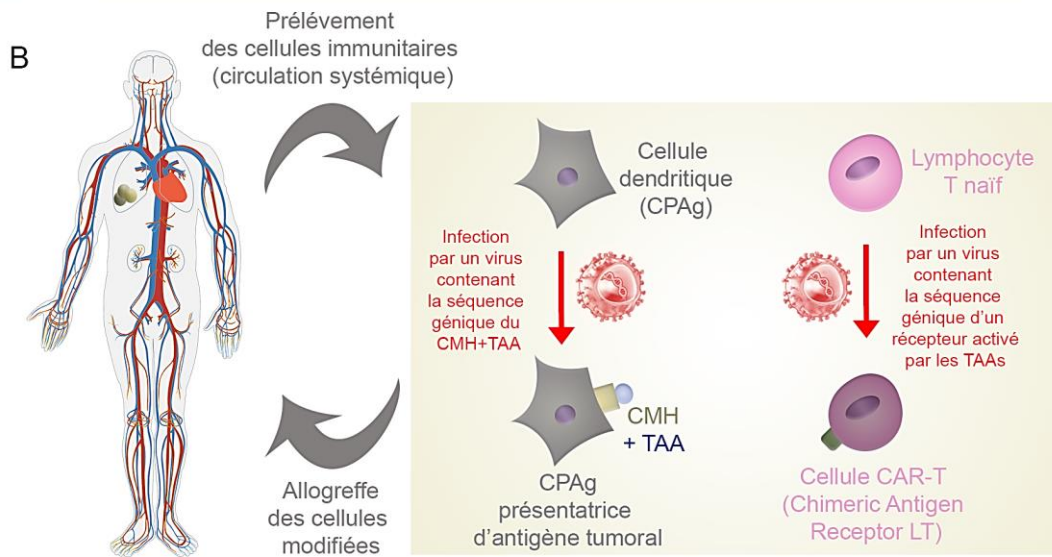
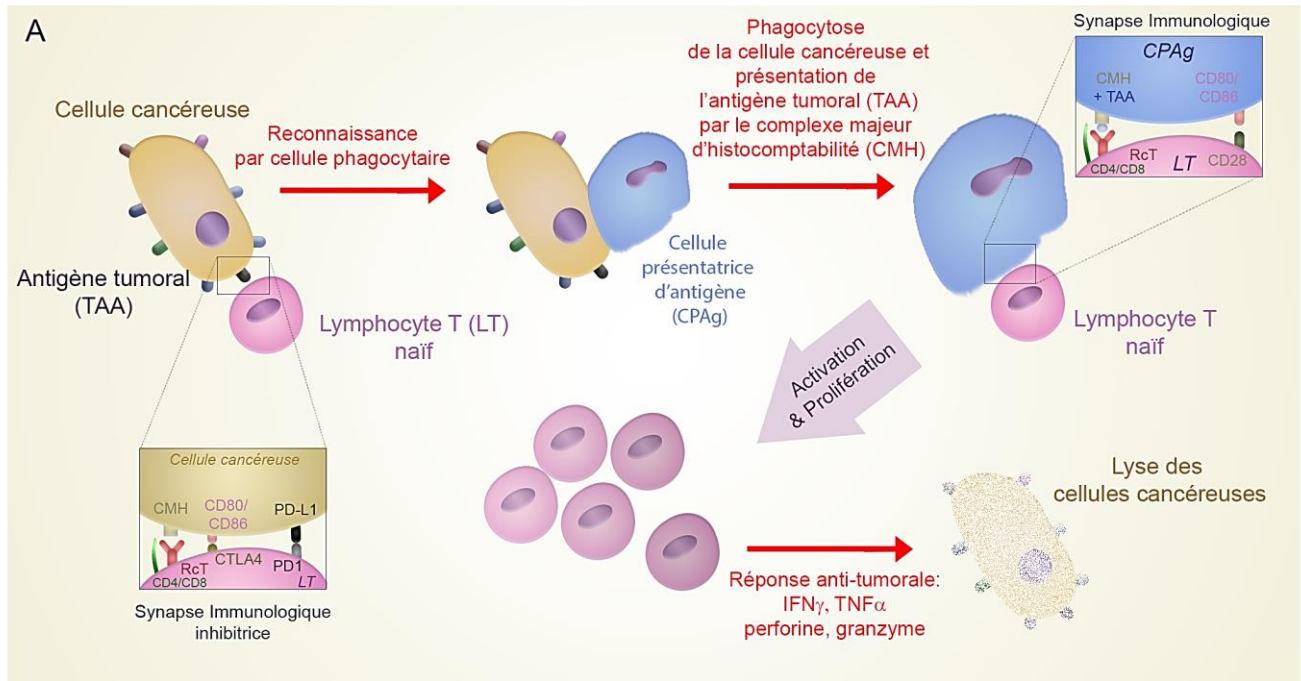
La radiothérapie est principalement utilisée comme traitement palliatif lorsque la tumeur est inopérable. Un traitement radiothérapeutique induit de nombreux effets secondaires et entraîne une fibrose du tissu pulmonaire. Contrairement à la chimiothérapie, l'adjonction de la radiothérapie à la chirurgie n'améliore pas la survie des patients à cause des nombreux effets secondaires toxiques affaiblissant les patients <sup>130</sup>.

#### 5. Immunothérapie

L'immunothérapie, utilisée avec succès dans le traitement de divers cancers, apparaît aujourd'hui comme le traitement le plus prometteur dans la lutte contre les CBNPC dû à leur forte charge mutationnelle produisant une grande variété d'antigènes. L'immunothérapie consiste à activer le système immunitaire du patient pour tuer les cellules tumorales. Comme je l'ai précédemment mentionné, les cellules cancéreuses interagissent avec les cellules immunitaires par le biais de récepteurs membranaires créant une synapse immunologique (Figure 14A). Au sein de cette synapse, il existe des interactions moléculaires qui favorisent la réponse immunitaire anti-tumorale et d'autres récepteurs appelés points de contrôle ("checkpoint") qui modulent l'activation des cellules immunitaires afin de limiter la durée et l'intensité de la réaction immune. Ces points de contrôle induisent une tolérance permettant à la cellule tumorale d'échapper à l'immunosurveillance de l'organisme. C'est le cas des récepteurs CTLA4 (Cytotoxic T-Lymphocyte Associated protein 4), PD1 (Programmed Cell Death Protein 1) exprimés à la surface des lymphocytes T ou de PD-L1 (ligand de PD1) exprimé à la surface des cellules cancéreuses.

L'un des aspects de l'immunothérapie consiste à utiliser des anticorps comme le nivolumab (anticorps anti-PD1) ou des inhibiteurs chimiques afin de bloquer ces interactions et stimuler l'immunité anti-tumorale <sup>65,131</sup>. D'autres outils thérapeutiques combinent des approches de thérapies géniques et cellulaires. Les cellules CAR-T (Chimeric Antigen receptor Lymphocyte T cell) sont des lymphocytes T de patients modifiées in vitro à l'aide d'un virus comportant la séquence génique de récepteurs

membranaires reconnaissant les antigènes spécifiques des tumeurs (TAAs, Tumor-Associated Antigen) (Figure 14B) <sup>132-134</sup>. Ces cellules sont ré-injectées au patient par allogreffe afin de cibler spécifiquement les cellules tumorales. Une stratégie semblable utilise les CPAg. Les cellules dendritiques (DC) de patients prélevées dans la circulation sanguine sont génétiquement modifiées à l'aide d'un virus contenant la séquence génique de TAAs. Les DC présentent à leur surface les TAAs qui permettent la reconnaissance et l'activation des lymphocytes T <sup>131,135,136</sup>.



**Figure 14:** La réponse immune anti-tumorale et les stratégies de l'immunothérapie.

**A.** Les cellules cancéreuses sont phagocytées par des cellules dendritiques ou des macrophages qui présentent des antigènes spécifiques des cellules tumorales (TAA, Tumor-Associated Antigen) associés au complexe majeur d'histocompatibilité (CMH). Les cellules présentatrices d'antigènes (CPAg) vont s'associer au sein d'une synapse immunologique avec les récepteurs des lymphocytes T naïfs. L'intégration des signaux permet l'activation et la prolifération des lymphocytes T anti-tumoraux. A l'inverse, la synapse immunologique entre les lymphocytes T et les cellules cancéreuses est inhibitrice. **B.** Allogreffe de cellules immunitaires (CPAg ou LT) modifiées génétiquement par un vecteur viral afin d'induire une réponse immunitaire anti-tumorale.

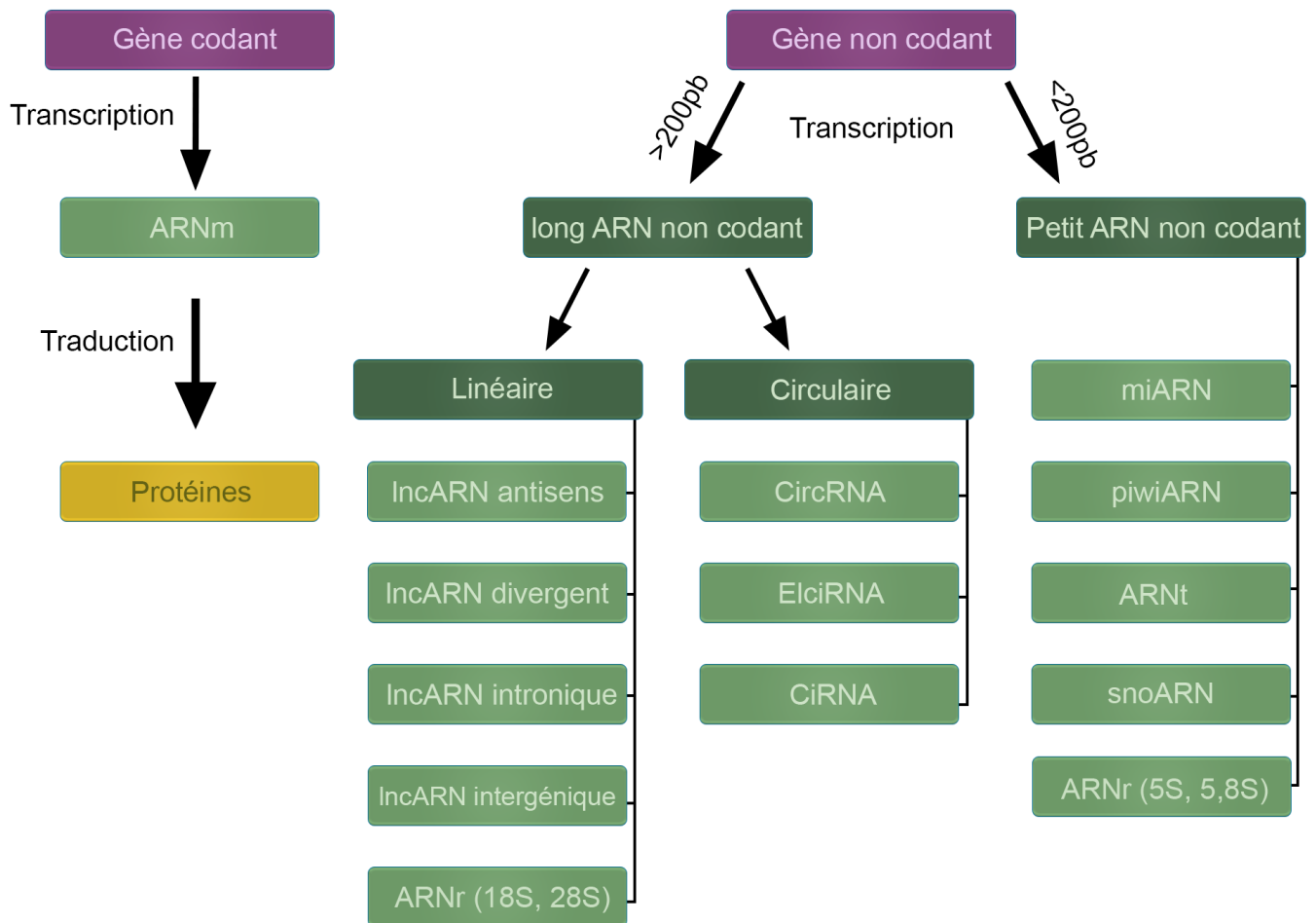
## Chapitre 3 : Classification des ARNs non codants

### A. De « l'ADN poubelle » aux ARNs non codants

La structure de l'acide désoxyribonucléique (ADN) a été mise en évidence en 1953 par Francis Crick et James Watson à l'aide des travaux de Rosalind Franklin qui a mis en évidence sa structure hélicoïdale <sup>137,138</sup>. Francis Crick a également formulé le dogme central de la biologie moléculaire selon lequel l'ADN est segmenté en unités fonctionnelles appelées « gènes » qui sont transcrits en acides ribonucléiques (ARN). Dans les années 1950, on considérait que seulement 2% du génome était fonctionnel et permettait la transcription d'ARNs dits messagers (ARNm). Cela représentait environ 15 000 gènes codants des protéines. Les 98% restant du génome ont été nommés « ADN poubelle » dans lequel ont été mis en évidence des séquences permettant la transcription d'ARNs impliqués dans la traduction des ARNm en protéines tels que les ARNs ribosomiques (ARNr) et les ARNs de transfert (ARNt) (Figure 15) <sup>139</sup>.

En 1990, une nouvelle famille de transcrits non-codants a été découverte par des approches de génétique chez *C. elegans* : les microARNs (miARN) <sup>140,141</sup>. Ces ARNs de 20 nucléotides sont synthétisés dans le noyau sous forme de précurseurs de grande taille, puis maturés et exportés dans le cytoplasme où ils se fixent par complémentarité de séquence à des ARNm cibles afin d'induire leur dégradation. C'est ainsi, que la notion d'ARNs régulateurs de l'expression génique a été mise en évidence.

Par la suite, en 2004, les consortiums ENCODE (projet ENCyclopedia Of DNA Element) et FANTOM (projet Fonctionnal ANnotation Of The Mammalian Genome) ont vu le jour afin de cartographier toutes les unités fonctionnelles du génome humain. Les résultats ont permis l'identification de 30 000 à 40 000 gènes transcrits en ARNs dont 25 000 gènes transcrits en ARNs non codants regroupant les piwiARNs (piARN), les snoARNs, les ARNs Y et les longs ARNs non codants (LncARNs). Les autres séquences génomiques humaines sont en grande majorité des séquences répétées, transposables et des séquences régulatrices de l'expression telles que les promoteurs des gènes, les séquences enhancers qui activent l'expression des gènes et les séquences insulateurs qui répriment l'expression des gènes <sup>142,143</sup>. Le consortium ENCODE a récemment publié les résultats de sa phase III de production avec presque 6000 expériences de séquençage d'ARNs provenant de lignée cellulaire, de cellules primaires ou de tissus humains (n= 4 834) et murins (n=1 158). Ces expériences permettent de préciser l'interaction de protéines avec les différentes classes d'ARNs, l'expression de transcrits dans les tissus et les cellules primaires ainsi que les modifications et la fixation des facteurs des transcriptions sur la chromatine (ChIP, Chromatine Immunoprecipitation) <sup>144</sup>.



**Figure 15:** Classification des acides ribonucléiques codants et non codants chez les mammifères.

On distingue les ARNs codants permettant la synthèse des protéines et les ARNs non codants partagés en longs ARNs non codants et en petits ARNs non codants. Les longs ARNs non codants regroupent les antisens (NAT, Natural Antisense Transcript), les divergents, les introniques et les intergéniques en fonction de leur position génique par rapport aux gènes codants et également certains ARN ribosomiques (ARNr 18S et 28S). Parmi les petits ARNs non codants, on distingue les microARNs et les piwiARNs qui sont des régulateurs de l'expression génique, et les ARN ribosomiques (ARNr 5S et 5,8S), les ARNt et les snoARN qui régulent la traduction des ARNm en protéines. Adapté de <sup>145</sup>.

## **B. Les ARNs impliqués dans la traduction**

### **1. Les ARNs ribosomiques**

Une fois les ARNm transcrits et épissés, ils sont pris en charge par les ribosomes à l'extérieur du noyau pour être traduits en protéines. Les ribosomes sont des complexes ribonucléoprotéiques constitués par des ARNs ribosomiques (ARNr) et des protéines. Les gènes des ARNr sont localisés dans le nucléole des noyaux des cellules eucaryotes. Il existe plus de 200 copies des gènes d'ARNr. Les ARNr sont transcrits par l'ARN polymérase I (ARN pol I) puis maturés par des protéines snoRNP (small nucleosomal RiboNucleoProtein). Les différents types d'ARNr permettent de réaliser les différentes sous-unités des ribosomes, la petite sous-unité 40S est composée des ARNr 18S et de 33 protéines différentes tandis que la grande sous-unité (60S) est composée des ARNr 28S, 5S, 5,8S et de 49 protéines. La petite sous-unité du ribosome a un rôle de lecture de l'ARNm et de validation de l'appariement des ARN de transfert (ARNt). La grande sous-unité, également appelée peptidyltransférase crée les liaisons peptidiques entre les différents acides aminés afin de produire une protéine. Les ARNr de cette grande sous-unité ont donc une activité enzymatique et sont appelés ribozymes (Figure 16) <sup>146</sup>.

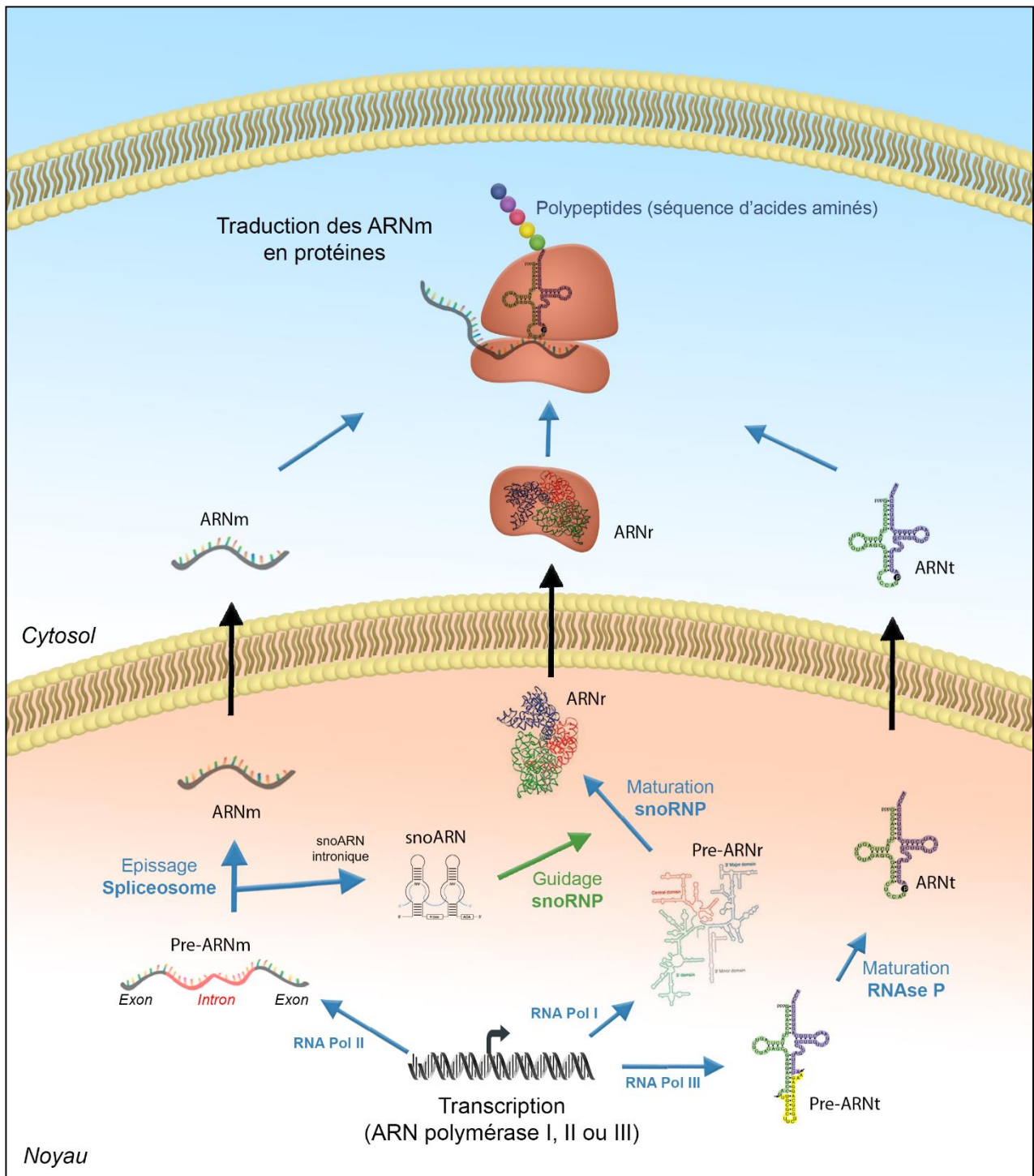
### **2. Les snoARNs**

Les snoARNs sont de petits ARNs nucléolaires des cellules eucaryotes. Leur fonction est liée à la maturation des ARNr. Ils permettent de guider par complémentarité de séquences les protéines snoRNP au niveau des ARNr et d'induire des modifications chimiques de type 2'-O-méthylation pour les snoARNs de type C/D et pseudouridylation pour les snoARNs de type H/ACA (Figure 16) <sup>147</sup>.

### **3. Les ARNs de transfert**

Les gènes des ARNt appartiennent aux génomes mitochondriaux et nucléaires et sont transcrits par l'ARN polymérase III (ARN pol III), puis maturés en 5' par la ribonucléase P et sur l'extrémité 3' par l'ARNt nucléotidyl-transférase ou "CCase" qui intervient pour que la séquence se termine par le codon CCA. Les ARNt ont une structure canonique en « feuille de trèfle » qui se transforme en « L » une fois associée avec le ribosome. Il existe 20 ARNt isoaccepteurs différents qui se fixent spécifiquement aux 20 acides aminés existants. Les ARNt associés à un acide aminé sont nommés aminoacyl-ARNt et se lient aux ribosomes ou une extrémité de la séquence des ARNt porte un anti-codon permettant la reconnaissance d'un codon de l'ARNm à traduire en protéine (Figure 16) <sup>148</sup>.





**Figure 16:** ARN ribosomique, ARN de transfert et snoARN sont les ARN régulateurs de la traduction des ARNm en protéines.

Certains ARNs sont essentiels à la traduction des ARNm en protéines. Les ARNr transcrits par l'ARN polymérase I et les ARNt transcrits par l'ARN polymérase III sont maturés au sein du noyau puis exportés dans le cytoplasme afin de former les ribosomes permettant la traduction des ARNm en protéines. Les snoARNs transcrits par l'ARN polymérase II participent à la maturation des ARNr. Adapté de <sup>149</sup>.

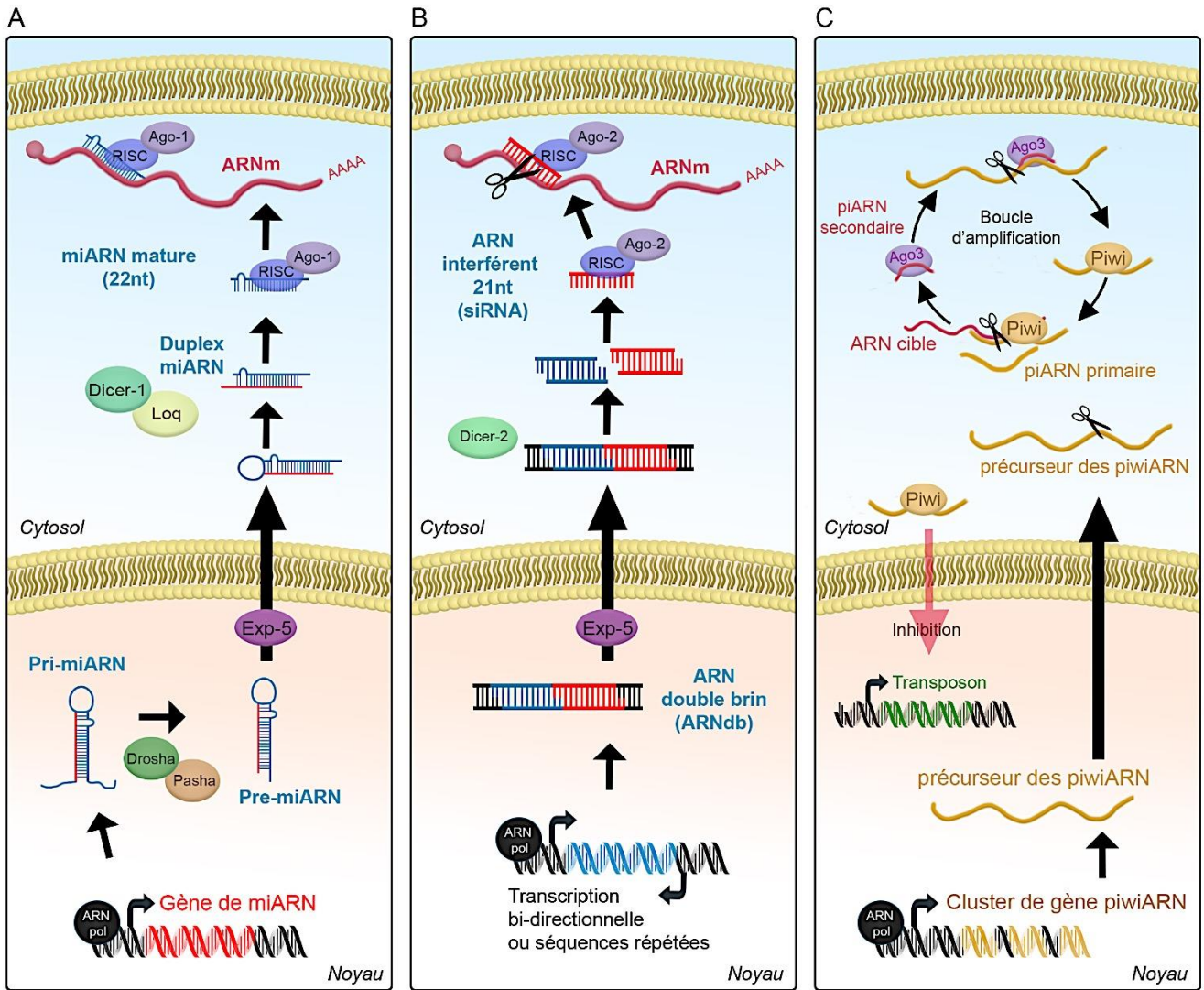
## C. Les petits ARNs régulateurs

### 1. L'ARN interférence et les microARNs

Le mécanisme d'ARN interférence impliquant de petits ARNs non codants a été découvert en 1990 par l'équipe de Jorgensen qui voulait renforcer la couleur violette de pétunias en transduisant le gène de la *chalcone synthetase* (CHS). Cependant, les résultats obtenus ont été contraires à leurs attentes puisque les fleurs ayant intégré le transgène étaient blanches et présentaient 50 fois moins d'activité enzymatique ainsi qu'une inhibition de 90% de l'ARN de la CHS. Les auteurs se sont interrogés sur l'existence d'un mécanisme de co-suppression (ou d'interférence) de l'ARN du transgène sur le transcrit endogène <sup>140</sup>.

En 1993, une autre étude sur le nématode *C.elegans* a permis d'identifier un petit ARN non codant de 22nt, lin-4, capable de se fixer par complémentarité de séquence sur le 3'-UTR de l'ARNm lin-14. L'invalidation de lin-4 augmente l'expression de la protéine lin-14 et bloque la croissance larvaire des nématodes. Ces expériences ont mis en évidence un mécanisme d'interférence post-transcriptionnel de l'ARNm par la fixation d'un petit ARN nommé par la suite micro-ARN (miARNs) <sup>141</sup>.

Quelques années plus tard, en 2004, les mécanismes de biosynthèse, de maturation et le mode d'action des miARNs ont été démontrés. Les gènes des miARNs sont transcrits par l'ARN polymérase II sous forme de précurseurs de grande taille, les pri-miARNs <sup>150</sup>. Ces pri-miARNs subissent une première étape de maturation nucléaire par un complexe « microprocesseur » formé des enzymes Drosha et DGCR8 afin de produire des pré-miARNs de 70 nucléotides repliés en structure tige-boucle et ayant la forme d'une épingle à cheveux. Les pré-miARNs sont ensuite exportés dans le cytosol par transport actif GTP-dépendant via la protéine exportine 5. Ils sont maturés par la protéine Dicer qui coupe la structure tige-boucle afin d'obtenir de petits ARNs double brin (miARNdb) pris en charge par le complexe multiprotéique RISC (RNA-induced Silencing Complex) contenant notamment la protéine Agonote Ago2 qui va séparer les deux brins (miR-5p et miR-3p). Chacun des brins peut guider le complexe RISC vers un ARNm cible par appariement imparfait au niveau d'une séquence de 7-8nt dans le 3'-UTR du gène appelée séquence « seed » afin de bloquer son expression par divers mécanismes post-transcriptionnels (Figure 17A) <sup>150,151</sup>. Les microARNs sont présents chez tous les eucaryotes et leurs séquences sont conservées traduisant également des fonctions conservées <sup>152</sup>. Actuellement, les miARNs sont connus pour leurs implications dans diverses pathologies dont les cancers et sont utilisés en tant que marqueurs pronostics <sup>153</sup>. Au sein de notre équipe, nous avons mis en évidence l'implication de miR-210 dans les mécanismes de radioresistance et de switch métabolique des cellules cancéreuses hypoxiques <sup>154,155</sup>.



**Figure 17:** Les différentes voies de synthèses des microARNs, siARNs et piwiARNs.

Les mécanismes moléculaires régissant la synthèse des microARNs et des siARNs présentent des similitudes, mais ils se distinguent par les protéines impliquées. **A.** Les miARNs sont transcrits sous forme de pri-miARNs maturés dans le noyau par le complexe Drosha/Pasha (DGCR8) en pré-miARNs ayant une structure tige-boucle. Le pré-microARN est ensuite exporté dans le cytoplasme par l'exportine 5 et maturé par le complexe Dicer1/Loq. Le microARN mature est ensuite recruté par le complexe RISC/Ago-1 et guidé sur des séquences partiellement complémentaires de l'ARNm pour bloquer la traduction et/ou induire la dégradation du transcrit. **B.** Les siARNs proviennent d'une molécule d'ARN double brin (ARndb) transcrite de manière endogène ou introduite par un virus ou par transfection. L'ARndb est transloqué dans le cytoplasme par l'exportine 5 puis maturé par la protéine Dicer2 en petits ARNs double brin. Les siARNs matures sont ensuite recrutés par le complexe RISC/Ago-2 et guidés sur les séquences cibles complémentaires de l'ARNm afin de le couper. **C.** Les piARNs sont synthétisés sous la forme de pré-piARNs dans le noyau, exportés dans le cytoplasme et segmentés avant de s'apparier à un transcrit cible. L'appariement avec l'ARNm cible entraîne leur dégradation et la formation de petits ARNs régulateurs, les piARNs qui sont ensuite transloqués dans le noyau afin d'inhiber la rétrotransposition de séquences génomiques et l'atteinte du matériel génomique germlinal. Adapté de <sup>156</sup>

Depuis, les mécanismes d'invalidation par ARN interférence ont été largement étudiés que ce soit dans l'étude fonctionnelle des miARNs endogènes ou par l'utilisation des small interfering RNA (siRNA) et des *short-hairpin*-RNA (shRNA) en tant qu'outils biotechnologiques pour l'invalidation d'ARNs cibles (Figure 17B) <sup>157</sup>.

## 2. Les piARNs et la régulation des éléments transposables

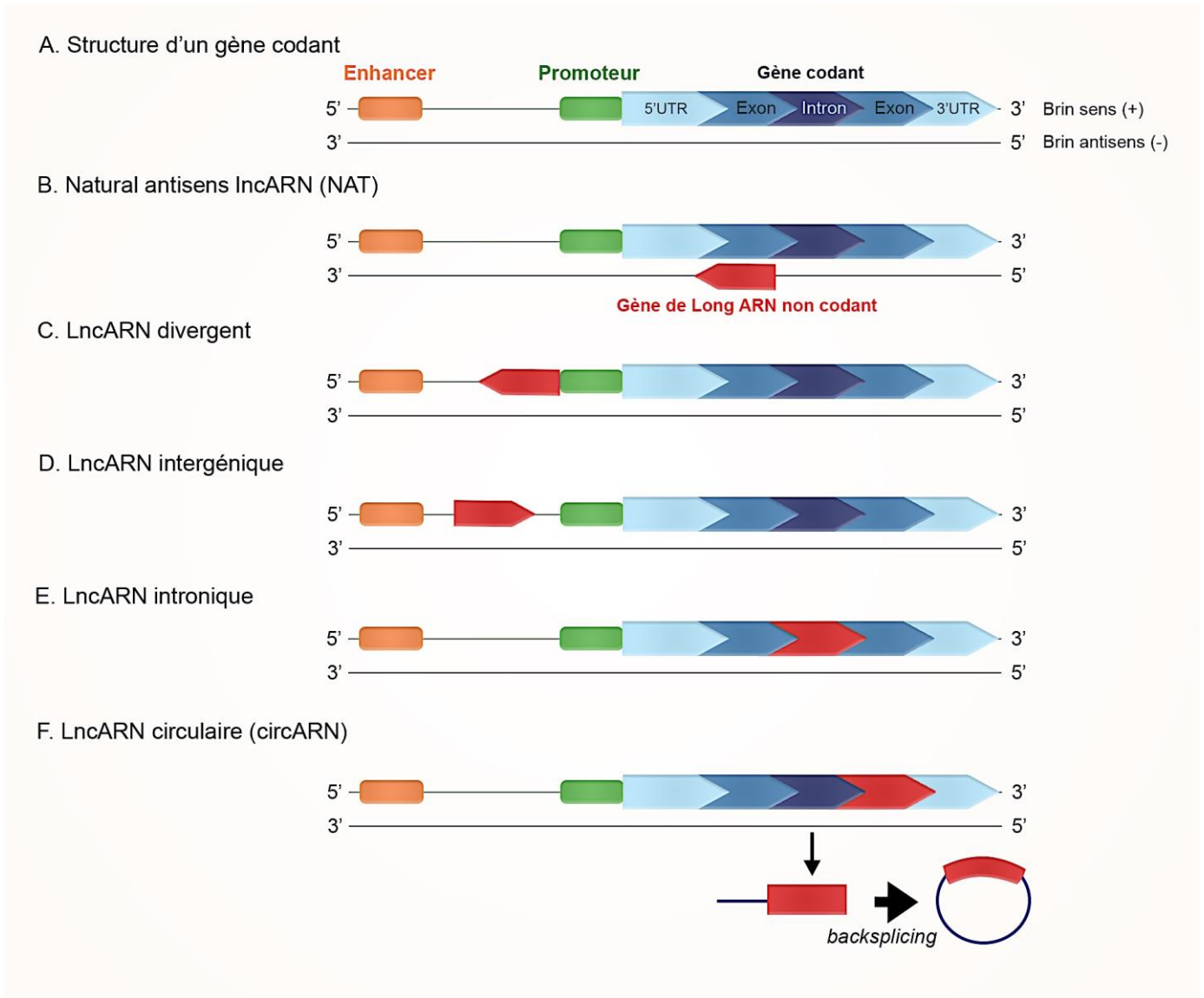
Les piARNs (PIWI-interacting RNA) sont également de petits ARNs régulateurs de 22 à 30nt qui agissent sur la régulation des éléments transposables dans les gamètes. Les piwi-ARNs sont produits et maturés au sein du noyau. Ils possèdent un groupement 2'-O-méthyl en 3' leur permettant de se fixer aux protéines PIWI. Ces complexes piARN-PIWI se fixent ensuite par complémentarité de séquences sur les éléments mobiles appelés transposons afin d'inhiber par un mécanisme complexe le déplacement de ces séquences et garantir l'intégrité génomique des gamètes (Figure 17C) <sup>158</sup>.

## D. Les longs ARNs non codants (lncARNs)

En parallèle des études sur les microARNs, en 1990, une équipe s'intéresse à un ARN de 2,5kb exprimé uniquement par l'allèle maternel du gène *H19* du chromosome 7 de la souris et dont l'inactivation augmente la taille des embryons murins <sup>159</sup>. Une seconde équipe détermine que la fonction de ce gène est liée uniquement à la molécule d'ARN produite et décrit ainsi une nouvelle classe de molécule : Les longs ARNs non codant (lncARNs) <sup>160,161</sup>. Les lncARNs sont des transcrits de plus de 200 nucléotides ne codant pour aucune protéines ou peptides fonctionnels. Il existerait plus de 15 000 gènes transcrits en lncARNs dans le génome humain. Ils sont majoritairement transcrits par l'ARN polymérase II (ARN pol II) avec quelques exceptions comme les ARNs B2-SINE synthétisés par l'ARN pol III <sup>162</sup>.

Les loci des lncARNS sont conservés chez les mammifères mais la séquence primaire des transcrits est peu conservée au cours de l'évolution avec seulement 30% d'homologie de séquence entre les orthologues murins et humains <sup>163</sup>. Toutefois, la fonction des lncARNs est majoritairement liée à leur structure secondaire, notamment aux structures tiges-boucles qui servent de plateformes d'ancrage pour différentes protéines (RBP, RNA Binding Protein) portant l'activité enzymatique des complexes macromoléculaires ribonucléoprotéiques <sup>164</sup>. Ces structures secondaires sont plus fréquemment conservées que la séquence primaire comme cela a été démontré pour le lncARN COOLAIR chez les végétaux permettant ainsi une conservation inter-espèces de la fonction du lncARN <sup>165</sup>.

Les longs ARNs non codants ont une grande diversité de fonctions et de modes d'actions qui dépendent de leurs associations avec des protéines partenaires. Plusieurs modes de classification des lncARNS ont été proposés en se basant notamment sur leur fonction ou leur mode d'action, mais la plus probante repose sur leur localisation génique par rapport aux gènes codants voisins <sup>161,166</sup>. Cette classification basée sur les annotations GENCODE sépare les gènes des lncARNs en 5 familles : les antisens, les divergents, les intergéniques, les introniques et les ARNs circulaires (Figure 18).



**Figure 18:** Classification des longs ARNs non codants en fonction de leur localisation génique par rapport aux gènes codants.

On distingue les longs ARNs non codants (lncARNs) intergéniques, les lncARNs enhancers, ceux associés à un promoteur, les lncARNs divergents, les lncARNs antisens transcrits en orientation inverse d'un gène codant et les lncARNs introniques. Adapté de <sup>167</sup>.

### 1. Les LncARNs antisens

Les longs ARNs non codants antisens ou "Natural Antisens Transcripts" (NATs) sont positionnés sur le brin opposé d'un gène codant <sup>168</sup>. Les séquences des NATs peuvent être chevauchante avec l'ensemble du transcrit codant ou avec une fraction de sa séquence incluant sa région promotrice (PROMPTs), un exon, un intron ou encore les séquences UTRs, ce qui modifie leur mode d'action. En effet, les NATs sont séparés en deux sous-familles, les *cis*-NATs qui affectent l'expression du gène du même locus car ils ont une bonne complémentarité de séquence ou les *trans*-NATs qui affectent un gène à une autre localisation génique. Ils agissent en empêchant la transcription, en déstabilisant les ARNm ou encore en favorisant leur traduction comme cela est le cas des SINEUPs (SINEB2 sequence to UP-regulate translation) <sup>169-171</sup> (Figure 18B).

### 2. Les LncARNs divergents

Les gènes des lncARNs divergents partagent les séquences promotrices d'un gène codant ou d'un autre gène non codant mais en orientation inverse (Figure 18C). Ils correspondent à environ 15% des lncARNs. Les promoteurs des gènes divergents sont généralement très conservés au cours de l'évolution. Ils contiennent de grandes régions riches en îlots CpG et sont associés à des gènes du développement et de la différenciation <sup>172,173</sup>.

### 3. Les LncARNs intergéniques

Les gènes des lncARNs intergéniques (lincARN) ont leur propre locus en dehors de ceux des gènes codants. Par conséquent, ils possèdent leurs propres unités transcriptionnelles (Figure 18D). Ils constituent la classe prédominante des lncARNs avec près de 10 000 gènes différents, soit près de 60% des lncARNs. Les lncARNs intergéniques exercent une grande diversité de fonctions par des mécanismes agissant au niveau transcriptionnel ou post-transcriptionnel.

### 4. Les LncARNs introniques

Enfin, les gènes des lncARNs introniques se trouvent dans les introns des gènes codants. On dénombre environ 100 gènes permettant de transcrire un lncARN intronique (Figure 18E). Ils peuvent servir de précurseurs de miARNs, d'éponges à miARNs, moduler l'épissage d'ARNm ou réguler la transcription de l'ARNm d'un gène voisin.

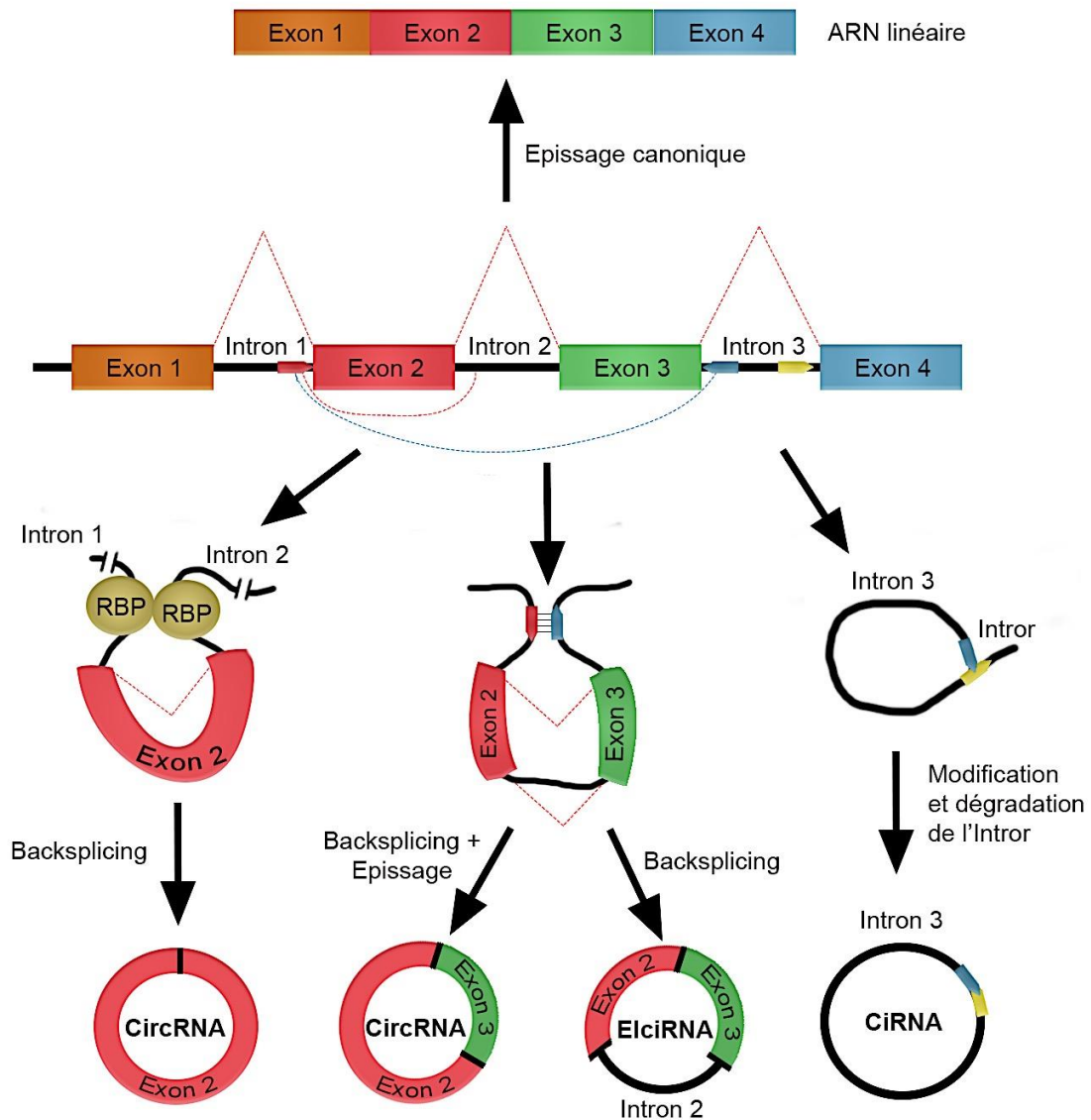
## 5. Les lncARNs circulaires

La dernière famille de lncARNs est constituée par les gènes qui produisent des ARNs circulaires. Ces gènes peuvent être introniques, intergeniques, antisens ou divergents. Ils se caractérisent par un mécanisme d'épissage circulaire appelé « back-splicing » (Figure 18F).

En fonction de leur composition exonique et intronique, on distingue différents sous-types d'ARNs circulaires, les circARNs constitués exclusivement d'exons, les ElciARNs comprenant des exons et des introns et les CiARNs provenant uniquement de séquences introniques (Figure 19).

Le « back-splicing » nécessite des séquences consensus correspondant à 7 nucléotides GU à l'extrémité 5' et 11 nucléotides C à l'extrémité 3' du gène. Ces ARNs circulaires s'accumulent fréquemment dans le noyau où ils ont une fonction similaire aux lncARNs introniques mais ils peuvent également être exportés dans le cytoplasme <sup>174</sup>. Ils possèdent des répétitions de séquences de reconnaissance des microARNs qui leur permet de les séquestrer ou peuvent se lier à des protéines afin de les stabiliser ou les séquestrer <sup>175,176</sup>.





**Figure 19:** Voies de synthèse des ARNs circulaires par épissage circulaire ou "back-splicing". Les ARNs circulaires proviennent d'un épissage non canonique appelé back-splicing qui permet la formation de structures circulaires. Ils sont classifiés en sous familles en fonction de leur origine. On distingue les circRNAs issus de séquences exoniques, les ElciRNAs provenant de séquences introniques et exoniques et les CiRNAs issus de structures introniques. Adapté de <sup>177</sup>.

## Chapitre 4 : Synthèse et mode d'action des longs ARNs non codants

Les longs ARNs non codants correspondent à une famille de molécules bien distinctes des ARNm et des petits ARNs régulateurs. Ils possèdent des caractéristiques métaboliques tant au niveau de leur synthèse que de leur stabilité qui les différencient des autres classes de transcrits. Ils se caractérisent également par une grande diversité de modes d'actions intrinsèquement liée aux RBP avec lesquelles ils s'associent.

### A. Synthèse et maturation des lncARNs

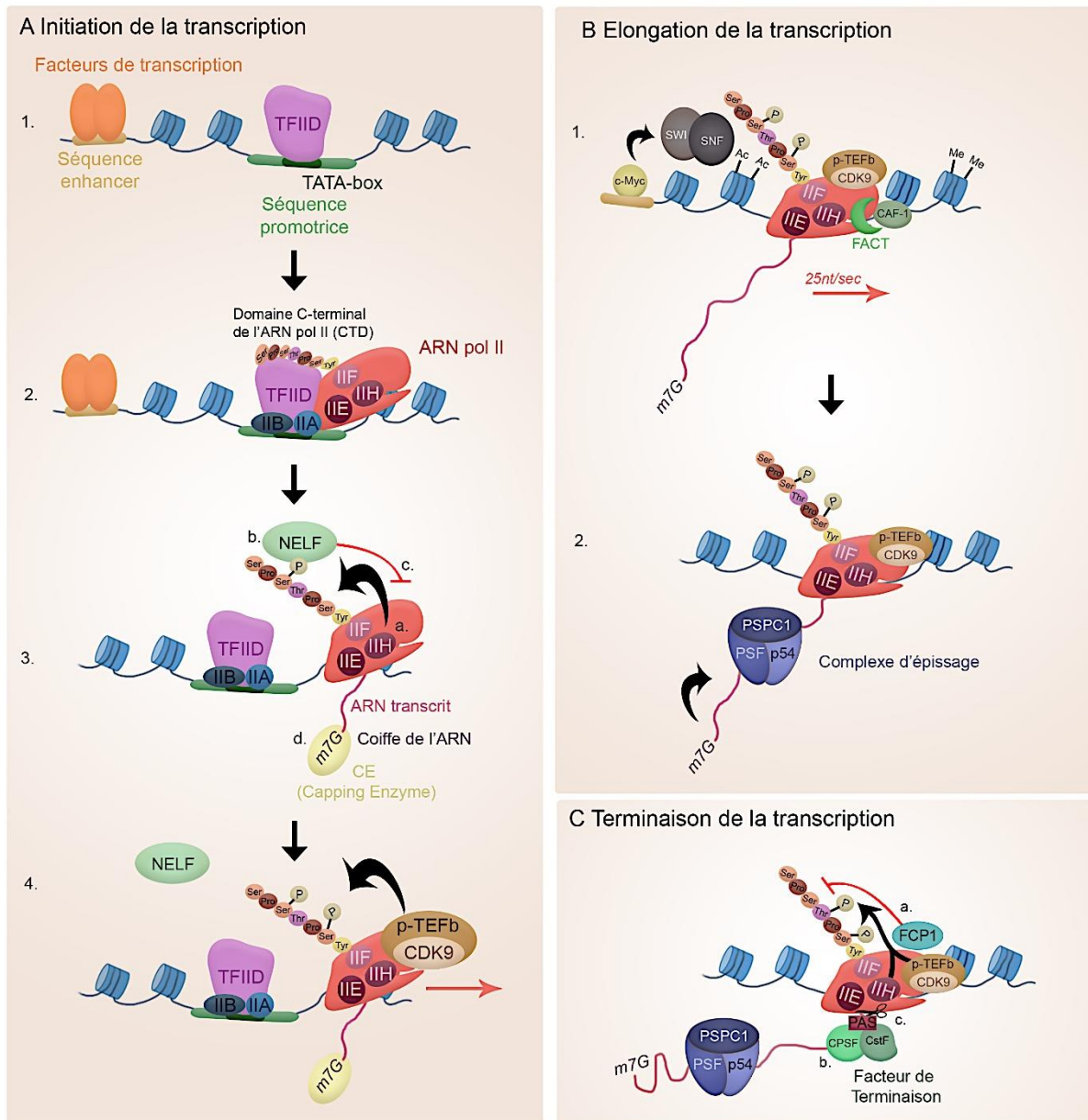
Par comparaison aux ARNm, les lncARNs sont plus faiblement exprimés dans les cellules eucaryotes mais leur profil d'expression apparaît plus spécifique d'un type cellulaire, d'un tissu ou d'une pathologie donnée. Ces différences peuvent être liées à des particularités dans les mécanismes de synthèse, de maturation et de stabilisation/dégradation de ces transcrits <sup>178</sup>.

#### 1. Transcription des lncARNs

La transcription des gènes en ARN se déroule en trois étapes, l'initiation, l'élongation et la terminaison. La capacité de transcription d'un gène est liée à la présence de séquences régulatrices en amont, en aval et dans les introns de ce gène. Les séquences régulatrices peuvent être inhibitrices ou activatrices. Les séquences activatrices sont soit des enhancers, des séquences d'environ 500 pb permettant le recrutement de facteurs de transcription, soit des promoteurs qui se situent directement en 5' du gène et font environ 60pb.

##### *i. Initiation de la transcription*

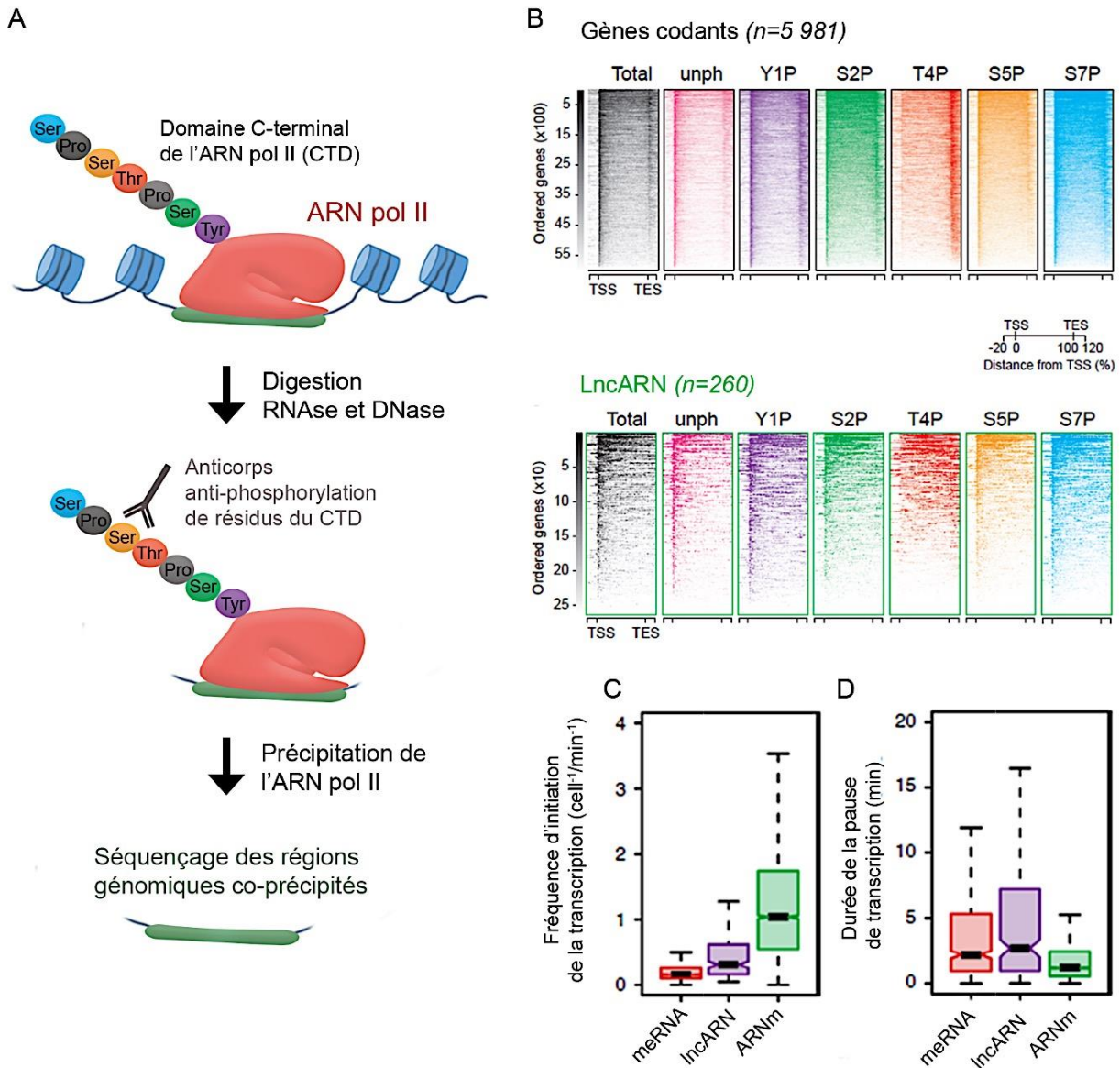
Les séquences promotrices sont composées de séquences consensus appelées TATA box. Cette séquence permet le recrutement du complexe d'initiation de la transcription TFIID (Transcription Factor IID) constitué de la TBP (TATA-binding protein) et de protéines associées, les TAFs (TBP-Associated Factors). TFIID facilite la fixation de l'ARN pol II et d'autres protéines de l'initiation de la transcription telles que le complexe médiateur, TFIIA, -B, -D, -E, -F et -H <sup>179</sup> (Figure 20A). La fixation du complexe kinase TFIIH (ou CDK7, cyclin-dependent kinase 7) permet la phosphorylation du domaine C-terminal de l'ARN pol II (CTD) composé de 52 répétitions du motif Tyr<sub>1</sub>-Ser<sub>2</sub>-Pro<sub>3</sub>-Thr<sub>4</sub>-Ser<sub>5</sub>-Pro<sub>6</sub>-Ser<sub>7</sub> <sup>180</sup>. La phosphorylation du résidu Ser<sub>5</sub> permet le recrutement de la protéine NELF (Negative elongation factor) à 20-30 pb du promoteur qui stoppe la progression de l'ARN pol II. Cette pause permet de recruter les enzymes impliquées dans la formation de la « coiffe » à l'extrémité 5' de l'ARN (Capping Enzyme, CE) qui modifie notamment le premier nucléotide en 7-méthylguanosine (7mG). Ensuite, la fixation de la kinase CDK9 (cyclin dependent kinase 9) appartenant au complexe p-TEFb phosphoryle la Ser<sub>2</sub> du domaine CTD pour décrocher la protéine NELF et libérer l'ARN pol II <sup>181</sup>.



**Figure 20:** Mécanismes d'initiation, de pause et d'élongation transcriptionnelle de l'ARN Pol II. La transcription est un mécanisme complexe comprenant trois étapes : l'initiation, l'élongation et la terminaison. **A.** L'initiation commence par la fixation du complexe TFIID sur la TATA-box du promoteur, facilitée par les facteurs de transcription se liant sur les enhancers (1). Grâce au TFIID, l'ARN polymérase II (ARN pol II) se fixe ainsi que de nombreux co-facteurs (2). Parmi ces co-facteurs, TFIIF/CDK7 phosphoryle la Ser<sub>5</sub> du domaine C terminal de l'ARN pol II (CTD) permettant le recrutement de NELF qui bloque l'ARN pol II. Cette pause permet aux enzymes de la coiffe (CE) de modifier l'extrémité 5' de l'ARN néo-synthétisé (3). Ensuite le recrutement par les enzymes de la coiffe de p-TEFb contenant la kinase CDK9 phosphoryle la Ser<sub>2</sub> du CTD et décroche NELF (4). **B.** L'élongation est facilitée par les modifications d'histones remodelant la chromatine (1). Conjointement à l'élongation, les protéines de la coiffe recrutent le complexe d'épissage (spliceosome) afin de relarguer les introns de l'ARN (2). **C.** A la fin de la transcription, la phosphatase FCP1 est recrutée et déphosphoryle la Ser<sub>5</sub> tandis que CDK7/CDK9 phosphoryle la Thr<sub>4</sub> du CTD. La phosphorylation de Thr<sub>4</sub> et les complexes d'épissages permettent le recrutement des facteurs de terminaison (CPSF et CstF) qui coupent l'ARN au niveau du signal de polyadénylation (PAS).

L'efficacité des étapes d'initiation de la transcription, liées à la fixation de l'ARN pol II et à sa phosphorylation au niveau des résidus Ser<sub>5</sub> et Ser<sub>2</sub> du domaine CTD, est différente dans le cas des lncARNs. Deux équipes ont analysé les transcrits naissants à l'échelle génomique par des approches de NET-seq (Native Elongating Transcript Sequencing). Ils ont réalisé des expériences d'immunoprécipitations de la forme non phosphorylée et des différents résidus phosphorylés (Tyr<sub>1</sub>, Ser<sub>2</sub>, Thr<sub>4</sub>, Ser<sub>5</sub>, Ser<sub>7</sub>) du CTD de l'ARN pol II couplées à l'identification par séquençage des régions génomiques co-précipitées (Figure 21A). La forme non phosphorylée de l'ARN pol II a permis de déterminer le taux de recrutement de l'ARN pol II sur les promoteurs tandis que les différentes phosphorylations du domaine CTD a permis de quantifier l'efficacité de différentes étapes de la transcription de ces gènes<sup>178,182</sup>. Ces expériences ont démontré que l'ARN pol II est moins recrutée par les promoteurs des lncARNs que par ceux des gènes codants (Figure 21B). D'autre part elle est moins phosphorylée. Ces défauts de recrutement se traduisent par une plus faible fréquence d'initiation de la transcription et une pause plus longue avant l'élongation de la transcription (Figure 21CD). Cette pause trop longue de l'ARN pol II peut conduire à son décrochage et à l'arrêt de la transcription<sup>182,183</sup>. Ainsi, dans le cas des lncARNs, les phosphorylations du CTD sur la Thr<sub>4</sub> et les problèmes d'efficacité de remodelage de la chromatine peuvent entraîner des terminaisons intragéniques, générant de multiples isoformes d'un transcrit n'ayant pas tous une fonction au sein de la cellule. D'ailleurs, la génération de ces multiples isoformes ainsi que leurs instabilités compliquent les annotations des gènes des lncARNs et actuellement de nombreuses bases de données ne détiennent pas les mêmes annotations. Actuellement, L'identification des isoformes majoritaires des lncARNs devient possible par les techniques de capture des ARNs à l'aide d'oligonucléotides spécifiques et séquençage de fragments longs (« long read ») des ARNs capturés<sup>184</sup>.

Le décrochage de l'ARN pol II durant la transcription du gène permet son recrutement sur des enhancers intragéniques, ce qui permet la transcription de différents variants. Ces séquences régulatrices permettent le recrutement de l'ARN pol II à l'intérieur du gène et la production d'un ARN incomplet, présentant de nombreux exons, pouvant être polyadénylé et appelé multiexonic RNA (meRNA) (Figure 22E). Ces enhancers ne sont pas sensibles aux mêmes signaux de transcription que les promoteurs mais utilise la même machinerie de transcription<sup>182</sup>. Ces meRNAs ne sont pas traduits en protéines et appartiennent donc bien à la famille des lncARNs<sup>185</sup>.



**Figure 21:** Différences dans les mécanismes de transcription des lncARNs et des ARNm.

A. Schéma expérimental de NET-seq par co-immunoprécipitation de séquences génomiques à l'aide d'anticorps dirigés contre des formes phosphorylées du domaine CTD de l'ARN pol II. B. Profil de phosphorylation des résidus du domaine CTD durant la transcription de gènes codants et de gènes non codants. C. Les données expérimentales, montrent une plus faible fréquence d'initiation de la transcription des lncARNs par rapport aux ARNm. D. Augmentation du temps de pause lors de l'initiation de la transcription des lncARNs par rapport aux ARNm. (Adapté de <sup>178,182</sup>)

L'initiation de la transcription est également liée à l'état de compaction de la chromatine. En effet, l'ADN est compacté sous forme de nucléosomes composés d'environ 146 pb d'ADN enroulées autour d'un coeur de 8 protéines d'histones (H2A, H2B, H3 et H4). La compaction de l'ADN module son accessibilité et est liée aux modifications post-traductionnelles des histones. Plus précisément, les marques épigénétiques des histones modifient l'accessibilité du promoteur et le déplacement du complexe de transcription sur le gène. Ainsi, la méthylation des résidus de lysine 9 et 36 de l'histone 3 (H3K9me, H3K36me) inhibent la transcription tandis que l'acétylation de la lysine 27 et la triméthylation de la lysine 4 de l'histone 3 (H3K27Ac, H3K4me3) activent la transcription. Ces modifications post-traductionnelles impliquent différentes enzymes appelées histones méthyltransférases (HMT), histones déméthylases (HDM), histones acétyltransférases (HAT) et histones désacétylases (HDAC). Ces enzymes entrent dans la composition des complexes SWI/SNF ou CBP/P300 qui décompactent la chromatine et favorisent la transcription ou inversement du complexe polycomb 2 (PRC2) qui compacte l'ADN et inhibe la transcription. Ces complexes sont recrutés au niveau des promoteurs et/ou des enhancers pour permettre la fixation du complexe TFIID et la phosphorylation du résidu Ser<sub>5</sub> du domaine CTD pour permettre l'élongation de la transcription.

Les promoteurs de la majorité des lncARNs sont fréquemment ciblés par les HMT Swr1, Isw2, et Rsc, réprimant la fixation de l'ARN pol II et limitant ainsi l'initiation de leur transcription. Le cas des lncARNs divergents est différent. Leurs promoteurs contiennent des séquences riches en îlots CpG mais également des marques épigénétiques H3K79me<sub>2</sub> indiquant une transcription active de leur locus <sup>172,186</sup>.

## *ii. Elongation de la transcription*

L'élongation de la transcription est un phénomène très efficace puisque l'ARN pol II transcrit 25nt/s. Cette efficacité est liée aux protéines d'assemblage et de désassemblage des nucléosomes telles que FACT (Facilitate Chromatin Transport) et CAF-1 (Chromatin Assembly Factor 1), ainsi qu'à d'autres signaux tels que la fixation des facteurs de transcription Myc ou p53 sur des enhancers distaux et intragéniques qui favorise le remodelage chromatinien (Figure 20B) <sup>179,187</sup>. Ces mécanismes favorisent l'expression des ARNm et des lncARNs <sup>186</sup>. Concomitamment à cette étape d'élongation, les ARNs sont maturés par épissage. Les complexes d'épissage (spliceosome), constitués de protéines d'épissage telles que PSPC1, p54, PSF et d'ARNs non codants, se fixent au niveau des Ser<sub>5</sub> et Ser<sub>2</sub> phosphorylées du domaine CTD et aux protéines de la coiffe <sup>180</sup>. Ces complexes reconnaissent des séquences d'épissages flanquants les exons de l'ARN néo-synthétisé afin de libérer les introns de la forme mature. Dans le cas des lncARNs, les défauts de phosphorylation impactent la qualité de

l'épissage de ces transcrits et favorisent l'émergence de variants d'épissage contenant des introns  
178,188.

### *iii. Terminaison de la transcription*

En fin d'élongation, la phosphatase FCP1 est recrutée et déphosphoryle la Ser<sub>5</sub> tandis que la Thr<sub>4</sub> est phosphorylée par les kinases CDK7 et CDK9. Ce switch de phosphorylation permet le passage de la phase d'élongation à la phase de terminaison liée au recrutement de protéines de maturation de l'extrémité 3', la polyadénylation. Les complexes d'épissages présents sur le transcrit et la Ser<sub>2</sub> phosphorylée du complexe CTD permettent le recrutement des facteurs de polyadénylation : CPSF (Cleavage and polyadenylation specificity factor) et CstF (Clivage stimulation Factor). Ces facteurs clivent l'ARN au niveau d'un signal de polyadénylation « AAUAA » (PAS) puis ils modifient l'extrémité 3' (Figure 22A). En cas de défaut de polyadénylation ou d'épissage du transcrit, les différents facteurs de maturation de l'ARN recrutent des protéines de l'exosome tels que l'exoribonucléase 5'-3' XRN2 qui dégrade le transcrit (Figure 22E) <sup>180</sup>.

Contrairement aux ARNs codants dans lesquels la forme phosphorylée de la Thr<sub>4</sub> du CTD est associée à l'extrémité 3' du gène et donc à la terminaison de la transcription, les expériences de NET-seq ont révélé que dans le cas des lncARNs cette modification de l'ARN pol II est associée à l'ensemble de la séquence du gène traduisant des terminaisons internes et un défaut de polyadénylation (Figure 21B) <sup>178</sup>. De plus, la fixation des facteurs de polyadénylation par le biais de la Thr<sub>4</sub> phosphorylée au cours de la phase d'élongation entraîne des clivages et des dégradations par l'exosome durant la transcription <sup>178,180</sup>.

Les faibles niveaux d'expression des lncARNs et leur maturation anormale en termes d'épissage ou de polyadénylation s'expliquent par des défauts de remodelage de la chromatine et par des phosphorylations particulières du domaine CTD.

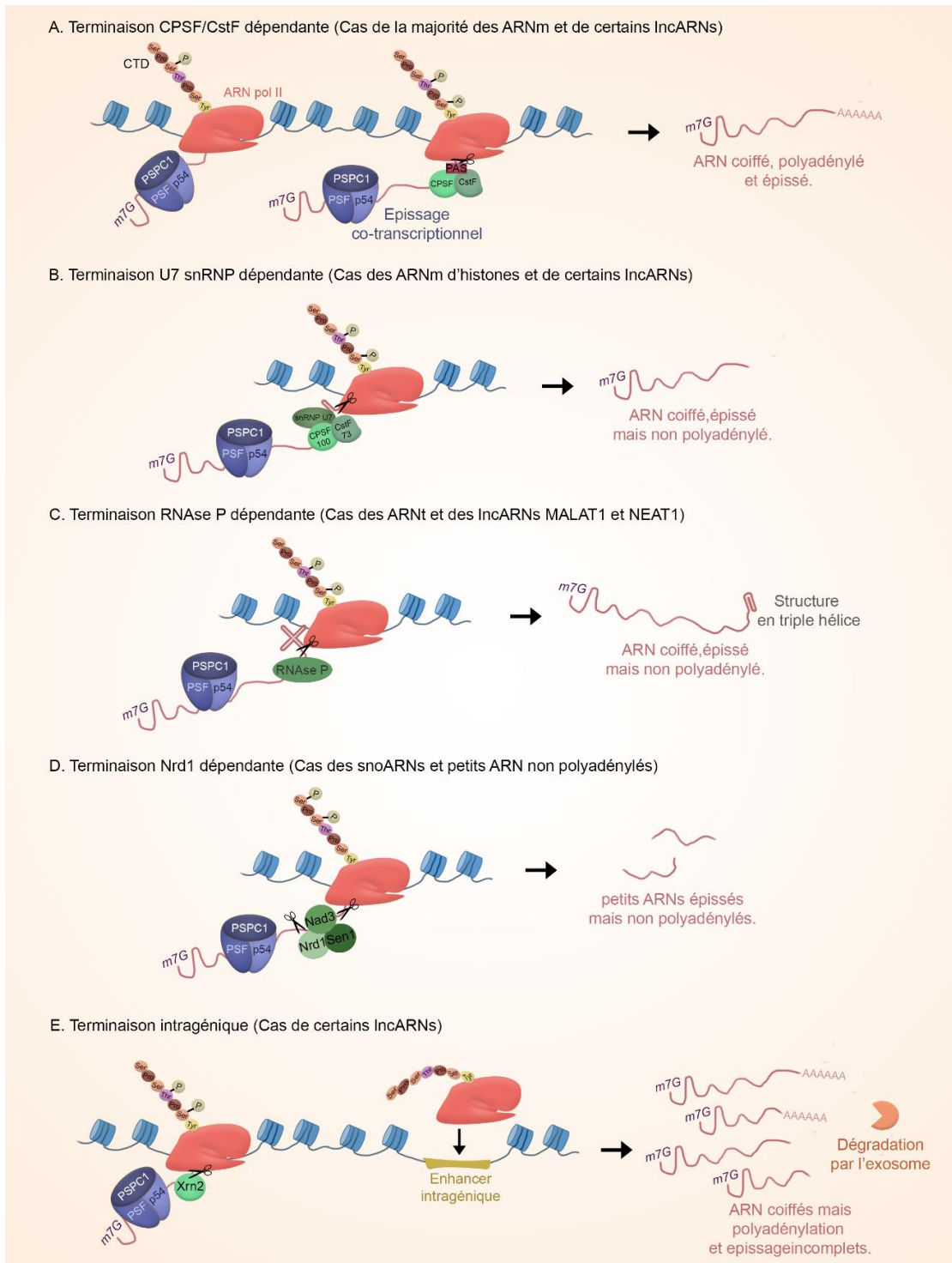
Cependant, certains lncARNs exercent des fonctions essentielles dans la cellule et utilisent d'autres mécanismes de maturation et de terminaison. Ainsi, d'autres mécanismes de terminaison permettent la maturation d'ARNs fonctionnels polyadénylés ou non. Parmi les gènes codants, les ARNm des histones sont une exception puisqu'ils ne sont pas polyadénylés. Ces ARNm, comme certains lncARNs, possèdent une structure secondaire en tige-boucle à l'extrémité 3' qui permet le recrutement de la protéine U7 snRNP et augmente la phosphorylation du CTD sur la Thr<sub>4</sub>. Cela favorise la fixation des protéines CPSF73 et CPSF100 qui clivent le 3' de l'ARNm avant polyadénylation (Figure 22B) <sup>178,180</sup>.

Un autre mécanisme de terminaison particulier concerne les lncARNs MALAT1 et NEAT1 qui possèdent une structure riche en tiges-boucles en forme de trèfle à l'extrémité 3' ressemblant au motif des ARNt. Cette structure particulière recrute la RNase P qui coupe l'extrémité 3' et permet la réorganisation de la structure sous forme d'une triple hélice (Figure 22C). Cette triple hélice permet à ces lncARNs de jouer un rôle d'échaffaudage dans la formation des granules nucléaires <sup>178,186</sup>.

Dans le cas des snoARNs et d'autres transcrits de petites tailles, il existe un mécanisme de maturation non dépendant de la queue polyadénylée. Ce mécanisme repose sur la fixation d'un complexe composé des protéines Nrd1-Nad3-Sen1 au niveau des Ser<sub>5</sub> et Ser<sub>7</sub> phosphorylées qui permettent le clivage de l'ARN en fin de transcription en absence de polyadénylation (Figure 22D)

<sup>178,186</sup>.





**Figure 22:** Différents mécanismes de terminaison de la transcription et de maturation des ARNs.

**A.** La terminaison dépendante des protéines CPSF et CstF est la terminaison majoritaire des ARNm et d'une partie des lncARNs et implique la reconnaissance du signal de polyadénylation (PAS). **B.** La terminaison des ARNm d'histones et de certains lncARNs, liée à une structure « Tige boucle » de reconnaissance de la protéine U7 snRNP, libère un ARN non polyadénylé. **C.** Les ARNt et les lncARNs MALAT1 ou NEAT1 ont une structure « en tréfle » reconnue par la RNase P permettant un clivage en 3'. **D.** Les petits ARNs sont reconnus par le complexe Nrd1. **E.** La plupart des lncARNs ont des terminaisons intragéniques et sont clivés par l'exonucléase XRN2 qui conduit à leur dégradation par l'exosome. Ils ont également des enhanceurs intragéniques. Les gènes de lncARNs produisent beaucoup d'isoformes différentes, plus ou moins épissées et plus ou moins polyadénylées.

## 2. Transport des lncARNs, localisation subcellulaire et potentiel codant

### *i. Transport des lncARNs et localisation subcellulaire*

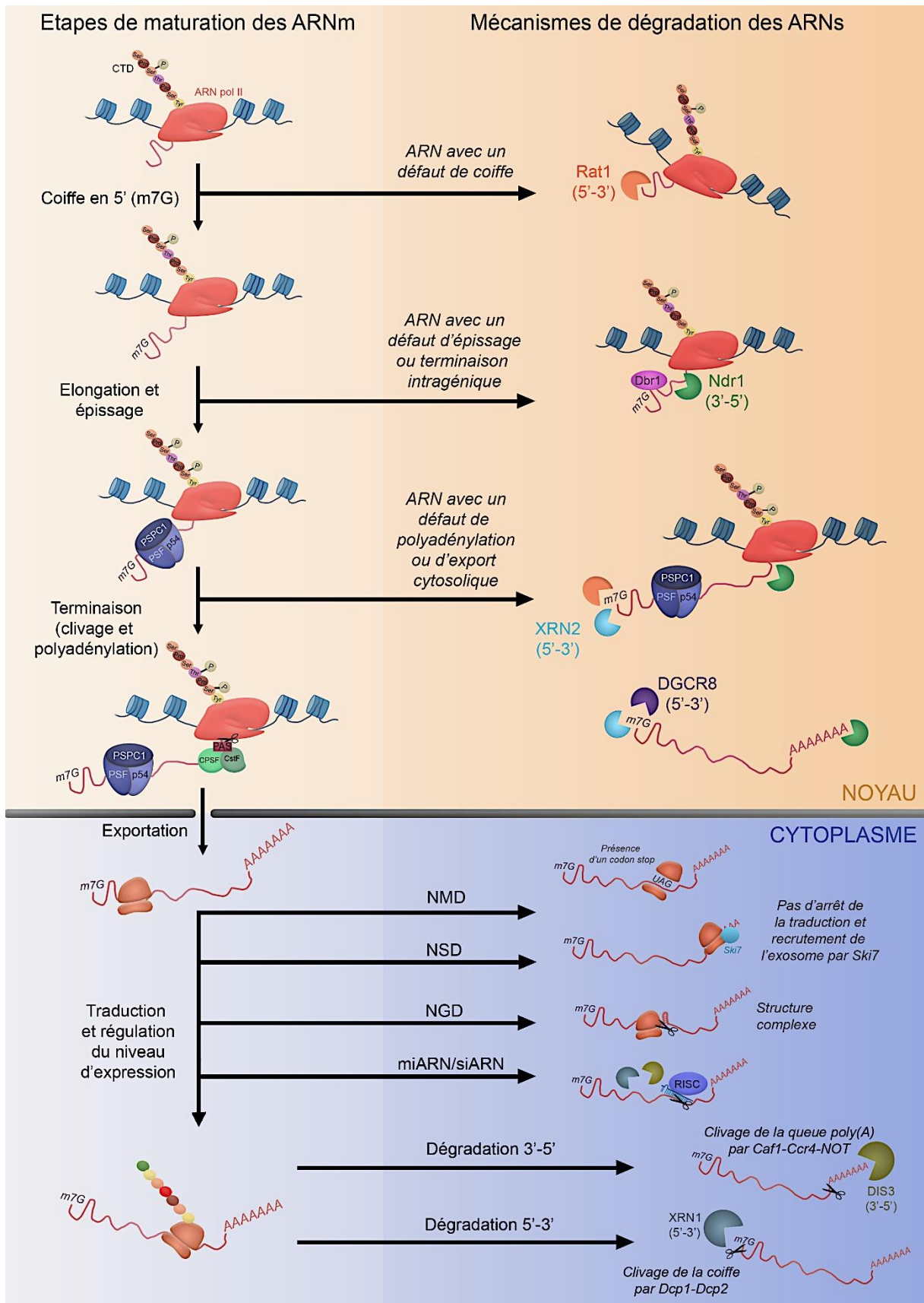
Suite à la transcription, les ARNm sont transloqués dans le cytoplasme afin d'être traduits en protéines (Fig. 16). Ce transport s'effectue grâce aux protéines Mex67/Mtr2 qui se fixent au cours de la transcription via leur interaction avec les complexes d'épissage. Puis suite au clivage de l'ARNm, elles se lient au pore nucléaire, notamment à la nucléopore, afin de permettre la translocation dans le cytoplasme <sup>181</sup>.

Les lncARNs possèdent pour leur part une grande diversité de structures, de fonctions et de modes d'action qui sont directement liés à leur localisation subcellulaire. Cette localisation dépend de leur taille, de la complexité de leur structure, et de la présence de séquences répétées. Les lncARNs sont majoritairement nucléaires et sont pris en charge durant leur transcription par des protéines de stabilisation ou des protéines impliquées dans la médiation de leur fonction <sup>186</sup>. C'est notamment le cas des lncARNs de grande taille (>3kb), très structurés et riches en séquences répétées qui sont quasi exclusivement nucléaires <sup>189</sup>. L'exemple le plus connu concerne NEAT1 qui est pris en charge durant sa transcription par le spliceosome dont il est l'un des constituants (Figure 28) <sup>190</sup>.

L'exportation nucléaire des lncARNs cytosoliques, implique le plus souvent d'autres protéines que celles prenant en charge les ARNm, telles que l'exportine 5 (EXP5) et la petite protéine G RAN en association avec des RBP. Ces protéines reconnaissent des motifs spécifiques ou des structures secondaires (tiges boucles), comme dans le cas des protéines ILF3 ou ILF2 <sup>191</sup>.

### *ii. Le potentiel codant des lncARNs*

Des études de co-précipitation d'ARNs cytosoliques avec la machinerie de traduction ont mis en évidence des interactions entre les lncARNs et les ribosomes pouvant soulever la question de leur potentiel codant <sup>192</sup>. De plus, des analyses bio-informatiques ont détecté la présence de petits cadres de lecture (small Open Reading Frame, sORF) permettant potentiellement la production de petits peptides. Dans le contexte particulier des lncARNs, on parle de lncORFs <sup>193</sup>. Ces lncORFs permettent la production de petits peptides de quelques dizaines d'acides aminés traduits à partir de séquences répétées ou de motifs aléatoires reconnus par les ribosomes. Pour corroborer ces observations, des études d'alignements de séquences, de détection de peptides par spectrométrie de masse et de traduction in vitro ont mis en évidence que la plupart de ces peptides n'existent pas, sont instables ou ne sont pas fonctionnels dans la cellule <sup>194</sup>. Toutefois, d'autres études récentes ont démontré l'existence d'une centaine de peptides provenant des lncARNs qui ont une localisation subcellulaire précise et seraient impliqués dans la prolifération cellulaire par un mécanisme encore inconnu <sup>195,196</sup>.



**Figure 23:** Mécanismes de dégradation des ARNs au cours des différentes étapes de leur synthèse et en fonction des compartiments subcellulaires. Adapté de <sup>197</sup>

### 3. Stabilité et dégradation des lncARNs.

Les défauts de maturation des lncARNs favorisent leur instabilité et leur dégradation. La régulation de la stabilité des ARNs est liée à leur association à des protéines de stabilisation ou de transport se fixant sur différents motifs souvent riches en séquences répétées. Par exemple, les protéines ARE-BP (AU-Rich Element Binding Protein) se fixent sur les motifs riches en séquences AU. Dans cette famille d'une vingtaine de membres, on peut citer PABP1 qui stabilise l'ARN en se fixant sur la queue poly(A), Tia-1 qui favorise l'inclusion d'exons et TTP qui induit la dégradation de l'ARN<sup>198-201</sup>. Certaines RBP agissent sur les transcrits en empêchant la fixation de protéines de dégradation. C'est le cas de HuR qui stabilise les transcrits en bloquant la protéine de déstabilisation AUF1 (hnRNPd) se fixant également sur les séquences riches en AU<sup>202</sup>.

Les mécanismes de dégradation des lncARNs sont identiques à ceux des ARNm bien que leur fréquence diffère (Figure 23). Les voies de dégradation des ARNs dépendent du sens de dégradation du transcrit (5'-3' ou 3'-5') et du compartiment subcellulaire impliqué<sup>197</sup>. Pour les lncARNs, le premier mécanisme de dégradation est lié aux défauts de transcription. En effet, les terminaisons intragéniques permettent le recrutement de l'ARN en cours de synthèse par les exonucléases XRN2 et DGCR8 de l'exosome qui dégradent le transcrit dans le sens 5'-3'<sup>178,180</sup>. D'autres exonucléases interviennent dans le cas de défaut de maturation nucléaire des ARNs, comme par exemple l'exonucléase Rat1 qui dégrade en 5' les transcrits ayant un défaut de coiffe, ou la protéine Dbr1 en cas de défaut d'épissage des transcrits. Le complexe de terminaison Nrd1 des ARNs polyadénylés est également une exonucléase dégradant les transcrits non fonctionnels dans le sens 3'-5'.

Les transcrits cytosoliques, incluant les lncARNs, sont dégradés, soit par les systèmes de dégradation NMD (Nonsense-mediated mRNA decay), NSD (Non-stop decay) ou NGD (No-Go decay) liés à un défaut de traduction, soit par différentes exonucléases régulant leur niveau d'expression. Dans le cas de l'exonucléase XRN1 qui dégrade de 5' en 3', les ARNs sont préalablement pris en charge par les protéines de clivage de la coiffe, Dcp1/Dcp2. Dans la situation inverse de dégradation de 3' en 5', les ARNs sont dépolyadénylés par les protéines Caf1/CCR4/Not puis dégradés par DIS3. Les ARNs peuvent également être la cible de miARNs ou de siARNs qui induisent leur dégradation. Dans ce cas, le clivage de l'ARN par le complexe RISC entraîne la prise en charge de l'extrémité 5' par la protéine XRN1 et de l'extrémité 3' par l'exosome composé de la protéine DIS3<sup>197</sup>.



## B. Modes d'action des lncARNs

Les lncARNs médient leur fonctions par une très grande diversité de mécanismes d'action qui dépendent de leur localisation subcellulaire, de leur séquence primaire, de leur structure secondaire et de leur propension à former des complexes macromoléculaires en s'associant à des protéines effectrices ayant une activité enzymatique et/ou en s'appariant à des séquences génomiques ou à d'autres ARNs <sup>203,204</sup>. La caractérisation de ces modes d'action constitue une problématique centrale dans le domaine des lncARNs notamment par l'identification de leurs partenaires protéiques et de leurs gènes cibles (ADN ou ARN) à la lumière de récents développements technologiques combinant des approches biochimiques et des analyses à haut débit par spectrométrie de masse ou par séquençage.

L'affinité et la cinétique des interactions entre un lncARN et ses partenaires moléculaires peuvent rendre difficile la précipitation des complexes macromoléculaires en conditions natives. La création de liaisons covalentes entre les différents constituants de ces complexes macromoléculaires par le biais d'irradiation aux U.V. ou l'utilisation de molécules chimiques (formaldéhyde, glutaraldéhyde...) est donc souvent un prérequis dans ces différentes approches <sup>205</sup>.

### 1. Approches technologiques spécifiques à l'étude du mode d'action des lncARNs

#### *i. Caractérisation des séquences génomiques interagissant avec des lncARNs*

Dans le cas des lncARNs nucléaires, certains transcrits régulent l'expression génique en se liant à des séquences génomiques. Afin d'identifier ces sites d'interactions, des techniques ont été développées telles que le ChIRP-seq (Chromatine Isolation by RNA purification) <sup>206</sup>. Cette approche consiste à utiliser des sondes oligonucléotidiques antisens biotinylées afin de précipiter la chromatine associée à un lncARN d'intérêt. Les séquences génomiques co-précipitées sont ensuite analysées par séquençage d'ADN à haut débit et alignement sur le génome de référence. Il existe des variantes de cette technique comme le dChIRP (domain ChIRP) qui prend en compte la structure secondaire prédictive des lncARNs pour désigner des sondes antisens contre des domaines potentiellement fonctionnels du transcrit <sup>207</sup>. Le CHART (Capture Hybridization Analysis of RNA Targets) est également dérivée du ChIRP <sup>208</sup>. L'accessibilité des sondes antisens sur les lncARNs est préalablement validée par cartographie par digestion des duplex ARN-ADN par la RNase H dans les conditions de précipitation (Figure 24A). Enfin, le GRID-seq (Global RNA Interactions with DNA by Deep sequencing) consiste à créer des liaisons covalentes entre l'ARN et les séquences génomiques à l'aide d'oligos biotinylés permettant la formation de triplex ARN-ADN double brin <sup>209</sup>. Ces complexes macromoléculaires sont ensuite précipités et analysés par séquençage d'ADN (Figure 24A).

### *ii. Caractérisation des ARNs partenaires des lncARNs*

Les lncARNs nucléaires et cytosoliques peuvent également se lier à d'autres transcrits. Le ChIRP est également utilisé pour identifier ces interactions tout comme d'autres techniques telles que le RAP (RNA antisense purification) et le LIGR-seq (LIGation of interacting RNA followed by high-throughput sequencing) (Figure 24B). Le RAP consiste à précipiter les lncARNs à l'aide d'une sonde antisens biotinylée de 120 nt puis d'analyser par RT-qPCR ou séquençage d'ARN les différents ARNs capturés <sup>211</sup>. Le LIGR-seq consiste à créer des liaisons covalentes entre les molécules d'ARNs par irradiation aux U.V. puis de digérer les séquences d'ARN libres à l'aide de la RNase R, et de séquencer les duplex d'ARNs <sup>212</sup>.

### *iii. Identification des partenaires protéiques des lncARNs*

Les lncARNs nucléaires et cytosoliques interagissent également avec différentes RBP (RNA binding protein) régulant le métabolisme des ARNs et/ou impliquées dans la médiation de leur fonction. Afin d'identifier ces partenaires protéiques, on peut utiliser les techniques de ChIRP ou de CHART couplées à une analyse par spectrométrie de masse (Figure 24C). Le "RNA pull-down" consiste à utiliser un lncARN transcrit *in vitro* puis biotinylé pour précipiter des RBP présentes dans des extraits cellulaires <sup>213</sup>. Les protéines précipitées sont séparées sur gel d'électrophorèse et identifiées par Western blot ou par spectrométrie de masse. Le site d'interaction à l'échelle du nucléotide peut également être déterminé par la formation de liaisons covalentes, la digestion chimique de l'ARN et l'analyse par spectrométrie de masse de la zone d'interaction non digérée <sup>214</sup>.

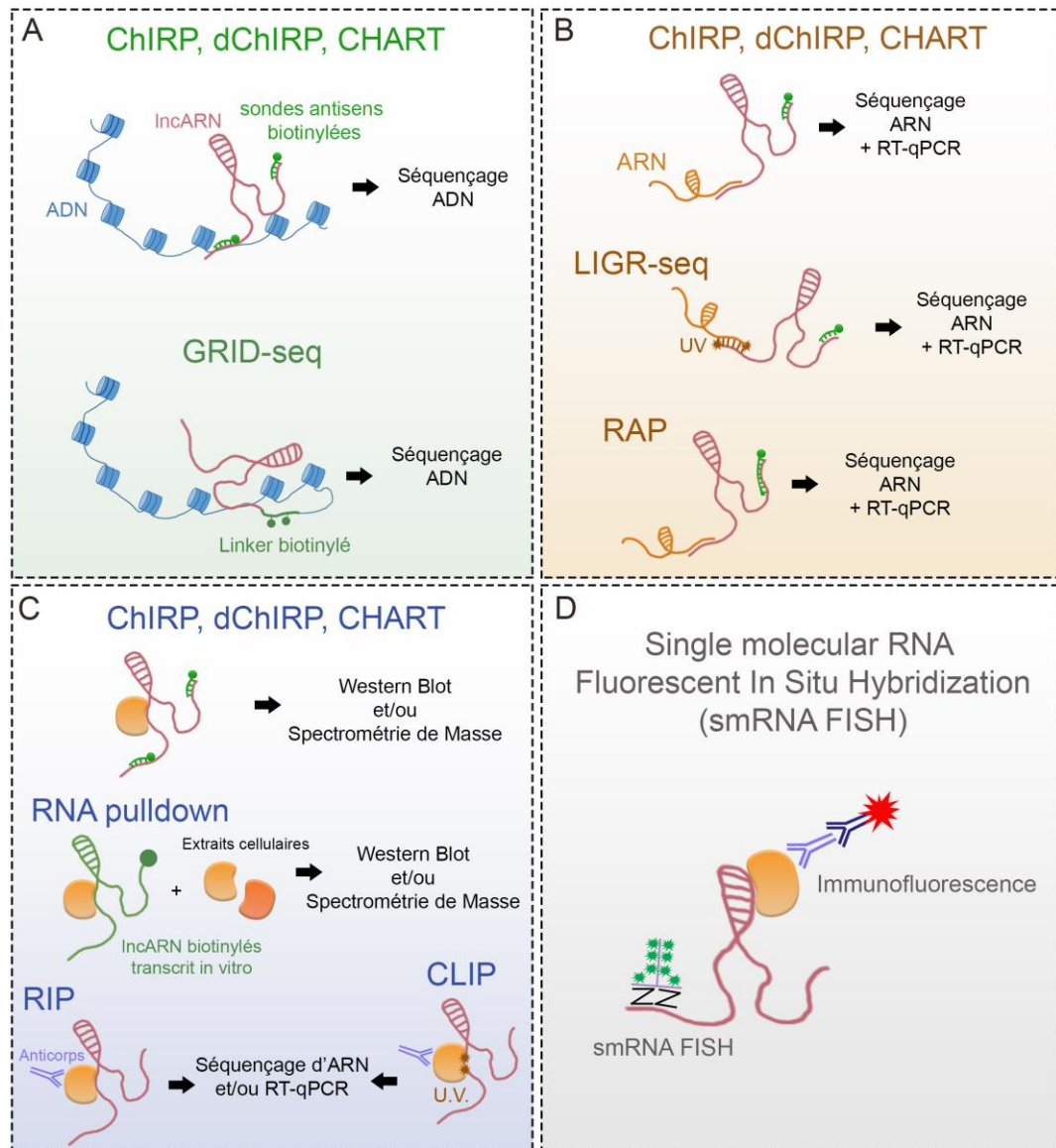
L'interaction entre un transcrit d'intérêt et ses RBPs partenaires peut également être mise en évidence par des approches inverses consistant à immunoprécipiter les protéines à l'aide d'anticorps. Dans le cas du RIP (RNA Immunoprecipitation), l'immunoprécipitation est réalisée sur les complexes natifs d'un extrait cellulaire ou subcellulaire, alors que dans le cas du CLIP (Crosslinked Immunoprecipitation), les constituants des macrocomplexes sont préalablement "crosslinkés". Les complexes macromoléculaires précipités sont ensuite dénaturés et la présence du lncARN d'intérêt est révélée soit par RT-qPCR, soit par séquençage d'ARNs <sup>215,216</sup>.

### *iv. Approches d'imagerie dédiées à la visualisation des interactions entre les lncARNs et leurs RBPs partenaires*

Les interactions mises en évidence par des approches biochimiques peuvent être confirmées par des co-marquages des transcrits et de leurs interactants protéiques. Ces approches permettent également de préciser la localisation tissulaire ou subcellulaire des complexes, ainsi que leur dynamique dans des processus cellulaires ou des pathologies.

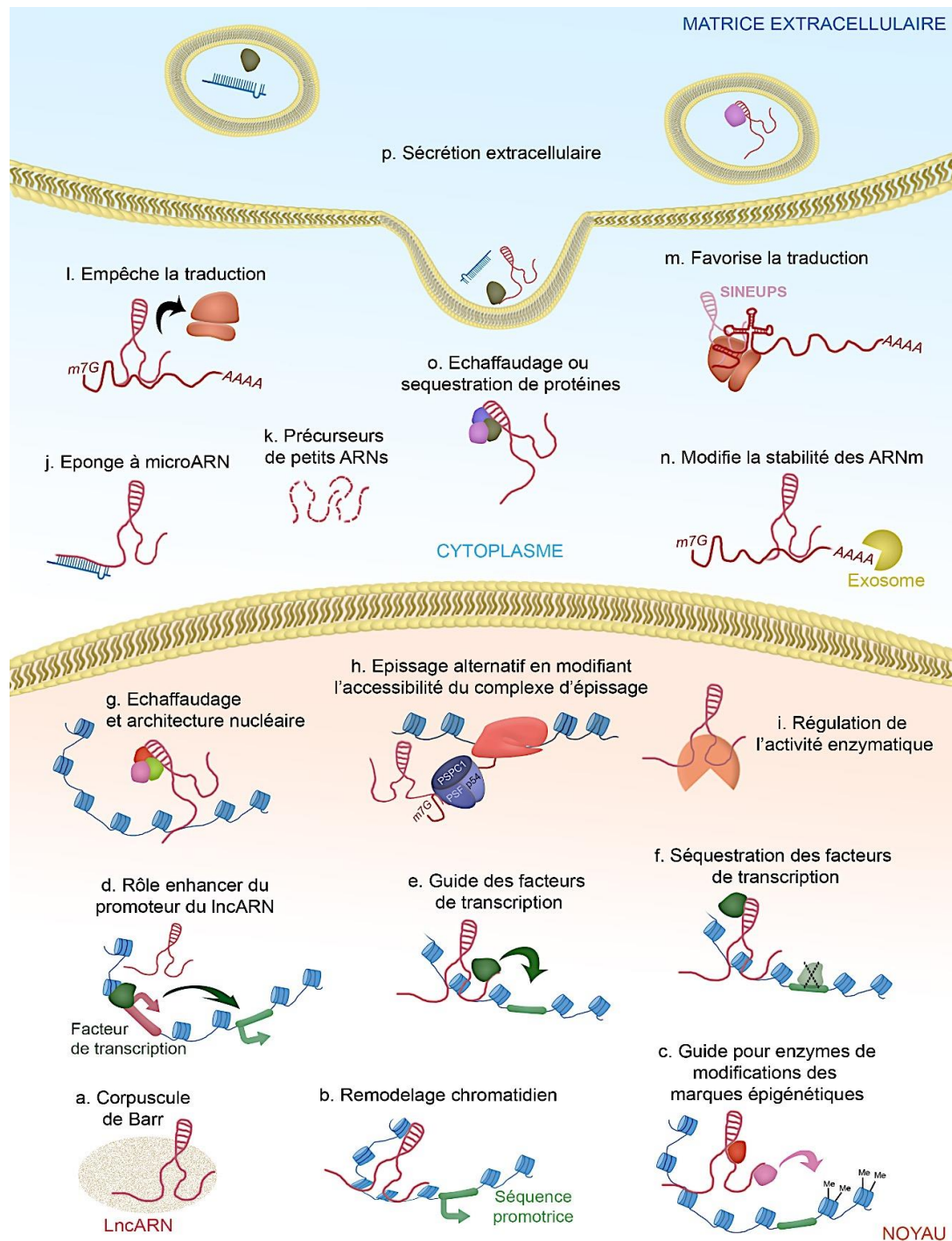


Les lncARNs étant faiblement exprimés dans les cellules, des approches d'hybridation in situ d'ARNs à l'échelle d'une molécule unique (Single Molecular RNA Fluorescent In Situ Hybridization, smRNA FISH) ont été développées <sup>217</sup>. A l'instar du ChIRP, cette technique utilise des lots de sondes oligonucléotidiques antisens spécifiques du transcrit d'intérêt dont l'appariement permet la fixation ultérieure d'un arbre moléculaire d'amplification du signal fluorescent garantissant également la spécificité du signal (Figure 24D). Cette technique peut être couplée à l'immunomarquage de protéines permettant la visualisation des complexes ribonucléoprotéiques sur cellules et coupes de tissus.



**Figure 24:** Méthodologies appliquées à l'étude du mode d'action des lncARNs. Techniques utilisées pour l'identification : A. des séquences génomiques interagissant avec des lncARNs, B. d'ARNs partenaires et C. d'interactants protéiques des lncARNs. D. Techniques de visualisation de la co-localisation de lncARNs et de protéines.





**Figure 25 :** Mécanismes moléculaires d'action des lncARNs en fonction de leur localisation subcellulaire.

Les lncARNs exercent une grande diversité de fonctions liées en premier lieu à leur localisation subcellulaire. On distingue les lncARNs nucléaires (a-i) et les lncARNs cytosoliques (j-p). Les lncARNs nucléaires peuvent : engendrer de manière directe ou indirecte des modifications épigénétiques (a-c), induire l'expression par leurs séquences enhanceurs (d) stimuler l'expression des gènes en favorisant l'accessibilité de facteurs de transcription à des séquences régulatrices (e), ou l'inhiber par séquestration de facteurs de transcription (f). Ils peuvent aussi jouer le rôle d'échaffaudage pour des protéines de structure (g), des facteurs impliqués dans la maturation des ARNs (h) ou des enzymes (i). Les lncARNs cytosoliques peuvent avoir des rôles de séquestration des microARNs (j), servir de précurseurs aux microARNs (k), empêcher la traduction (l) ou au contraire la favoriser (m). Ils peuvent également avoir un rôle de dégradation ou stabilisation d'ARNm (n), altérer la localisation ou l'activité de protéines (o) ou être sécrétés dans le milieu extracellulaire avec différents co-facteurs (p). (Adapté de

210)

## 2. Modes d'action des lncARNs nucléaires

Nous aborderons dans un premier temps les lncARNs nucléaires qui exercent soit des rôles structuraux soit des fonctions de régulation de l'expression génique.

### *i. Adressage de facteurs de transcription et modifications épigénétiques*

Xist est le premier lncARN eucaryote dont le mode d'action a été décrypté. Il participe à l'inactivation d'un des chromosomes X dans l'ensemble des cellules chez les mammifères femelles, le Corpuscule de Barr, d'où son nom : X inactive specific transcript (Figure 25-a). Il participe au mécanisme de compensation de dose pour éviter que les femelles ne produisent deux fois plus de certaines protéines codées par des gènes du chromosome X que les mâles. Xist s'apparie au niveau de séquences génomiques complémentaires et s'associe via sa structure secondaire à des modificateurs épigénétiques de l'organisation de la chromatine (Figure 25-c). Les protéines partenaires permettent d'inhiber durablement l'expression des gènes. Précisément, Xist est recruté par la protéine SAF-A (Scaffold Attachment Factor A). Il se lie ensuite au complexe SHARP qui désacétyle les histones à proximité. SHARP permet également l'association avec la protéine hnRNPK qui inhibe l'acétylation des histones et avec les complexes PRC2 qui méthylent la lysine 27 de l'histone 3. La propagation des méthylations d'histones permet le recrutement de l'ADN méthyltransférase (DNMT) qui méthyle durablement les îlots CpG (Figure 26A) <sup>203,218–220</sup>.

Un autre exemple d'extinction d'expression génique en trans par régulation épigénétique concerne le lncRNA HOTAIR (HOX Transcript Antisens RNA) qui recrute le complexe PRC2 au niveau du promoteur du gène du développement HoxD <sup>221</sup>. Cependant, de récents travaux remettent en cause ce mode d'action car HOTAIR est bien capable de lier PRC2 avec une forte affinité mais la délétion de PRC2 n'affecte pas l'action répressive de HOTAIR <sup>222</sup>. Ces données soutiennent également la réévaluation du paradigme fonctionnel entre d'autres lncARNs et PRC2 <sup>223</sup>.

Les lncARNs régulent également l'expression génique en favorisant la fixation d'activateurs ou de répresseurs transcriptionnels au niveau des promoteurs de gènes (Figure 25-e). Ainsi, le lncARN CCND1 réprime le facteur de transcription CBP-p300 en permettant la fixation d'une protéine répresseur TLS sur le complexe de transcription <sup>224</sup>. Inversement, THRIL recrute la protéine hnRNPL au niveau du promoteur du gène du TNF $\alpha$  afin d'activer son expression suite à une infection par un pathogène (Figure 26B) <sup>225</sup>.

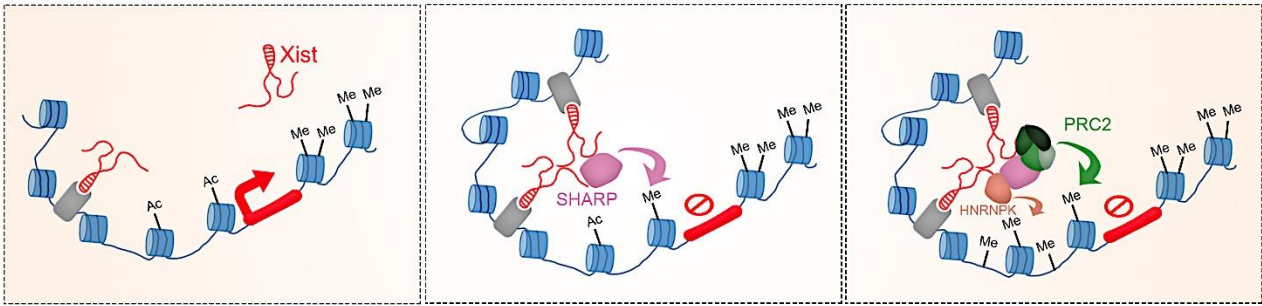
## *ii. Séquestration de protéines nucléaires*

D'autres lncARNs régulent l'activité transcriptionnelle des gènes mais via un mécanisme d'action différent puisqu'ils n'adressent pas les facteurs de transcription ou de remodelage chromatinien au niveau des séquences promotrices mais ils les séquestrent en jouant le rôle de "decoy" (Figure 25-f).

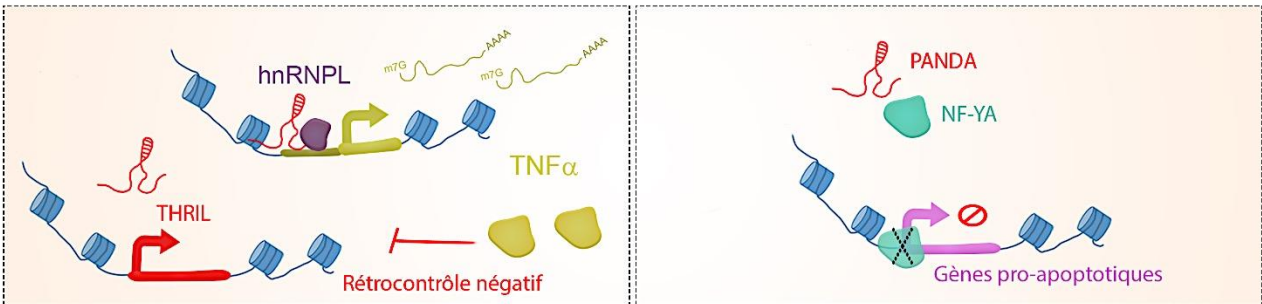
Ainsi, le lncARN PANDA (P21 Associated NcRNA DNA damage Activated) réprime l'apoptose des cellules cancéreuses par la séquestration du facteur de transcription NF- $\kappa$ B qui contrôle l'expression de gènes pro-apoptotiques tels que FAS, APAF1 ou BIK (Figure 26B) <sup>226</sup>.

Un autre exemple concerne le lncARN LET (Low Expressed in Tumour) réprimé par l'hypoxie. Il séquestre la protéine ILF3/NF-90 qui est impliquée dans la formation des granules de stress, des p-bodies et dans la translocation nucléo-cytoplasmique des ARNs <sup>227</sup>.

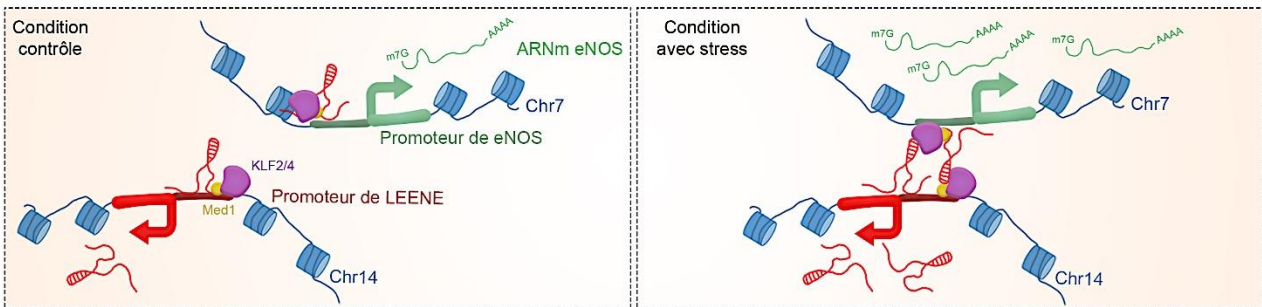
### A Inhibition des gènes du chromosome X par le lncARN Xist



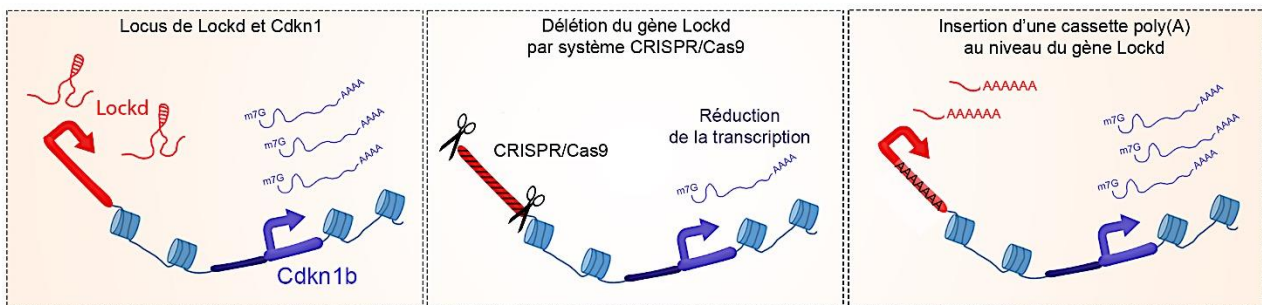
### B Adressage ou séquestration des facteurs de transcription par les lncARNs THRIL et PANDA



### C Remodelage chromatinien par le lncARN LEENE



### D Régulation de l'expression de Cdkn1 par le eRNA Lockd



**Figure 26:** Modes d'actions des lncARNs nucléaires Xist, THRIL, LEENE et Lockd sur la transcription des gènes codants.

Les lncARNs nucléaires agissent par différents mécanismes sur la transcription des gènes. **A.** Xist inhibe la transcription en remodelant l'organisation de la chromatine et en servant de plateforme d'échaffaudage aux complexes modificateurs de la chromatine (SHARP, hnRNPK et PRC2). **B.** THRIL guide hnRNPL sur le promoteur du TNF $\alpha$  pour favoriser la transcription et subit un rétrocontrôle négatif par la protéine produite tandis que PANDA séquestre le facteur de transcription NF-YA empêchant l'expression de gènes pro-apoptotiques. **C.** LEENE agit en condition de stress en rapprochant le promoteur du transcrit eNOS au niveau de son propre promoteur afin d'augmenter la transcription de ce gène. **D.** Lockd est un enhancer-RNA. Ses séquences promotrices agissent comme des enhancers pour le gène codant à proximité Cdkn1b. L'ajout d'une cassette poly(A) en aval du premier exon de Lockd n'influe pas sur la transcription de Cdkn1b.

### iii. Modifications spatiales de la chromatine

Outre leurs capacités de liaison à des RBPs, les lncARNs peuvent également se fixer à différentes séquences génomiques distantes afin de modifier l'organisation spatiale de la chromatine (Figure 25-b). C'est le cas du lncARN LEENE dont le mode d'action fut déterminé par ChIRP et RIP. LEENE agit par le biais des facteurs de transcription KLF2/4 et Med1. Il induit sa propre expression sur le chromosome 14 ainsi que celle de eNOS sur le chromosome 7, par appariement sur des séquences génomiques induisant un rapprochement des séquences promotrices (Figure 26C) <sup>228</sup>.

Cette réorganisation spatiale de la chromatine peut également se faire par la liaison d'un transcrit à différentes protéines associées à la chromatine. Ainsi, une autre facette du mode d'action de Xist consiste à se fixer sur des protéines SAFB/hnRNPU disposées à différents endroits sur le chromosome X et permettre un rapprochement de ces segments génomiques appelés SAR (Scaffold-Attached Region) provoquant une inhibition quasi complète du chromosome X <sup>203,229</sup>.

### iv. Les lncARNs Enhancers (eRNA)

Le mode d'action des eRNA dépend de la transcription active de leur gène et non pas de la molécule d'ARN synthétisée. Ces lncARNs dont la fonction est liée à l'activité de leur promoteur, c'est à dire à la fixation de protéines activatrices de la transcription sur leurs séquences régulatrices, sont appelés ARNs Enhancers (eRNA) (Figure 25-d). L'un des premiers exemples décrit dans la littérature est le lncARN Lockd (LncRNA downstream of Cdkn1b). La délétion complète du gène Lockd par le système d'édition génomique CRISPR/Cas9 inhibe l'expression du transcrit CDKN1. Toutefois, l'insertion d'une cassette de polyadénylation à l'intérieur du premier exon du transcrit Lockd altère à 90% l'expression du lncARN mais n'interfère pas avec l'expression du transcrit CDKN1. Ainsi, Il a été démontré que le gène Lockd exerce une régulation *en cis* sur le transcrit CDKN1 liée à une activité enhancer de sa région promotrice (Figure 26D) <sup>230</sup>.

Le lincRNA-Cox2 est un autre exemple de eRNA qui contrôle par un mécanisme semblable *en cis* l'expression du transcrit Ptgs2 (Cox2) impliqué dans la synthèse des prostaglandines <sup>231</sup>. Cependant, le lincRNA-Cox2 agit également par un mécanisme *en trans* lié cette fois à l'action de sa molécule d'ARN sur l'expression de gènes liés à la réponse immunitaire innée chez la souris <sup>232</sup>.

#### v. *Maturation co-transcriptionnelle des ARNm*

Certains lncARNs impactent la maturation des ARNm que ce soit au niveau de la maturation 5' de la coiffe, de l'épissage ou de la maturation 3' par polyadénylation (Figure 25-h). Un exemple de la littérature concerne le lncARN CCAT2 (Colon Cancer Associated Transcript 2) induit dans le cancer du côlon. CCAT2 régulent l'expression d'un variant d'épissage en se fixant au niveau des séquences poly(A) interne à l'intron 14 de l'ARNm de la Glutaminase (GLS) dans les cellules cancéreuses humaines. La fixation du transcrit permet le recrutement sur ce site du complexe d'épissage CFIm25/CFIm68. Ce complexe permet l'épissage alternatif de l'ARNm de la GLS et l'expression d'une isoforme fonctionnelle augmentant le métabolisme glutaminergique. L'inhibition du transcrit CCAT2 in vitro et in vivo altère le métabolisme, la migration et la prolifération des cellules cancéreuses.

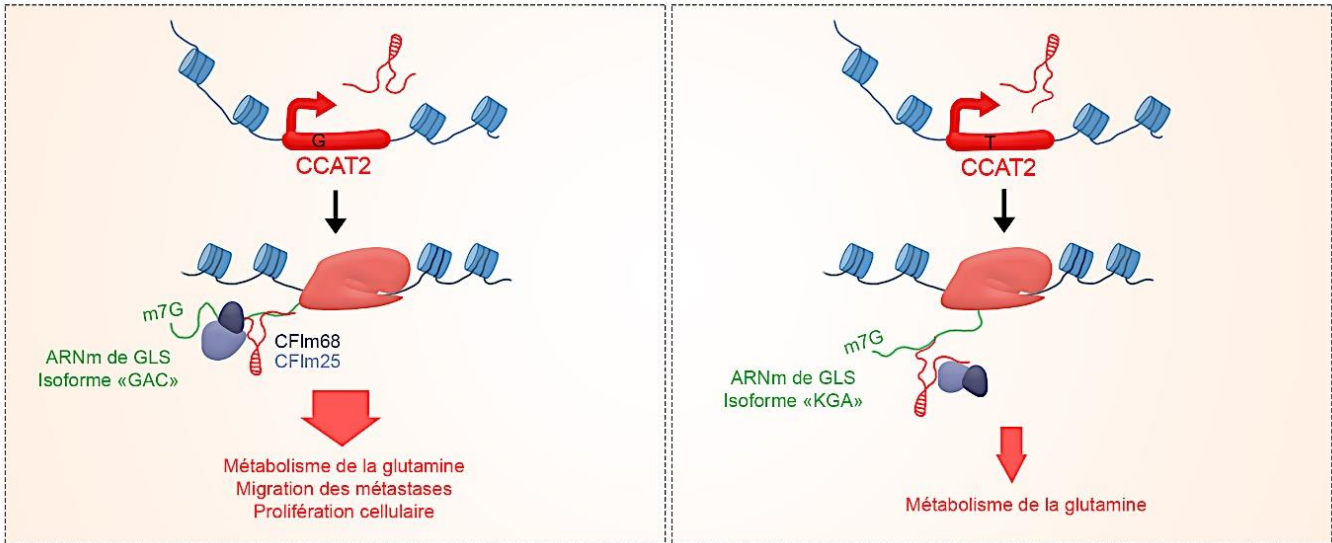
Un aspect intéressant du mode d'action de CCAT2 est qu'il possède différentes isoformes en fonction de l'allèle l'exprimant. En effet, il existe deux versions du gène CCAT2 qui diffèrent par le polymorphisme d'un nucléotide G/T (SNP, Single Nucleotide Polymorphism). Ce polymorphisme altère la structure secondaire du transcrit CCAT2 et modifie sa capacité d'interaction avec le complexe d'épissage CFIm25/CFIm68 et pourrait être la cause de sa surexpression dans les cancers du côlon (Figure 27A) <sup>233</sup>.

Une autre classe d'ARN est connue pour son action sur l'épissage alternatif : les ARNs circulaires (circARNs). Ils ont la particularité d'être co-transcrits avec des ARNm et peuvent être constitués uniquement de séquences introniques, de séquences exoniques ou d'un mélange de séquences introniques et exoniques (Figure 19). Le circARN circMbl est constitué de l'exon 2 et des deux introns flanquants de la protéine d'épissage muscleblind (MBL) (Figure 27B).

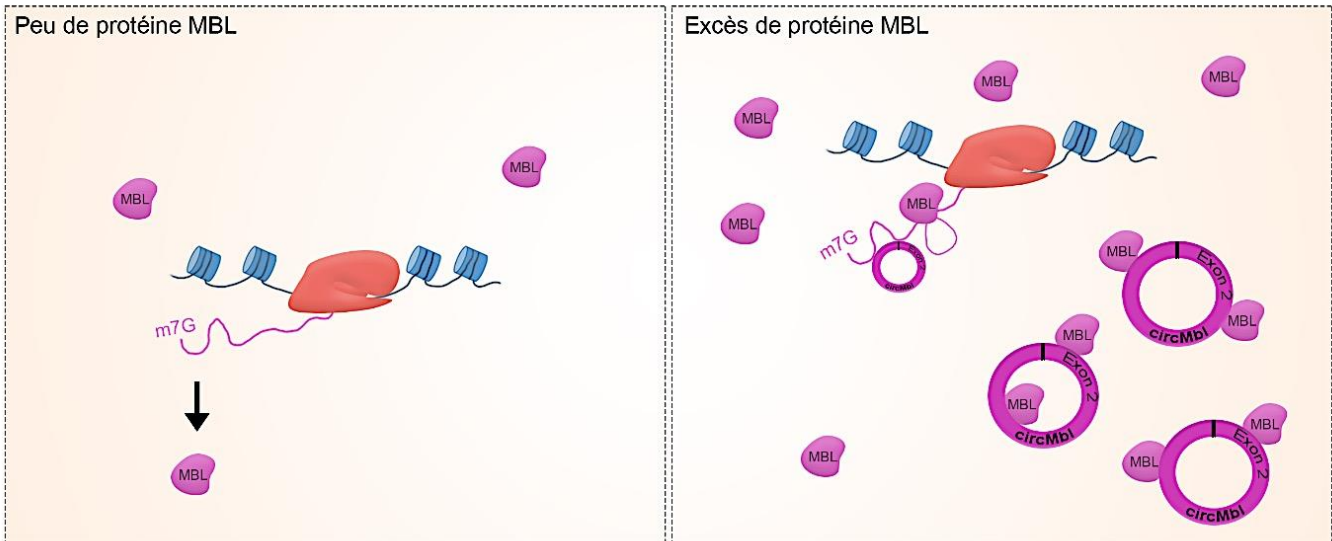
. Lorsque la protéine MBL est en excès, elle favorise l'expression du transcrit circMbl qui entre en compétition avec l'ARNm linéaire au niveau des sites d'épissage, empêchant la transcription de l'ARNm (Figure 25-h). Il s'agit d'un mécanisme de co-transcription compétitif. D'autre part, circMbl peut également séquestrer la protéine MBL dans le noyau pour limiter son action sur d'autres transcrits (Figure 25-i) <sup>234</sup>.



### A Epissage alternatif de l'ARNm de la GLS par le lncARN CCAT2



### B Compétition transcriptionnelle de l'ARNm de MBL et du circMbl



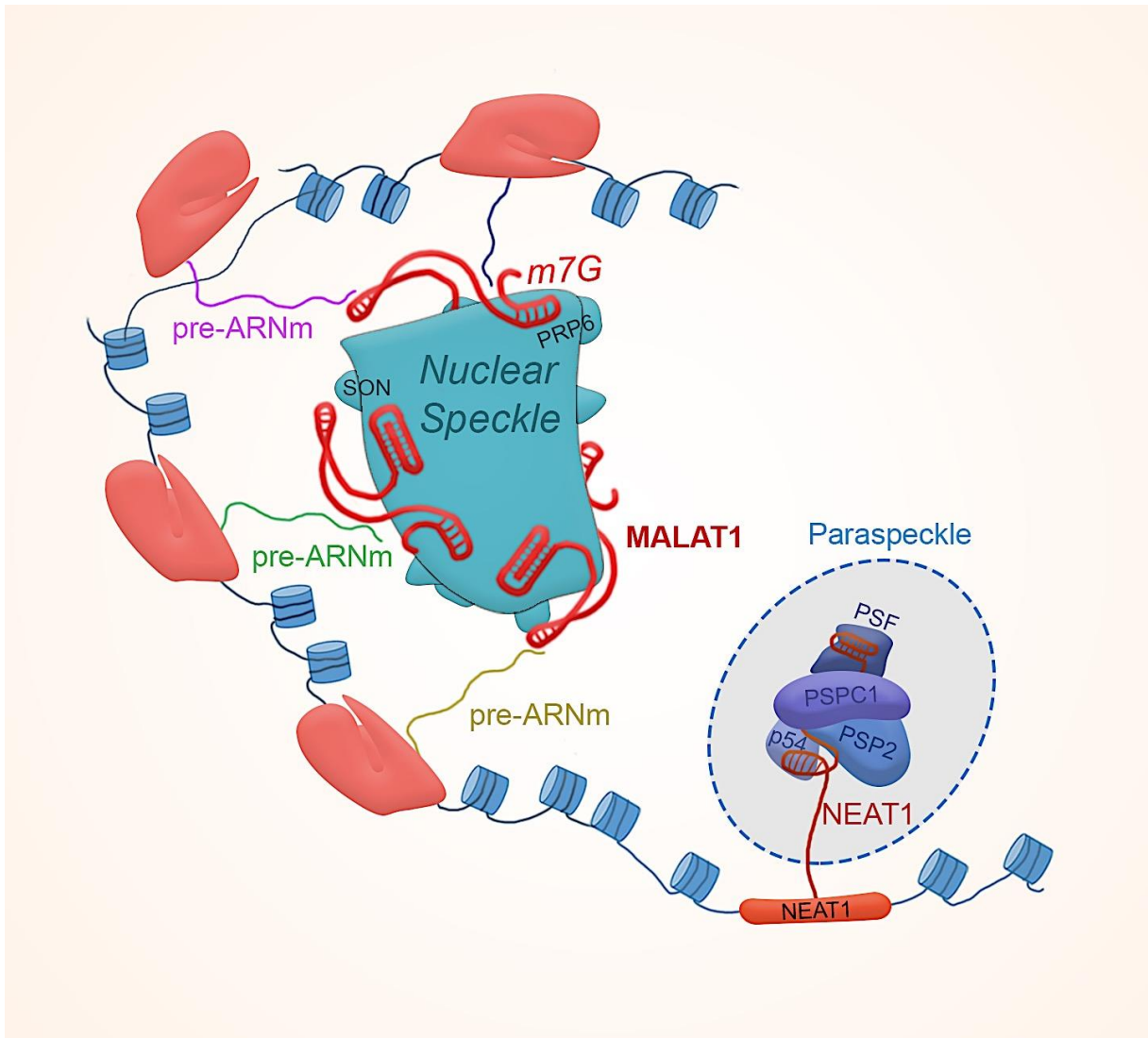
**Figure 27:** Modes d'actions des lncARNs CCAT2 et circMbl sur la maturation co-transcriptionnelle des ARNm. Les lncARNs nucléaires peuvent impacter la maturation des transcrits. **A.** Le lncARN CCAT2, en fonction de l'allèle qui l'exprime, impacte l'épissage de l'ARNm de la protéine GLS en guidant les facteurs d'épissage CFIm68/CFIm25. L'isoforme « GAC » de la GLS code pour une enzyme plus efficace dans le métabolisme de la glutamine et favorise ainsi la migration des métastases et la prolifération des cellules cancéreuses. **B.** Le circARN circMbl est induit par la protéine MBL par épissage alternatif et entre en compétition pour la transcription de l'ARNm de MBL. De plus, circMbl séquestre la protéine MBL pour bloquer son activité enzymatique.

#### vi. Les lncARNs structuraux ou architecturaux (arcARN)

Certains lncARNs sont impliqués dans la formation de sous-structures nucléaires ayant des fonctions de rétentions d'ARNm, de centres d'épissage ou de modifications des ARNs. Contrairement aux autres lncARNs, ces transcrits possèdent de forts niveaux d'expression (Figure 28). Ils sont de grandes tailles et sont riches en séquences répétées indiquant des conformations d'ARN complexes, essentielles à des fonctions d'échafaudages (Figure 25-g). Les deux exemples les plus connus de la littérature sont les lncARNs NEAT1 et MALAT1 (NEAT2). NEAT1 s'assemble à plus de soixante protéines, dont des RBPs mais également des facteurs de transcription, pour constituer des structures appelées paraspeckles. Les paraspeckles se forment concomitamment à la transcription de NEAT1 notamment en réponse à divers stress (thermique, hypoxie, infection virale, réplication induite par un oncogène...) et en absence de transcription de ce lncARN, ils apparaissent dissociés dans le noyau <sup>190</sup>. Ainsi des expériences de ChIRP ont mis en évidence qu'au cours de sa transcription, l'extrémité 3' de NEAT1 s'apparie par complémentarité de séquence à son gène, puis les structures secondaires de son extrémité 5' forment des repliements permettant la fixation de diverses protéines des paraspeckles telles que *PSPC1*, *PSP2*, *PSF* et *p54* <sup>235</sup>. NEAT1 modifierait l'expression génique par son rôle de séquestration d'ARNm et de RBPs régulatrices, et par des modifications des ARNm. Il serait également impliqué dans la réponse anti-virale <sup>236</sup>.

MALAT1 est impliqué dans la formation d'autres granules nucléaires, les nuclear speckles. Ces granules sont des centres de modifications et d'épissages des ARNm grâce aux protéines SR (Serine-Arginine-Rich splicing factors) telles que PRP6 et SON. Ils contiennent également de nombreux facteurs de transcription qui permettent d'activer la transcription à proximité de ces complexes <sup>237</sup>. MALAT1 participe mais n'est pas essentiel à la formation de ces granules <sup>238</sup>. MALAT1 servirait plutôt d'échafaudage pour favoriser les interactions entre les promoteurs des gènes, les facteurs de transcriptions, les ARNm en cours de synthèse et les protéines d'épissages <sup>239</sup>.





**Figure 28:** Rôles structuraux nucléaires de MALAT1 et NEAT1 dans la formation des granules d'épissage (nuclear speckles) et des paraspeckles.

### 3. Modes d'action des lncARNs cytosoliques

Les lncARNs cytosoliques agissent sur la stabilité d'autres molécules d'ARNs, sur la traduction des ARNm ou sur l'échafaudage et la signalisation de protéines.

#### *i. Modes d'actions impliquant des interactions avec des acides nucléiques*

##### *a. ceARN (competing endogenous ARN) : Rôle d'éponges à microARNs*

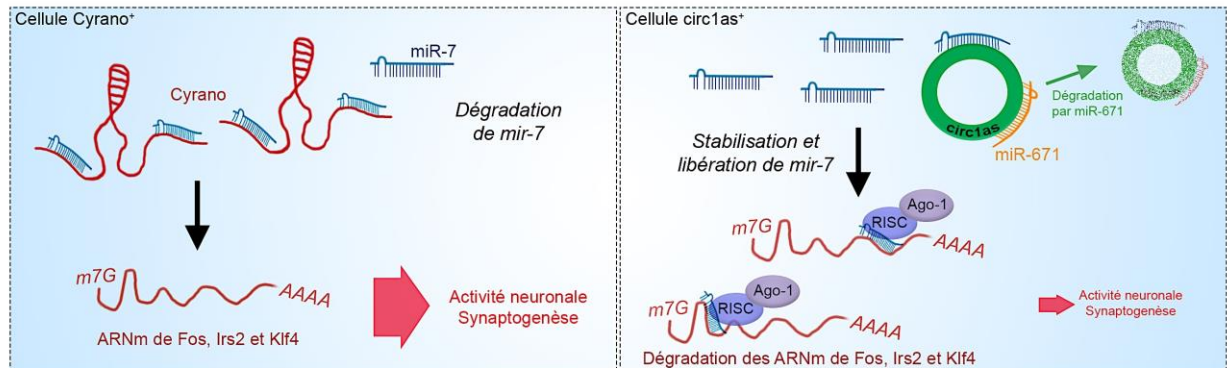
Les lncARNs peuvent se lier à d'autres familles d'ARNs tels que les microARNs par complémentarité de séquences (Figure 25-j). Un exemple de ce mode d'action concerne le lncARN Cyrano qui fixe miR-7 afin d'induire sa dégradation dans les neurones <sup>240</sup>. A l'inverse, la fixation de miR-7 sur le circARN CDR1as (CIRS-7) stabilise miR-7. En effet, la fixation de miR-7 sur CDR1as favorise la fixation de miR-671 qui dégrade efficacement le circARN libérant miR-7 qui va cibler des ARNm de gènes impliqués dans l'activité neuronale tels que Fos, Irs2 et Klf4 <sup>241</sup>. Ainsi l'expression de Cyrano favorise l'activité neuronale tandis que celle de CDR1as l'inhibe (Figure 29A). Dans les cellules cancéreuses, CDR1as a un rôle différent puisqu'il séquestre miR-7 sur ces 70 répétitions de séquences complémentaires, inhibant sa fonction répressive sur les ARNm de l'EGFR, d'IRS-1 et d'IRS-2, et favorisant ainsi la tumorigenèse <sup>242</sup>. Une sous-famille de lncARNs a été également associée à ce mode d'action, les pseudogènes. Les pseudogènes dérivent de séquences codantes ayant perdu leur potentiel codant au cours de l'évolution. Ils présentent des homologies de séquences avec les gènes codants dont ils sont issus, et agissent en tant que précurseurs d'ARNs interférents ou en tant qu'éponges à microARNs (ceARN) (Figure 25-j,-k ; Figure 29B) <sup>243-246</sup>. Un exemple bien connu de la littérature concerne le pseudogène PTENP1 qui séquestre les microARNs ciblant l'ARNm de PTEN. La protéine PTEN est un suppresseur de tumeur et par conséquent la séquestration des miR-17, miR-21, miR-214, miR-19 et miR-26 par PTENP1 permet de stabiliser son expression. Le pseudogène PTENP1 joue donc également un rôle de suppresseur de tumeurs (Figure 29B) <sup>247</sup>. Ce mode d'action est toutefois remis en question dû aux différences stoechiométriques entre les molécules d'ARN et de miARNs <sup>246</sup>.

##### *b. Contrôle de la stabilité des ARNm*

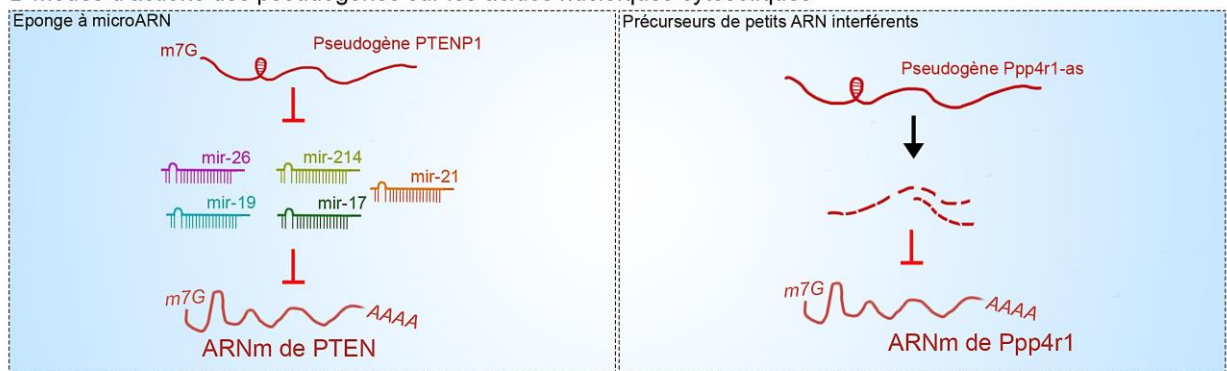
Les lncARNs exportés dans le cytoplasme peuvent s'apparier par complémentarité de séquences avec des ARNm et modifier leur stabilité en s'associant à des RBPs qui vont les stabiliser ou les dégrader (Figure 25-n). Le lncARN TINCR permet la fixation de la protéine Staufen1 (STAU1) au niveau d'interactions double-brins d'ARN, sur des séquences répétées de type Alu (AAAUAAA). La fixation de STAU1 induit la dégradation du dimère TINCR/ARNm de PGYLRP3 avec le concours des protéines UPF1 et UPF2 du système de dégradation NMD <sup>248</sup>. La dégradation des ARNm liée au recrutement de la protéine STAU1 est appelée STAU1-mediated RNA decay (SMD). Un autre exemple

du système SMD concerne la dégradation de l'ARNm de la SERPINE1 via la fixation du lncARN 1/2-sbsRNA1<sup>249</sup> (Figure 29C).

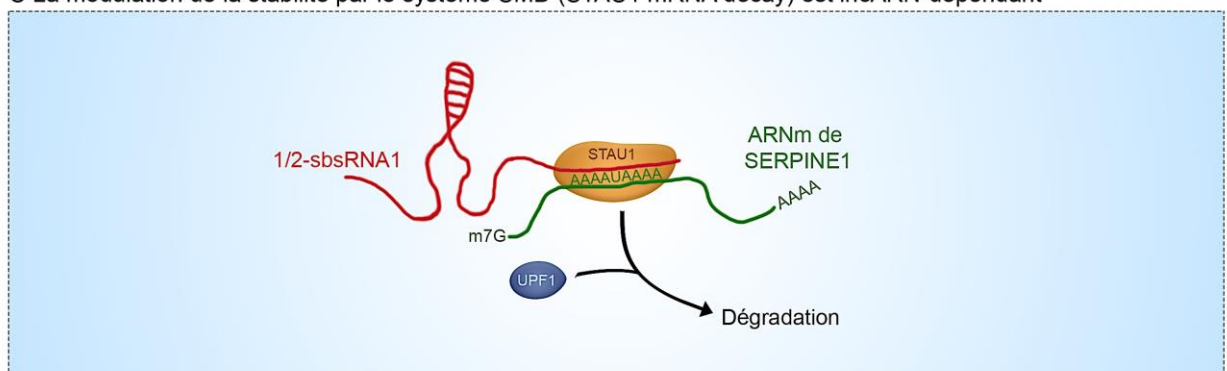
### A Régulation de l'activité neuronale via un mécanisme lncRNA-circRNA-microARN dépendant



### B Modes d'actions des pseudogènes sur les acides nucléiques cytosoliques



### C La modulation de la stabilité par le système SMD (STAU1 mRNA decay) est lncARN-dépendant



**Figure 29:** Mode d'actions de lncARNs interagissant avec d'autres transcrits cytosoliques.

Les lncARNs peuvent réguler l'expression génique en se liant par complémentarité de séquence à d'autres ARNs dans le cytoplasme. **A.** Ils peuvent se lier avec des microARNs, soit pour les séquestrer et les dégrader comme le lncARN Cyrano ou soit pour les stabiliser comme le circARN circ1as. **B.** Les membres de la sous-famille de lncARNs des pseudogènes ont des séquences très proches des ARNm et peuvent avoir un rôle d'éponges à microARN comme le pseudogène PTENP1. Ils peuvent également servir de précurseurs d'ARNs interférents comme le pseudogène PPP4r1-as. **C.** Enfin des lncARNs comme le lncARN 1/2-sbsRNA1, interviennent dans la stabilité des ARNm via le mécanisme de dégradation STAU1 mRNA decay (SMD).

## ii. Régulation de la traduction des ARNm

La traduction des ARNm en protéines est dépendante de la fixation des ribosomes et de différents co-facteurs. Les lncARNs en se fixant sur les transcrits permettent ou empêchent le recrutement de protéines impliquées dans la traduction.

### a. Inhibition de la traduction par les lncARNs

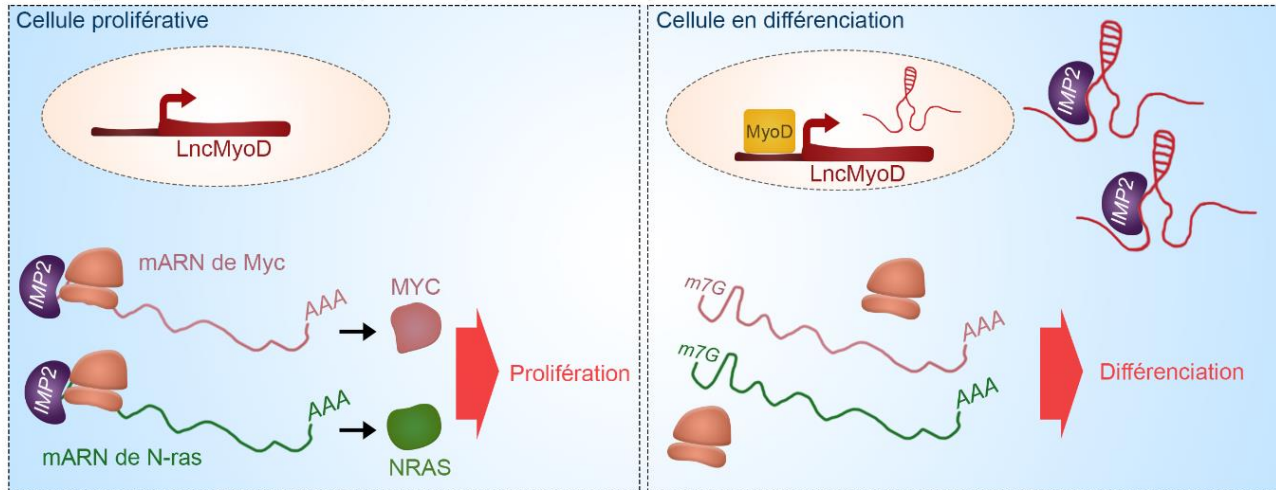
Les lncARNs peuvent séquestrer des co-facteurs de la traduction (Figure 25-l). C'est le cas de lncMyoD régulé par le facteur de transcription MyoD impliqué dans la myogenèse. Ce transcrit n'agit pas directement sur MyoD mais participe à la régulation d'autres protéines. lncMyoD recrute IMP2 (IGF2 mRNA-binding protein) qui favorise la traduction et le recrutement des ribosomes. Ainsi, le lncARN séquestre IMP2 et empêche la traduction de nombreux ARNm, tels que Myc et N-ras, durant les étapes de la différenciation musculaire (Figure 30A) <sup>250</sup>.

Un autre exemple de répression de la traduction concerne le lncARN PU.1 AS qui forme un hybride ARN-ARN avec l'ARNm de PU.1 empêchant la fixation du ribosome sur sa séquence et favorisant l'adipogénèse <sup>251</sup>.

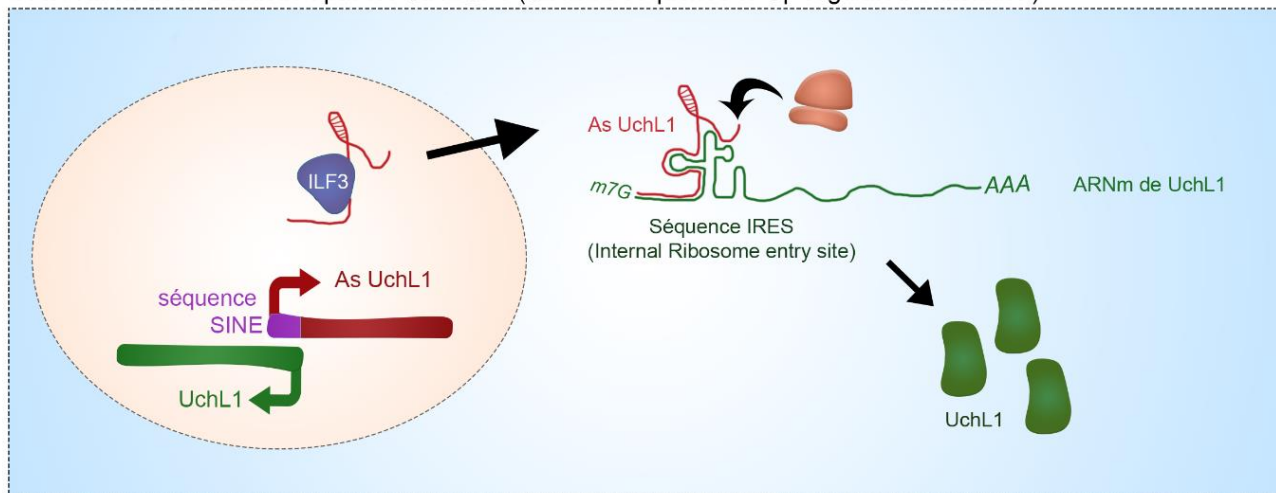
### b. Augmentation de la traduction par les lncARNs

Les lncARNs peuvent également avoir une fonction inverse en favorisant la traduction des ARNm (Figure 25-m). Ce mécanisme a été découvert par l'équipe de Stefano Gustincich qui a montré que la répression du lncARN antisens asUchl1 inhibe l'expression du transcrit Uchl1, impliqué dans la neurogenèse. AsUchl1 possède des séquences complémentaires au niveau du 5' du transcrit Uchl1. Ces séquences répétées correspondent à des éléments transposables appelés SINE B2. La surexpression d'asUchl1 ne modifie pas le niveau d'expression de l'ARNm Uchl1 mais augmente la quantité de protéine Uchl1 dans la cellule. Des expériences de délétions des régions SINE B2 ont démontrées qu'elles étaient indispensables à ce mécanisme post-transcriptionnel de l'asUchl1 sur le transcrit Uchl1. Les lncARNs antisens possédant ces séquences complémentaires en 5' ont été appelés SINEUPs (SINEB2 sequence to UP-regulate translation) <sup>170,171</sup>. Les SINEUPs sont transloqués dans le cytoplasme via le recrutement de la protéine ILF3 et se fixent au niveau des séquences IRES (Internal Ribosome Entry Sequence) de démarrage de la traduction. Ils favorisent ainsi la fixation du ribosome et la synthèse de protéines. Ce mode d'action est actuellement étudié pour ses applications thérapeutiques potentielles dans des pathologies liées à des défauts d'expression de protéines (Figure 30B) <sup>251</sup>.

A LncMyoD inhibe la traduction par la séquestration de la protéine IMP2



B Activation de la traduction par les SINEUPs (SINEB2 sequence to Up-regulated translation)



**Figure 30:** Régulation de la traduction par les lncARNs.

Les lncARNs modulent la traduction des ARNm en facilitant ou empêchant la fixation du ribosome sur le transcrit. **A.** LncMyoD est induit dans les cellules musculaires en différenciation et séquestre la protéine IMP2 qui a pour rôle de faciliter l'adressage du ribosome sur le transcrit à traduire. **B.** Les lncARNs antisens SINEUPs ont des séquences répétées de type SINE en 5' qui sont complémentaires de la séquence 5' du gène codant. Les SINEUPs sont transloqués dans le cytoplasme par la protéine ILF3 et facilitent la traduction des ARNm.

### *iii. Modifications post-traductionnelles, stabilité des protéines et signalisation cellulaire*

La fonction des lncARNs est donc intrinsèquement liée à leur capacité de recrutement de partenaires protéiques comme dans la modulation de la traduction et la stabilité des ARNm vu précédemment. Ces interactions protéine-ARN ont également d'autres rôles. Les lncARNs cytosoliques peuvent servir d'échafaudage et interagir exclusivement avec des protéines pour réguler leur stabilité, leur interaction avec d'autres protéines et leur signalisation (Figure 25-o).

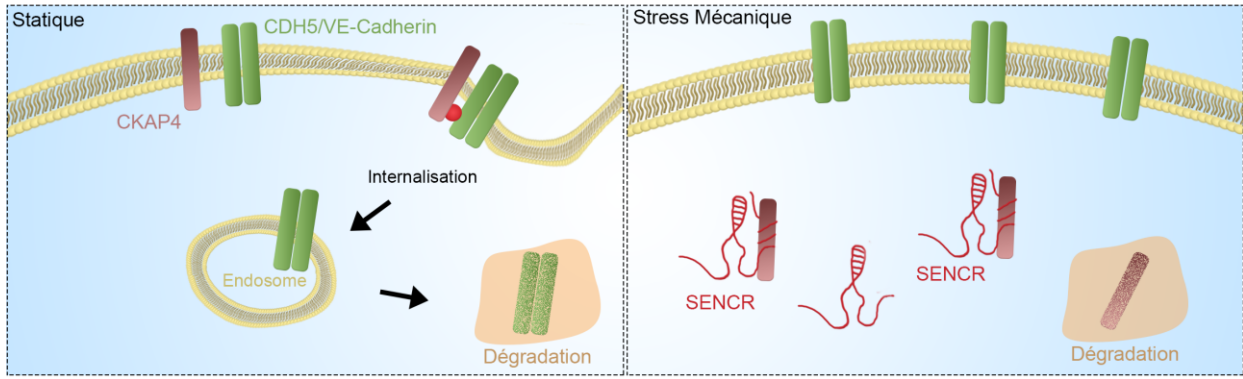
Dans les cellules endothéliales, le lncARN SENCER séquestre la protéine CKAP4 et induit sa dégradation suite à un stress mécanique, stabilisant la VE-Cadhérine/CDH5 et les contacts intercellulaires. Ainsi, en absence de SENCER, CKAP4 est exportée à la membrane ou elle induit l'endocytose de la VE-cadhérine (CDH5) et diminue l'adhésion entre les cellules et avec leur support (Figure 31A) <sup>253</sup>.

NORAD est un autre lncARN impactant la stabilité de sa protéine partenaire. Il induit la dégradation de la protéine PUMILIO impliquée dans la production de RLO et les anomalies chromosomiques <sup>254</sup>. Les souris déficientes pour Norad ont un vieillissement prématuré causé par l'hyperactivité de PUMILIO (Figure 31B).

Certains lncARNs impactent l'activité des protéines en empêchant la fixation de répresseurs protéiques. Par exemple, MetaLnc9/LINC00963 est induit dans les cancers et favorise l'apparition de métastases en empêchant la fixation d'une ubiquitine ligase sur la protéine PGK1 (Figure 31C) <sup>255</sup>. C'est également le cas du lnc-DC qui contrôle la différenciation des cellules de moelle osseuse en cellules dendritiques. Lnc-DC se fixe sur le facteur de transcription STAT3 et maintient la phosphorylation sur la tyrosine 705 en empêchant la fixation de la tyrosine phosphatase SHP1 (Figure 31C) <sup>256</sup>. Un autre exemple concerne le lncARN NKILA (NF-KappaB Interacting LncRNA) induit par le facteur de transcription NF-κB. La fixation de NKILA sur le complexe NF-κB/IκB (répresseur de NF-κB) bloque la phosphorylation de IκB par IKK et donc sa dissociation avec NF-κB qui est donc réprimé (Figure 31D). Ce lncARN tempère l'activation de NF-κB dans les cellules épithéliales mammaires soumises à une inflammation et inhibe la formation de métastases dans ces cancers mammaires <sup>257</sup>.



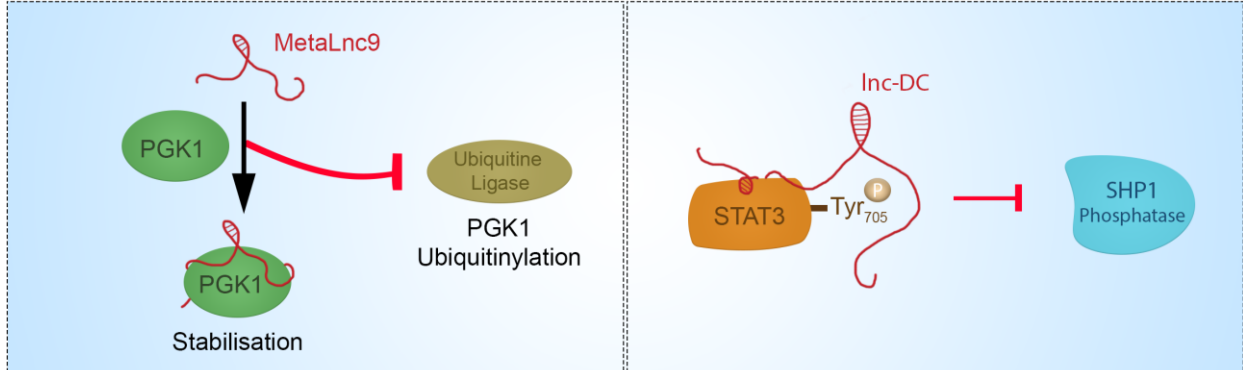
A Internalisation de la VE-cadhérine par la CKAP4 régulée par le lncARN SENCER



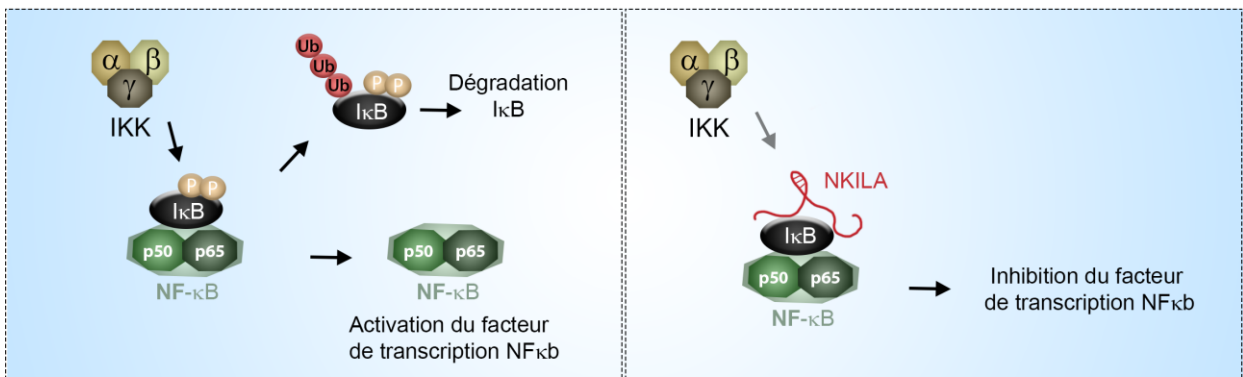
B Dégradation de la protéine PUMILIO par le lncARN NORAD



C Régulation des modifications post-traductionnelles des protéines par les lncARN MetaLnc-9 et lnc-DC



D Régulation de la voie de signalisation NF-κB par le lncARN NKILA



**Figure 31:** Modes d'action des lncARNs SENCER, NORAD, MetaLnc9, Lnc-DC et NKILA sur les modifications post-traductionnelles des protéines.

Les lncARNs via leur fixation sur les protéines peuvent moduler leur stabilité ou leur fonctionnalité. **A.** SENCER séquestre la protéine CKAP4 en condition de stress mécanique et empêche l'internalisation de la VE-cadhérine (CDH5). **B.** NORAD induit la dégradation de la protéine PUMILIO. **C.** MetaLnc9 empêche l'ubiquitinylation de la protéine PGK1 et favorise ainsi le métabolisme des cellules cancéreuses tandis que le lnc-DC empêche la fixation de la phosphatase SHP1 sur le facteur de transcription STAT3 qui reste actif. **D.** Enfin, NKILA empêche la phosphorylation de l'inhibiteur de NF-κB, IκB en limitant l'accès aux complexes kinases IKK.

#### iv. Formation des granules de stress et des p-bodies cytosoliques

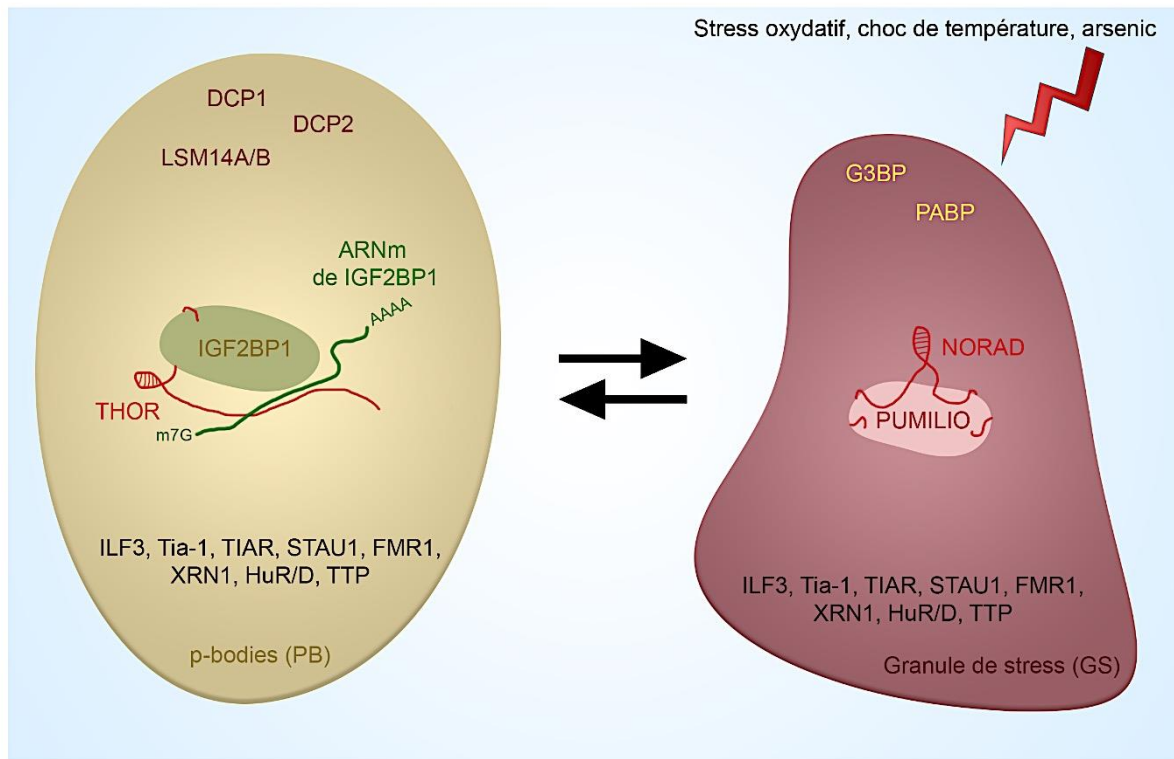
Les lncARNs cytosoliques peuvent également avoir un rôle structural dans la formation et le dynamisme des granules de stress, des vésicules d'endocytose et des p-bodies (processing bodies) (Figure 25-o). Le cytoplasme contient, outre les organites tels que l'appareil de Golgi, les mitochondries, le réticulum endoplasmique, les lysosomes et les endosomes, divers granules de stress (GS) et p-bodies (PB) correspondant à des agrégats denses d'acides nucléiques et de protéines (Figure 32) <sup>258</sup>.

Les PBs ont majoritairement une fonction de régulation du métabolisme des ARNs en les séquestrant et en favorisant leur interaction avec des protéines régulant la traduction, la dégradation et l'interférence ARN <sup>259</sup>. Les PBs sont constitués de protéines structurales comme LSM14A/LSM14B ou de protéines impliquées dans le transport telles qu'ILF3, ou dans la dégradation des ARNs comme STAU-1. Des ARNs appartenant à différentes classes entrent également dans la composition des PBs. Des analyses par hybridation *in situ* d'ARN à l'aide de sondes fluorescentes permettant une résolution à l'échelle d'une molécule unique (smRNA FISH) ont révélé que les miARNs non utilisés dans la cellule sont stockés au sein des PBs, ainsi que des ARNm dont la traduction est réprimée et certains lncARNs <sup>260</sup>. On observe d'ailleurs une anti-corrélation entre le potentiel traductionnel d'un ARN et sa localisation au sein des PBs <sup>259,260</sup>. Cette étude montre également que les lncARNs s'associent de façon très dynamique avec ces granules d'ARNs et se localisent préférentiellement à leur périphérie, favorisant leurs interactions avec différents constituants du cytoplasme <sup>260</sup>.

Le traitement de cellules avec des RNAses détruit l'assemblage des p-bodies, démontrant que les ARNs sont indispensables à ces structures <sup>261</sup>. De plus, la présence de différentes familles d'ARNs n'est pas anodine car les interactions existant entre ces ARNs et les protéines de ces structures impactent l'expression génique. Par exemple, le lncARN THOR s'associe à la protéine IGF2BP1 dans les PBs et stabilise l'ARNm de IGF2BP1 <sup>262</sup>.

Les granules de stress diffèrent sensiblement par la nature des protéines qu'ils contiennent. Les GS sont enrichis en protéines PABP et G3BP tandis que les PBs contiennent spécifiquement des protéines de clivage de la coiffe des ARNm, DCP1 et DCP2 <sup>258,263</sup>. Les molécules d'ARN stockées sont des complexes de pré-initiation de traduction bloqués en réponse à différents stress cytotoxiques (chaleur, stress oxydatif, arsenic). Les GS impactent différents processus physiologiques en modulant la stabilité et la traduction des ARNm tout comme les PBs. Norad est un exemple de lncARN qui affecte la stabilité d'une protéine au sein des GS. Ainsi, en réponse à un stress oxydatif, NORAD entraîne la dégradation de la protéine PUMILIO et réduit la production de RLO <sup>264,265</sup>. La différence majeure entre ces différents granules concerne leurs stimuli d'assemblage et leurs constituants <sup>266</sup>.

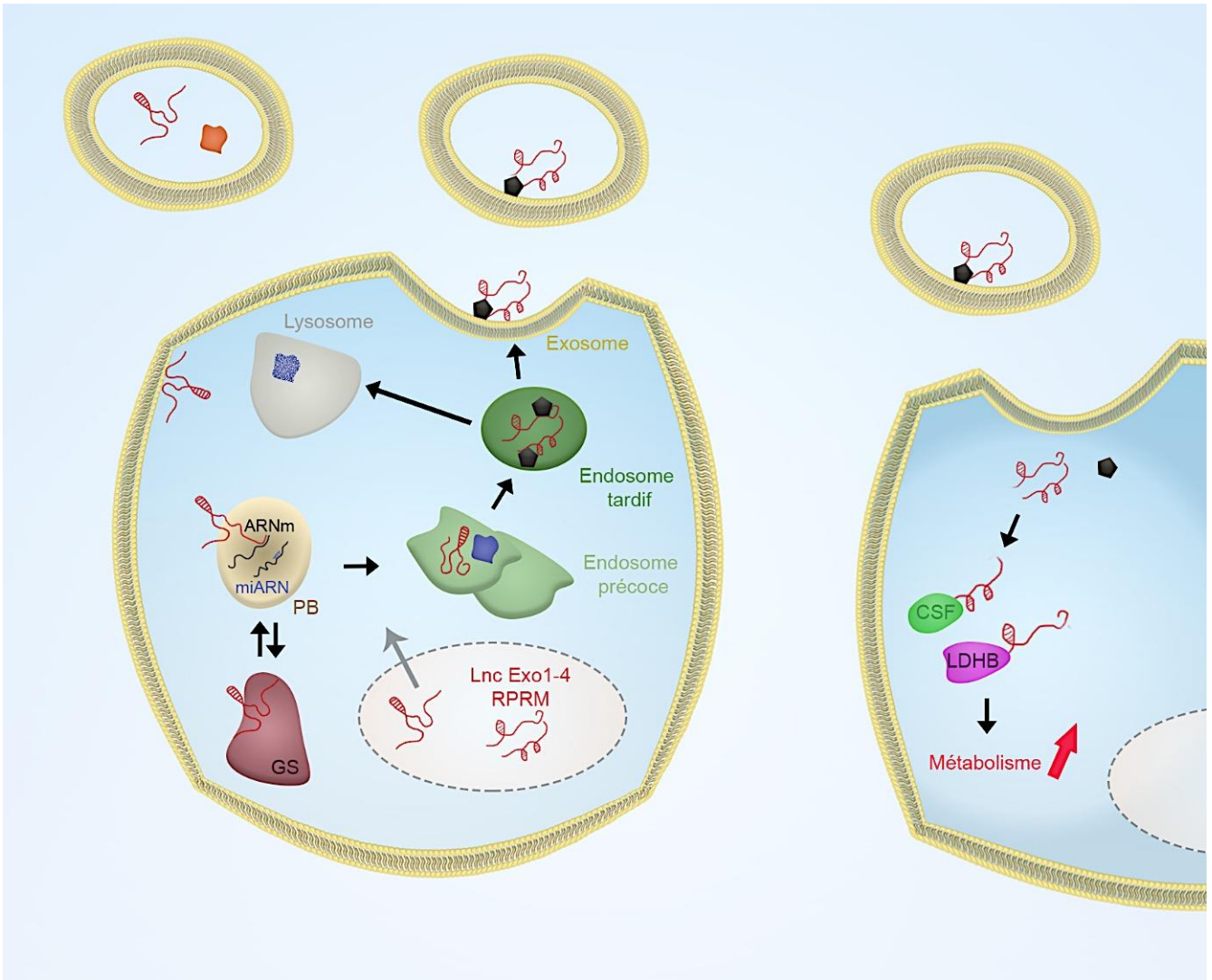




**Figure 32:** Interactions des lncARNs et de leurs partenaires protéiques au niveau des granules d'ARNs. Les granules d'ARNs du cytoplasme permettent de stabiliser les transcrits mais ils facilitent également la rencontre des lncARNs et de leur(s) partenaire(s) protéique(s). Les PBs sont présents en permanence dans les cellules et sont constitués de protéines spécifiques (en violet) tandis que les GS se forment en réponse à un stress exogène et possèdent également des protéines spécifiques (en jaune).

v. *Rôle des lncARNs dans la communication inter-cellulaire*

Les granules d'ARNs interagissent également avec les vésicules endosomales qui participent aux transports des ARNs et des protéines au sein de la cellule <sup>267</sup>. Les endosomes sont également impliqués dans la formation des exosomes, donc dans la sécrétion de protéines et d'ARNs qui participent au dialogue intercellulaire (Figure 25-p). Ainsi, les cellules cancéreuses de glioblastomes stimulent l'angiogenèse des cellules endothéliales voisines à l'aide d'ARNs contenus dans des microvésicules exosomales <sup>268</sup>. D'autres exemples, plus spécifiques aux lncARNs, existent dans la littérature <sup>269</sup>. Les lncARNs Gas5, BCYRN1, CCND1-ncRNA, RMRP et Lnc Exo1-4 ont été détectés dans des exosomes <sup>270,271</sup>. Plus précisément, Exo1-4 et RMRP sont sécrétés dans des exosomes puis capturés par endocytose par d'autres cellules dans lesquelles ils interagissent avec la cytokine CSF (Colony Stimulating Factor) et l'enzyme LHDB (Lactate déshydrogénase B) pour stimuler la viabilité cellulaire (Figure 33) <sup>271</sup>.



**Figure 33:** Trafic intracellulaire et sécrétions des lncARNs dans le compartiment extracellulaire.

## Chapitre 5 : Fonctions des longs ARNs non codants

Au cours de cette dernière décennie, de nombreuses études ont permis d'étayer considérablement nos connaissances sur la fonction des lncARNs dans la régulation de l'expression génique dans des contextes physiologiques mais également lors de leurs dérégulations dans diverses pathologies incluant les cancers.

### A. Rôles physiologiques des lncARNs

Les longs ARNs non codants sont pour la majorité faiblement exprimés dans les tissus ce qui complique leur exploration fonctionnelle. De récentes études révèlent cependant qu'ils exercent des fonctions très variées notamment au cours du développement embryonnaire et dans des processus de différenciation cellulaire (Figure 34).

#### 1. Rôles des lncARNs au cours du développement embryonnaire

Dans les cellules humaines, Xist est l'un des premiers lncARNs dont la fonction physiologique a été découverte <sup>272</sup>. Nous avons précédemment vu que Xist réprime les gènes du chromosome X et forme le corpuscule de Barr (cf. III.4.B.2) <sup>219</sup>. L'invalidation de *Xist* a démontré son rôle indispensable au développement des embryons femelles <sup>229</sup>. Plus précisément, Xist intervient dès les premières étapes de différenciation du trophoctoderme. Afin de mieux comprendre les effets de Xist sur la différenciation cellulaire durant l'embryogenèse, des chercheurs ont généré un KO conditionnel sous le contrôle du promoteur *Sox2*, induit lors de la gastrulation des embryons. A 8 jours de développement lors de la gastrulation, la délétion du gène n'impacte pas le développement puisque les embryons femelles arrivent à terme. Toutefois, ces souris femelles *Xist*<sup>-/-</sup> ont une létalité post-natale importante et un retard de croissance <sup>273</sup>. Le faible nombre de souris mutantes ayant survécu à l'âge adulte en absence de Xist, présente une compensation de dosage partielle des gènes du chromosome X. Le profil d'expression de Xist dans les différents types cellulaires est en mosaïque, c'est à dire qu'une partie des cellules expriment encore ce transcrit. Ces études démontrent l'importance de Xist au cours du développement. Nous avons également évoqué précédemment le rôle du lncARN H19 (cf. III.D) dans le développement embryonnaire et qui fut découvert parallèlement aux études sur Xist. Le lncARN H19 est exprimé sur l'allèle maternel et son expression réprime le gène *Igf2* (Insulin-like growth factor 2) en amont. Ainsi, l'invalidation de H19 induit la surexpression de la protéine Igf2, impliqué dans le métabolisme et la croissance cellulaire, et ainsi augmente la taille des embryons murins <sup>159,274</sup>.

D'autres lncARNs sont impliqués dans les mécanismes du développement embryonnaire. Les gènes *Hox* régulent la disposition spatiale de l'embryon. Ils sont segmentés en quatre clusters de gènes : *HoxA*, *HoxB*, *HoxC* et *HoxD*. Le lncARN HOTAIR, exprimé au niveau du cluster de gènes de

différenciation *HoxC*, favorise le développement des membres postérieurs des mammifères. HOTAIR inhibe les promoteurs du cluster de gènes *HoxD* via le recrutement des complexes polycomb (EZH2)<sup>221</sup>. D'autre part, le lncARN HOTTIP (HOXA Distal Transcript Antisense RNA) guide les remodeleurs chromatinien Wdr5/MLL (Mixed Lineage Leukemia) sur les promoteurs des gènes du cluster *HoxA*. Wdr5/MLL méthyle la lysine 4 de l'histone 3 et activent la transcription afin de favoriser le développement distal des membres<sup>275</sup>.

Des études plus récentes ont abordé à plus grande échelle l'étude fonctionnelle de lncARNs au cours du développement chez la souris. Les auteurs ont invalidé l'expression de 18 lncARNs candidats conservés chez l'homme (lnc-Brn1a/PANTR1, lnc-Brn1b/LINC01159, linc-Cox2, FabI/ELDR, linc-Enc1/LINC01157, Manr, Fendrr, Haunt, Hottip, Mdgt/HAGLR, Celr, Crnde, Spasm/KANTR, linc-Pint, linc-p21, linc-Ppara, Peril et Tug1) en remplaçant leurs gènes par des cassettes du gène *LacZ*, permettant la visualisation des profils d'expressions de ces transcrits *in vivo*<sup>276,277</sup>. Les résultats ont révélé que les délétions des lncARNs Peril, Mdgt et Fendrr altèrent le développement embryonnaire et entraînent de fortes mortalités péri- et post-natales. Peril est fortement exprimé dans les cellules embryonnaires cérébrales et faiblement dans les cellules différenciées. Sa délétion inhibe le développement cérébral et induit une létalité néonatale de près de 50%. Mdgt est majoritairement testiculaire mais également exprimé dans l'estomac, le colon, le thymus et le cerveau. Son invalidation induit un retard de croissance chez les souris femelles (60%) et mâles (32%) qui persistent à l'âge adulte. Enfin, le KO du transcrit pulmonaire Fendrr est létal car il altère l'organisation tissulaire du poumon, du cœur et du tractus gastro-intestinal. Deux autres candidats, linc-Pint exprimé dans le cerveau et les reins, et lnc-Brn1b dans le système nerveux, impactent à un niveau moindre la croissance embryonnaire<sup>277</sup>. L'invalidation constitutive des 13 autres lncARNs n'a aucun impact sur la survie ou le développement des embryons. Ces résultats illustrent la difficulté que représente la caractérisation fonctionnelle de cette famille de transcrits.

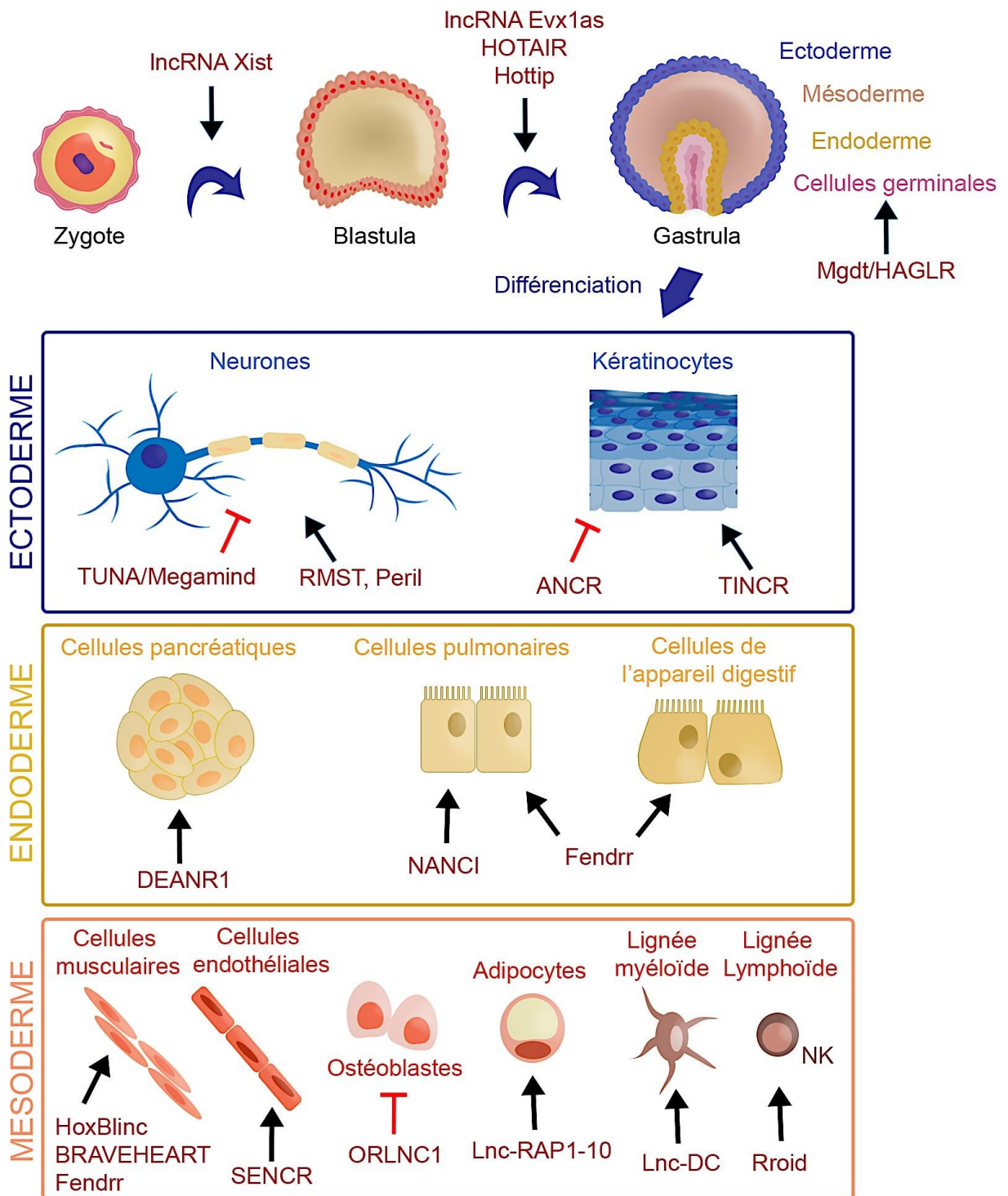
En accord avec ces observations, le phénotype des souris déficientes pour NEAT1, indispensable à la formation des paraspeckles, est peu marqué. Ces animaux sont viables, et seules les femelles présentent une baisse de fertilité par altération de l'ovulation, ainsi, que des difficultés à allaiter leur progéniture à cause d'anomalies d'organisation des glandes mammaires<sup>190,278,196</sup>.

## 2. Rôles des lncARNs dans des processus de différenciation cellulaire

Les lncARNs ont un rôle dans la différenciation des cellules souches en cellules différenciées (Figure 34)<sup>279</sup>. Les cellules souches pluripotentes du blastocyste se différencient lors de la gastrulation en ectoderme, endoderme et mésoderme, qui sont les cellules progénitrices de nombreux types cellulaires dans les tissus différenciés.

Les précurseurs des neurones et des kératinocytes de la peau sont d'origine ectodermique. Dans le cas de la différenciation neuronale, un crible fonctionnel a été effectué sur 1280 lncARNs différenciellement exprimés entre les neurones et les cellules progénitrices chez l'homme et la souris. Les auteurs ont réalisé une librairie d'ARNs interférents, constituée de 3 shARNs par lncARN, qu'ils ont transduit dans une lignée de cellules souches embryonnaires murines (mESc)<sup>280</sup>. Cette lignée exprime une protéine de fusion constituée du facteur de transcription Oct4, marqueur des cellules pluripotentes, couplé à la GFP. La librairie des différents shARN a été transduite à faible multiplicité d'infection afin de permettre l'inhibition d'un seul lncARN par cellule. Les cellules ont ensuite été clonées et le niveau de GFP déterminé par cytométrie en flux. Les clones ayant un faible niveau de GFP, traduisant leur différenciation, ont été séquencés pour identifier la construction shARN responsable de cet effet. Les résultats ont montré que seulement 20 lncARNs testés ont un impact sur la différenciation, ce qui traduit encore une fois la difficulté d'entreprendre des études fonctionnelles sur ces transcrits. Parmi ces candidats, TUNA/Megamind (Tcl1 Upstream Neuron-Associated lincRNA) est impliqué dans le maintien de la pluripotence. Il agit sur le remodelage chromatinien en interagissant avec les protéines hnRNPK et PTBP1 et inhibe la transcription des gènes de la différenciation *Nanog* et *Sox2*. RMST (Rhabdomyosarcoma 2-associated transcript) est un autre exemple qui guide le facteur de transcription SOX2 sur les promoteurs de gènes de la différenciation. Plus précisément, l'inactivation de RMST bloque la différenciation en neurones dopaminergiques<sup>281</sup>.

Au niveau épidermique, des analyses de données de RNA-seq de cellules progénitrices et de kératinocytes ont permis d'identifier 258 lncARNs différenciellement exprimés. Parmi ces transcrits, ANCR (anti-differentiation ncRNA) inhibe la différenciation des kératinocytes, tandis que TINCR (differentiation-induced ncRNA) la favorise<sup>248,282</sup>. En effet, l'inactivation d'ANCR induit une différenciation ectopique de l'épiderme visualisée par l'augmentation de l'expression des gènes *KRT1*, *KRT10* et *GHRL3*, marqueurs des couches kératinocytaires suprabasales différenciées de l'épiderme. A l'inverse, l'interférence de TINCR bloque la différenciation épidermique car ce lncARN module la stabilité des ARNm impliqués dans ce processus via le système SMD (cf. III.4.B.3).



**Figure 34:** Rôle de lncARNs dans le développement et la différenciation cellulaire. Les lncARNs sont impliqués dans les différentes étapes du développement embryonnaire que ce soit lors de la formation de la blastula (lncRNA Xist<sup>229</sup>), de la gastrula (Evx1as<sup>283</sup>, HOTAIR<sup>221</sup>, Hottip<sup>275</sup>) ou des différents feuillet embryonnaires (Mgd/HAGLR<sup>276</sup>). Les lncARNs sont également impliqués dans la différenciation cellulaire que ce soit au niveau des cellules neuronales (TUNA<sup>280</sup>, RMST<sup>281</sup>, Peril<sup>276</sup>), kératinocytes (ANCR<sup>282</sup>, TINCR<sup>248</sup>), pancréatiques (DEANR1<sup>283</sup>), pulmonaires (NANJI<sup>285</sup>, Fendrr<sup>276</sup>), digestives (Fendrr<sup>276</sup>), musculaires (HoxBlin<sup>286</sup>, BRAVEHEART<sup>287</sup>, Fendrr<sup>288</sup>), endothéliales (SENCR<sup>289</sup>), osseuses (ORLNC1<sup>289</sup>), adipocytaires (Lnc-RAP1-10<sup>291</sup>), myéloïde (Lnc-DC<sup>256</sup>) ou lymphoïdes (Rroid<sup>292</sup>).

Les cellules progénitrices d'origine endodermique se différencient en cellules pulmonaires, pancréatiques ou gastriques. DEANR1 (Definitive Endoderm-Associated lncRNA 1) est un lncARN spécifique de l'endoderme dont la répression par shARN inhibe la différenciation des cellules progénitrices en cellules pancréatiques<sup>284</sup>. DEANR1 favorise la fixation des facteurs de transcription SMAD2/3 sur le promoteur du facteur de transcription FOXA2 qui contrôle ce processus de différenciation<sup>284</sup>. NANCI (Nkx2.1-Associated Noncoding Intergenic RNA) est un autre lncARN qui régule la différenciation pulmonaire de l'endoderme par le biais de la voie Wnt<sup>285</sup>.

D'autres transcrits sont impliqués dans la différenciation de cellules d'origine mésodermique comme les cellules musculaires, adipocytaires, squelettiques, vasculaires, myéloïdes et lymphoïdes. Dans ce contexte, SENCR (Smooth muscle and endothelial cell enriched) favorise la différenciation des cellules pluripotentes en cellules endothéliales. En effet, la surexpression de SENCR réprime les marqueurs de pluripotence (Oct4 et Nanog), de l'ectoderme (Nestin) et de l'endoderme (FOXA2), et augmente ceux du mésoderme (MESP1, MIXL1 et BRACHURY) et même la VE-Cadhérine (CDH5) et PECAM-1 (CD31), marqueurs de cellules endothéliales, 7 jours plus tard<sup>253,289</sup>. De même, la diminution du lncARN PUNISHER par interférence ARN, altère le réseau vasculaire des embryons de poissons zèbres, démontrant son implication dans la différenciation et le fonctionnement des cellules endothéliales<sup>293</sup>.

Dans le cas des cellules immunes, le lncARN Rroid favorise la différenciation des cellules progénitrices de la lignée lymphoïde en cellules Natural Killer (NK) en guidant le facteur de transcription STAT5 sur le promoteur du gène *Id2* qui inhibe la différenciation des lymphocytes B et T<sup>292</sup>. Par ailleurs, au sein de la lignée myéloïde, une étude d'expression différentielle a mis en évidence 24 lncARNs induit au cours de la différenciation des cellules monocytaires en cellules dendritiques (DC). Parmi ces transcrits, l'inactivation de lnc-DC altère ce processus de différenciation par un mécanisme que nous avons précédemment décrit (cf. III.4.B.3.)<sup>256</sup>.

Un autre exemple de lncARN impliqué dans la différenciation mésodermique concerne hoxBlncl qui contrôle la formation des cellules hématopoïétiques et cardiaques en régulant les voies Notch et Wnt, et les gènes de la différenciation *Hox*. Afin de démontrer sa fonction, les auteurs ont dans un premier temps inhibé hoxBlncl par ARN interférence. Toutefois, le faible niveau d'expression des lncARNs et les effets partiels et non spécifiques des ARNs interférents rendent difficile l'interprétation du phénotype induit par cette approche<sup>294</sup>. Ils ont donc inhibé constitutivement l'expression du transcrit en excisant 2kb du gène en amont de l'extrémité 5' du transcrit à l'aide du système CRISPR/Cas9. Cette inactivation a mis en évidence le rôle indispensable de hoxBlncl dans la différenciation des



cellules mésodermiques par le remodelage de la chromatine et l'activation des gènes des voies Notch et Hox <sup>286</sup>.

Nous avons précédemment abordé la létalité embryonnaire de l'inactivation du lncARN Fendrr qui impacte la formation du poumon, du cœur et du tractus digestif. Le poumon et le tractus digestif ont une origine majoritairement endodermique mais le mode d'action de Fendrr dans la formation de ces tissus est encore indéterminé. Par contre, son implication dans le développement cardiaque est mieux documentée. Fendrr agit sur la différenciation du mésoderme latéral en guidant PRC2 au niveau des promoteurs de gènes de la pluripotence tels que *Foxf1* <sup>288</sup>. Braveheart (Bvht) et Alien sont deux autres lncARNs impliqués dans la différenciation des cellules cardiaques <sup>293</sup>. La diminution du niveau de Bvht par shARN inhibe ce processus <sup>287</sup>. Bvht agirait également en guidant le complexe PRC2 sur les promoteurs de gènes cibles encore inconnus.

Des études ont également été réalisées pour déterminer l'impact des lncARNs dans la différenciation adipocytaire. L'analyse de leur profil d'expression à différents stades de développement des adipocytes (pré-adipocyte, adipocyte primaire et adipocyte en culture) a identifié 175 lncARNs modulés au cours de ce processus. Des études fonctionnelles réalisées par perte de fonction sur 26 candidats ont démontré l'importance de 10 lncARNs dans la différenciation des adipocytes, nommés Inc-RAP1-10 pour "Regulated in Adipogenesis", et dont que les modes d'action ne sont pas encore déterminés <sup>291</sup>. Enfin, dans le cadre de la différenciation des cellules mésenchymateuses en ostéoblastes, l'inactivation du lncARN ORLNC1 bloque ce processus et favorise l'adipogenèse. Il serait induit par le facteur de transcription PTEN et agirait en séquestrant le microARN miR-296 <sup>290</sup>.

## B. Rôles pathologiques des lncARNs

Bien que peu exprimés et instables, les lncARNs sont fréquemment altérés et fonctionnellement impliqués dans différentes maladies dégénératives, inflammatoires, génétiques ou cancéreuses.

### 1. lncARNs impliqués dans des pathologies non tumorales

Dans le cadre des maladies neurodégénératives, nous prendrons l'exemple de la maladie d'Alzheimer caractérisée par des dépôts de protéines  $\beta$ -amyloïdes qui entraînent la formation de plaques bloquant la transmission synaptique. Chez une personne saine, la protéine BACE1 (B amyloid cleavage enhanced 1) permet de dégrader les plaques amyloïdes. BACE1-AS est un lncARN antisens co-régulé avec l'ARNm de BACE1 qui permet de stabiliser le transcrite et empêche la formation des plaques amyloïdes <sup>295</sup>. Chez les patients, l'absence de ce transcrite participerait à l'installation de la pathologie.

Dans le cadre des maladies inflammatoires, les lncRNAs sont exprimés dans des types cellulaires spécifiques (monocytes, macrophages, neutrophiles, cellules NK, lymphocytes ou cellules dendritiques). Par exemple, lincRNA-Cox2 est exprimé dans les macrophages et participe à la mise en place de la réponse inflammatoire, pouvant promouvoir certaines pathologies telles que l'artériosclérose <sup>296</sup>. Son invalidation par ARN interférent réprime l'expression de l'IL6 et augmente celle de la chimiokine RANTES/CCL5, tandis qu'inversement sa surexpression augmente l'expression de l'IL6 et diminue celle de CCL5. L'IL6 est normalement produite par les macrophages et permet l'amplification de la réponse inflammatoire en induisant la prolifération des neutrophiles et des lymphocytes. LincRNA-Cox2 induit également le gène *Ptgs2* impliqué dans la voie de synthèse des prostaglandines <sup>231,232</sup>. D'autres lncARNs sont également impliqués dans des maladies inflammatoires telles que le lncARN THRIL dont nous avons parlé précédemment (cf. III.4.B.2.) et dont l'expression est augmentée dans la maladie de Kawasaki, l'arthrite rhumatoïde et la maladie de Crohn <sup>225</sup>.

Parmi les maladies d'origine génétique, on peut citer l'exemple de la Progeria ou maladie de Hutchinson-Gilford caractérisée par un vieillissement accéléré. Cette pathologie rare est due à une mutation de la lamine A (LMNA) qui génère une protéine tronquée, la Progerine. En modifiant la structure chromatinienne, la progerine entraîne de nombreux dommages à l'ADN. Elle induit également l'expression de lncARNs, les dilncARNs, impliqués dans ces dommages et qui sont les précurseurs de petits ARNs régulateurs, les DDRNAs. Les DDRNAs, les dilncRNAs et des protéines de la machinerie de réparation telles que 53BP1 forment des plateformes d'échafaudage qui inhibent la maintenance des télomères et entraînent les cellules vers la sénescence <sup>297,298</sup>.

## 2. Les lncARNs impliqués dans les processus de tumorigenèse

Des études transcriptionnelles de différentes cohortes de CBNPC ont mis en évidence la dérégulation des lncARNs dans ces cancers. De plus, on a constaté que de nombreux SNPs corrélés au cancer sont présents au sein des loci de lncARNs. Les profils de méthylation des loci de lncARNs sont également différents entre les tissus sains et les tissus tumoraux, expliquant en partie la dérégulation de leur expression dans différents types de cancers <sup>299</sup>. Les lncARNs induits dans les cancers peuvent jouer le rôle d'oncogène en favorisant la tumorigenèse, et inversement des lncARNs réprimés dans les tumeurs peuvent être des suppresseurs de tumeurs. Les lncARNs régulent divers processus participant au développement tumoral tels que l'immortalisation, la prolifération, ou l'invasion cellulaires. Ils influencent également la résistance aux traitements et les effets du microenvironnement tumoral (Figure 35) <sup>300</sup>.

### *i. Régulation de l'immortalité cellulaire*

L'immortalisation des cellules cancéreuses c'est à dire leur capacité illimitée de multiplication est liée à l'inhibition de processus de différenciation et à la conservation des télomères. Des études ont démontré le rôle de certains lncARNs dans ces processus d'immortalisation. L'un des exemples le mieux caractérisé concerne TERRA, exprimé dans différents cancers, qui permet le maintien des télomères par divers modes d'actions en *cis*, en formant un hybride ADN-ARN permettant le recrutement de complexes épigénétiques, ou en *trans* par la fixation du facteur de transcription CTCF sur différents promoteurs. Ainsi la fixation de TERRA permet le maintien des télomères et inhibe le processus de vieillissement cellulaire et la sénescence <sup>301,302</sup>. D'autres lncARNs oncogéniques participent au renouvellement des cellules cancéreuses sans agir sur les télomères. C'est le cas du lncARN TCF7 qui active la voie de signalisation Wnt impliquée dans la régénération des cellules cancéreuses du foie en guidant des remodeleurs de la chromatine sur les promoteurs de ses gènes cibles <sup>302</sup>.

### *ii. Contrôle de la prolifération cellulaire*

La prolifération anarchique des cellules cancéreuses implique l'altération de nombreuses voies de signalisation comme les voies activées par les facteurs de croissance, ou celles régulant les points de contrôle (checkpoints) du cycle cellulaire. Dans ce contexte, il existe dans la littérature de nombreux exemples de lncARNs régulant la croissance cellulaire.

Le lncARN THOR est induit dans de nombreux cancers comprenant les CBNPC et les mélanomes. Des approches de gain et de perte de fonction *in vitro* et *in vivo* (siARN, CRISPR/Cas9)

ont démontré son rôle oncogénique sur la prolifération des cellules cancéreuses. En effet, nous avons précédemment vu qu'il agit sur la régulation de la protéine IGF2BP1 (Figure 32) <sup>262</sup>.

Un autre exemple concerne le lncARN nucléaire LUNAR1 (Leukemia-Associated Non-Coding IGF1R Activator RNA 1) induit par la voie Notch1 dans la leucémie aigüe lymphoblastique T. L'inhibition de LUNAR1 par des ARNs interférents diminue la prolifération des cellules leucémiques. LUNAR1 est un oncogène qui guide des facteurs de transcription sur le promoteur du gène du récepteur de l'IGF1 (*IGFR1*) <sup>304</sup>.

La voie du proto-oncogène Myc est également exacerbée dans de nombreux cancers. Le lncARN PCGEM1 (Prostate Cancer Gene Expression Marker 1) est, comme son nom l'indique, induit dans les cancers de la prostate. Son inhibition par shARN diminue l'expression de nombreuses enzymes contrôlant le métabolisme cellulaire tels que la LDHA, GLUT1 et PGK1, ce qui par voie de conséquences impacte la prolifération des cellules. Des expériences de co-précipitation ont montré que PCGEM1 est un lncARN oncogénique qui favorise la fixation du facteur de transcription c-Myc sur les promoteurs des gènes régulant le métabolisme <sup>305</sup>. D'autres exemples de lncARNs impliqués dans la signalisation de c-Myc concernent MYMLR dans les cancers pulmonaires et PVT1 (Plasmacytoma Variant Translocation 1) dans différents cancers. MYMLR est un lncARN divergent, qui agit comme un oncogène en permettant un repliement de l'ADN pour activer la transcription de c-Myc et donc la prolifération cellulaire <sup>306</sup>. PVT1 est induit par la signalisation de Myc. Il guide le complexe répresseur PRC2 sur les promoteurs des inhibiteurs de kinases dépendantes des cyclines *p15/INK4B/CDKN2B* (Cyclin-dépendant kinase inhibitor 2B) et *p16/INK4A/CDKN2A* qui sont des checkpoints bloquant la progression dans le cycle cellulaire et induisant la sénescence. PVT1 est donc également un lncARN oncogénique favorisant la prolifération cellulaire <sup>306,307</sup>.

ANRIL (Antisens noncoding RNA in INK4 Locus), qui promeut la prolifération des cellules leucémiques, est un autre exemple de transcrite agissant sur les points de contrôle du cycle cellulaire. ANRIL recrute PRC2 sur le promoteur de son antisens, le gène *p15/INK4B/CDKN2B* <sup>309-312</sup>.

### *iii. Invasion cellulaire et processus métastatique*

La transformation métastatique des cellules tumorales des tissus épithéliaux nécessite une transition épithélio-mésenchymateuse (TEM) pendant laquelle les cellules perdent des caractéristiques épithéliales et acquièrent des propriétés de cellules mésenchymateuses. La TEM nécessite l'activation de nombreuses voies de signalisation dont la voie du TGF $\beta$ . Des études ont identifié des centaines de lncARNs induits par le TGF $\beta$  <sup>313,314</sup>.

Parmi ces lncARNs, certains sont suppresseurs de tumeurs et répriment la TEM. Par exemple, le lncARN TGF $\beta$ -AS1 (TGF $\beta$  antisens 1) exprimé dans de multiples cancers réprime la voie TGF $\beta$ . Il

guide les complexes PRC2 sur les promoteurs de gènes répondant au TGF $\beta$  et empêche la fixation des facteurs SMAD2/3/4 activant la transcription de ces gènes <sup>314</sup>. Le lncARN BANCR (BRAF-Activated non-coding RNA) est un autre exemple de lncARN anti-métastatique qui inhibe la migration et l'invasion cellulaires dans les cancers de l'endomètre et les CBNPC. Il agit sur l'expression transcriptionnelle de gènes de la TEM en inhibant les transcrits de la cadhérine N et des métalloprotéases (MMP2) et en activant l'expression de la cadhérine E <sup>315,316</sup>.

D'autres lncARNs sont oncogéniques et favorisent la TEM tels que le lncRNA-HIT (HOXA transcript induced by TGF $\beta$ ), le lncRNA-ATB (activated by TGF- $\beta$ ), HOTAIR, MALAT1 et MetaLnc9 <sup>317</sup>. LncRNA-HIT est induit dans les CBNPC. Son rôle dans le contrôle de la migration et l'invasion des cellules cancéreuses a été mis en évidence par des approches de gain et de perte de fonctions. LncRNA-HIT stabilise le facteur de transcription ZEB1 (Zinc finger E-box-binding homeobox 1) qui favorise l'expression des gènes de la TEM <sup>318</sup>. De même, la répression de lncRNA-ATB inhibe la migration et l'invasion des cellules tumorales hépatiques. Ce transcrit agit par deux mécanismes, en permettant d'une part la séquestration et la dégradation du microARN miR-200 ce qui stabilise les protéines ZEB1/ZEB2 et favorise la TEM et d'autre part, en stabilisant l'ARNm de IL-11 qui active la voie STAT3 et favorise la colonisation des cellules cancéreuses <sup>319</sup>.

La découverte de MALAT1 est liée à son expression corrélée aux métastases pulmonaires. Son invalidation *in vivo*, à l'aide d'oligonucléotides antisens de type "gapmers" dégradant les ARN nucléaires, a démontré son implication dans la formation de métastases des cancers pulmonaires par un mécanisme de régulation de l'épissage des transcrits <sup>319,320</sup>.

Enfin, HOTAIR réprime l'expression de la cadhérine E en guidant le complexe PRC2 sur son promoteur dans différents cancers tandis que MetaLnc9/Lnc000963 stabilise la protéine PGK1 et l'intégration des signaux des voies MAPK/ERK facilitant la transformation des cellules métastatiques <sup>255,322,323</sup>.

#### *iv. Résistance aux traitements et dommage à l'ADN*

Certaines cellules tumorales développent des résistances aux différents traitements anticancéreux par des mécanismes complexes pouvant impliquer des lncARNs. Les lncARNs peuvent être oncogéniques et favoriser la résistance aux drogues telles que le lncARN UCA1 (Urothelial Cancer Associated 1) dans les cancers de la vessie qui active la voie de signalisation Wnt/béta caténine et favorise la résistance au cisplatine <sup>324</sup>. Le lncARN HOTTIP rend les cellules cancéreuses pancréatiques résistantes à la gemcitabine en activant le transcrit HOXA13 et la voie Wnt tandis que le lncARN SNHG12 active la voie MAPK dans les glioblastomes entraînant une résistance au temozoladime <sup>325,326</sup>.

A l'inverse, certains lncARNs sont suppresseurs de tumeurs et réprimés dans les cellules cancéreuses résistantes. Par exemple, TUG1 dans les CBPC réprime l'expression de la protéine LIMK2b impliquée dans les mécanismes pro-apoptotiques lors de stress génotoxiques en guidant le complexe PRC2 sur son promoteur <sup>327</sup>.

Certaines drogues chimiothérapeutiques, comme le cisplatine, induisent des dommages à l'ADN. En réponse à ces stress génotoxiques, la voie p53 est activée et bloque la prolifération des cellules pour permettre la réparation des altérations mineures de l'ADN par divers mécanismes <sup>328</sup>. Lors d'altérations trop importantes, p53 active les voies de signalisations pro-apoptotiques entraînant la mort cellulaire. Cependant, nous avons vu que cette voie est fréquemment altérée dans différents cancers (cf.II.A.3). Plusieurs études ont démontré un lien entre cette voie de signalisation et la fonction de divers lncARNs. Ainsi, le lncARN PANDA (p21 Associated NcRNA DNA damage Activated) est activé par la voie p53 dans les fibroblastes et dans les CBNPC. PANDA est un lncARN oncogénique qui séquestre le facteur de transcription pro-apoptotique NF-YA et inhibe l'apoptose des cellules lors de stress génotoxiques <sup>226</sup>.

D'autres exemples impliquent plutôt des lncARNs suppresseurs de tumeurs. Par exemple, lincRNA-p21 est exprimé dans les lymphomes, les sarcomes et les fibroblastes soumis à des agents génotoxiques <sup>329</sup>. lincRNA-p21 est induit par la voie p53 et médie la répression transcriptionnelle p53-dépendante en guidant la protéine hnRPNK sur les promoteurs des gènes cibles entraînant l'apoptose des cellules cancéreuses. Par ailleurs, des expériences de KO conditionnels chez la souris ont montré que lincRNA-p21 agit en *cis* pour activer l'expression de son gène voisin p21Cip1/CDKN1A et contrôler le checkpoint G1/S <sup>330</sup>. D'autres exemples concernent le lncRNA MEG3 (Maternally expressed gene 3), LED (LncRNA activator of Enhancer Domain) et linc-Pint. MEG3 est réprimé dans les CBNPC et stabilise la protéine p53 en dégradant son inhibiteur MDM2, tandis que LED est induit par p53 dans de nombreux cancers et favorise sa fixation sur les enhanceurs de ses gènes cibles <sup>330,331</sup>. linc-Pint est induit par p53 dans les cancers colorectaux et participe à la réponse p53 en guidant le complexe PRC2 sur les promoteurs des gènes anti-apoptotiques <sup>333</sup>. Enfin, le lncARN Gas5 (Growth Arrest Specific 5), réprimé dans de nombreux cancers, active la voie de signalisation p53 et est également impliqué dans l'apoptose des cellules cancéreuses <sup>333</sup>.

### 3. Les lncARNs et le microenvironnement tumoral

Nous avons vu précédemment que le microenvironnement tumoral participe au développement de l'agressivité des tumeurs (cf. II.B). L'influence de ce microenvironnement peut mettre en jeu des lncARNs qui sont i) exprimés par les cellules cancéreuses et modulés par le stroma, ou ii) exprimés par les différents types cellulaires du stroma tumoral.

#### *i. Les hypoxaLinc, des lncARNs modulés par l'hypoxie*

Comme nous l'avons précédemment exposé (cf. III.2.B.4), l'environnement hypoxique est un facteur d'agressivité des cellules tumorales. Parallèlement aux gènes codants, les facteurs de transcription activés par l'hypoxie (HIFs, NF- $\kappa$ B, NRF2) induisent l'expression de divers ARNs non codants dont les microARNs. Au sein de notre équipe, nous avons mis en évidence l'implication de l'hypoxamiR miR-210 dans les mécanismes du switch métabolique et dans la radiorésistance des cellules cancéreuses pulmonaires en hypoxie (Figure 13)<sup>154,155</sup>. D'autres familles d'ARNs non codants sont également modulés par l'hypoxie, tels que les lncARNs dont NEAT1 qui est induit par le facteur de transcription HIF2 dans le cancer du sein<sup>335,336</sup>.

Il existe quelques exemples de lncARNs pro-oncogéniques induits par l'hypoxie qui agissent sur divers aspects de la tumorigenèse. Par exemple, le lncARN UCA1, dont nous avons déjà parlé, est induit par le facteur de transcription HIF1 et favorise la viabilité des cellules cancéreuses de la vessie<sup>337</sup>. H19 est également induit par HIF1 dans différents cancers dont les glioblastomes. Il stabilise l'expression de l'ARNm de la  $\beta$ -caténine et favorise la migration et l'invasion des cellules tumorales<sup>338</sup>. HINCUT1 active la prolifération des cellules cancéreuses du colon et impacte la maturation post-traductionnelle de protéines dans un contexte de stress cellulaire<sup>339</sup>. On peut encore citer, le lncARN EFNA3 induit par l'hypoxie dans différents cancers et dans les cellules endothéliales, dont la surexpression induit l'angiogenèse et la formation de métastases en stabilisant la protéine EFNA3 (Ephrin A3) par un mécanisme encore inconnu<sup>340</sup>.

Certains lncARNs agissent directement sur la stabilité du facteur de transcription HIF1. Par exemple, LINK-A est induit dans les cancers du sein triple-négatifs et son expression augmente l'expression de HIF1 $\alpha$  en normoxie. En réponse à HB-EGF, LINK-A sert de plateforme d'ancrage pour deux protéines kinases BRK (Breast Tumor Kinase) et LRRK2 (Leucine Rich Repeat Kinase 2) qui stabilisent HIF1 $\alpha$  par phosphorylation sur Tyr565 et Ser797, ce qui bloque son hydroxylation et activent son programme transcriptionnel<sup>341</sup>. Le lncARN-ROR induit par HIF1 en hypoxie dans les cancers du foie stabilise l'ARNm de HIF1 en séquestrant miR-145 tandis que LincRNA-p21 stabilise la protéine

HIF1 $\alpha$  en limitant son interaction avec l'ubiquitine ligase pVHL dans les cancers du poumon, de l'utérus et du sein <sup>342,343</sup>.

Enfin, certains lncARNs modulés par l'hypoxie sont des suppresseurs de tumeurs, c'est le cas du lncARN-LET (Low Expressed in Tumor) réprimé dans de nombreux cancers dont les cancers pulmonaires. L'hypoxie induit l'expression de HDAC3 qui désacétyle et réprime le promoteur de lncARN-LET favorisant la formation de métastases <sup>227</sup>. Un autre exemple concerne HITT (HIF1 $\alpha$  Inhibitor at translation level) réprimé par l'hypoxie via miR-205 dans les cancers du côlon et de l'utérus. En normoxie, HITT interfère avec la traduction de HIF1 $\alpha$  en séquestrant le régulateur traductionnel YB-1 <sup>344</sup>.

## *ii. Rôle des lncARNs dans le dialogue avec l'infiltrat immun*

L'infiltration de cellules inflammatoires au sein des tumeurs participe à la réponse anti-tumorale mais peut également être bénéfique aux cellules cancéreuses en favorisant leur échappement notamment via la sécrétion de médiateurs solubles. Parmi les molécules anti-tumorales, les cytokines pro-inflammatoires IL-1 $\beta$  et IL-6 (Interleukine 1 et 6) et la voie de signalisation STAT3 sont réprimées par les lncARNs linc-Cox2, DILC et TLSNC8 <sup>231,345,346</sup>.

La voie NF- $\kappa$ B est également impliquée dans les réponses inflammatoire et hypoxique. Elle est contrôlée par différents lncARNs tels que NKILA qui exerce un rétrocontrôle négatif en empêchant la phosphorylation et la dégradation de l'inhibiteur I $\kappa$ B <sup>257,347</sup>. Le lncARN Lethe, également induit via NF- $\kappa$ B par le TNF $\alpha$  et l'IL1 $\beta$ , régule négativement la voie NF- $\kappa$ B en empêchant la fixation du facteur de transcription sur les séquences promotrices de ses gènes cibles <sup>348</sup>.

La réponse immunitaire anti-tumorale est liée à l'infiltration de cellules immunitaires et plus particulièrement à l'infiltration des lymphocytes T cytotoxiques (LCT) CD8+. Les LCT sont régulés par différentes cytokines dont l'IL-2 sécrétée par les lymphocytes T régulateurs (Tregs). Dans le cas du cancer du côlon, le lnc-sox5 est un oncogène qui promeut l'expression de l'indoleamine 2,3-dioxygénase 1 (IDO1). IDO1 favorise la prolifération et la différenciation des LT en Tregs, ce qui a une action oncogénique en limitant l'infiltration des LCT <sup>348</sup>. A l'inverse, lnc-INSR stabilise le récepteur à l'insuline (INSR) qui favorise la prolifération des cellules immunitaires et la transformation des Treg en LCT dans les leucémies <sup>350</sup>.

L'échappement des cellules cancéreuses au système immunitaire dépend de l'expression de récepteurs de surface tels que le récepteur PD-L1/CD274. Dans ce contexte, le lncARN NKX2-1-AS1 est un lncARN suppresseur de tumeur dans les cancers pulmonaires qui inhibe l'expression de PD-L1 en empêchant la fixation du facteur de transcription NKX2-1 sur son promoteur <sup>351</sup>.

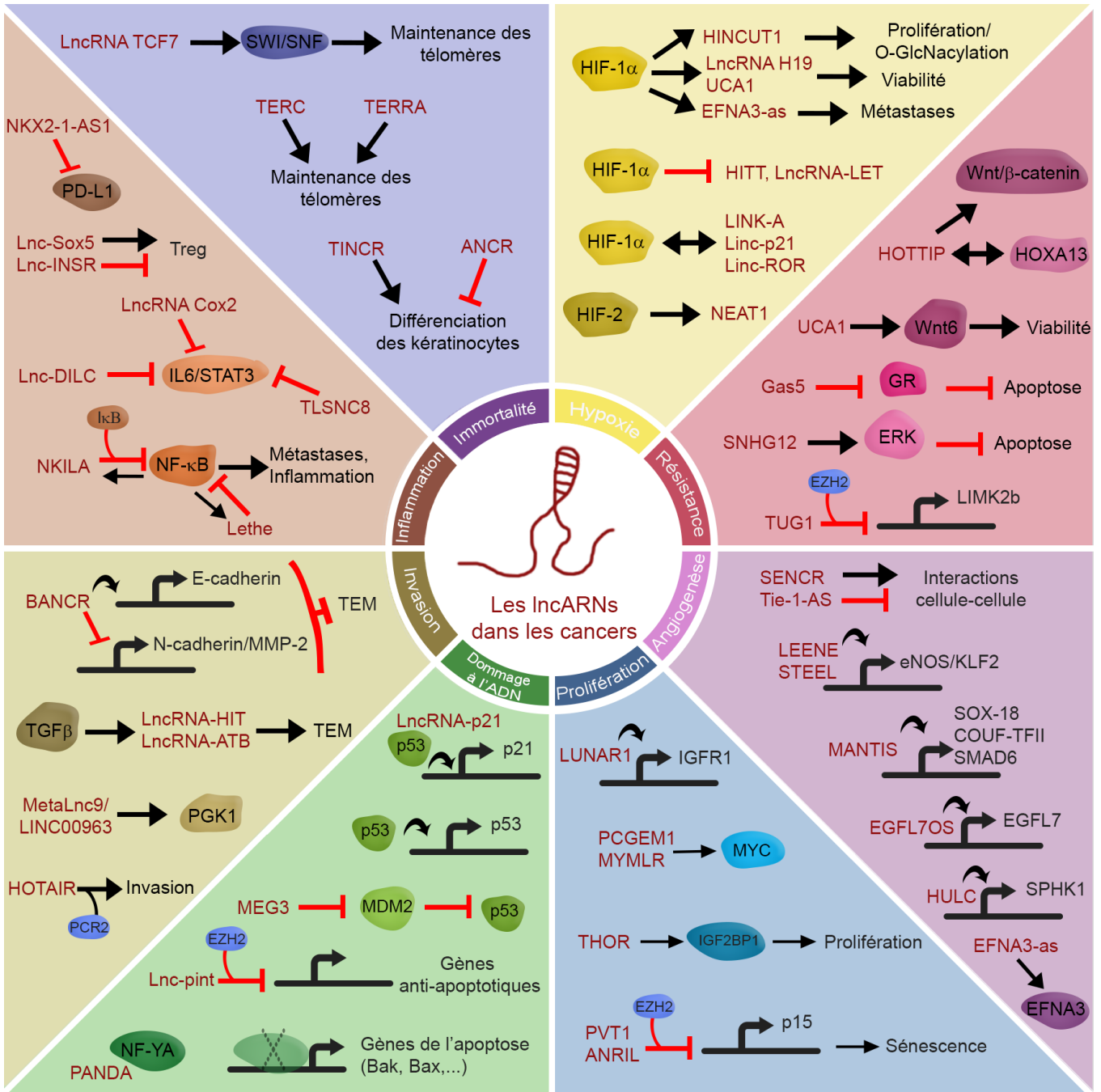


### iii. Les lncARNs favorisant l'angiogénèse

Le recrutement des cellules endothéliales et la formation de nouveaux vaisseaux sanguins au sein des tumeurs favorisent les apports d'oxygène et de nutriments et participent indirectement au développement tumoral. Ce phénomène d'angiogénèse tumorale est lié aux voies de signalisation induites par des facteurs de croissance comme le VEGF qui active les facteurs de transcription FOXO1, KLF2 et KLF4. Le lncARN HULC (Highly upregulated in liver cancer) est fortement exprimés dans les cellules cancéreuses du foie. Il guide le facteur de transcription E2F1 au niveau du promoteur de la Sphingosine kinase 1 (SPHK1) pour induire son expression et permettre ainsi la transmission des signaux du récepteur au VEGF <sup>352</sup>.

D'autres lncARNs impliqués dans l'angiogénèse sont exprimés dans les cellules endothéliales et sont corrélés à un mauvais pronostic chez les patients atteints d'un cancer. C'est le cas de MANTIS qui promeut l'angiogénèse dans les cancers pulmonaires en guidant le complexe épigénétique activateur de la transcription BRG1 sur le promoteur des gènes *SOX18*, *SMAD6* et *COUP-TFII* <sup>353</sup>. Par ailleurs, le lncARN enhancer LEENE (*lncRNA* that enhances eNOS expression) et le lncARN STEEL (spliced-transcript endothelial-enriched lncRNA) induisent l'expression du transcrit eNOS (endothelial Nitric Oxide Synthase) dans les cellules endothéliales ce qui stimule leur prolifération et leur migration <sup>228,354</sup>. Les lncARNs PUNISHER et EGFL7OS régulent les mêmes processus. Le mode d'action de PUNISHER n'est pas encore déterminé mais il est indispensable à la différenciation des cellules endothéliales et au maintien de leur prolifération <sup>293</sup>. EGFL7OS guide le facteur de transcription MAX sur le promoteur du gène voisin sur le brin opposé, *EGFL7* <sup>355</sup>. Tous ces lncARNs endothéliaux ont une action pro-tumorale car ils favorisent les mécanismes d'angiogénèse.

D'autres lncARNs modulent l'adhésion des cellules endothéliales à leur support ou leur contact cellule-cellule. La diminution de ces interactions bloque les mécanismes de néo-vascularisation et d'angiogénèse. Par exemple, le lncARN Tie1-AS se fixe sur l'ARNm de Tie1 afin d'induire sa dégradation. Tie1 étant indispensable à la formation de contact cellule-cellule, la surexpression de Tie1-AS entraîne des défauts de néovascularisation <sup>355</sup>. Tie1-AS est donc considéré comme un lncARN suppresseur de tumeur. Le contre-exemple concerne le lncARN oncogénique SENCER qui stabilise la VE-cadhérine et les interactions cellule-cellule en réponse à un stress mécanique (Cf : III. 4.B.3) <sup>253</sup>.



**Figure 35:** LncARNs impliqués dans les processus cellulaires de la tumorigenèse.

#### 4. Les lncARNs impliqués dans les cancers broncho-pulmonaires

Nous avons abordé l'implication de nombreux lncARNs dans la mise en place de l'agressivité des cancers. Certains de ces transcrits ont une fonction particulièrement importante dans les cancers broncho-pulmonaires et leur spécificité tissulaire et leur dérégulation dans ces tumeurs permet de distinguer les différents sous-types histologiques de cancers broncho-pulmonaires<sup>357</sup>. Nous allons plus particulièrement nous focaliser sur les transcrits dont la fonction et le mode d'action sont fortement documentés (Figure 36).

##### *i. Les lncARNs pulmonaires oncogéniques*

Les lncARNs qui favorisent la tumorigenèse sont des oncogènes. Leur expression est augmentée dans les cancers et est corrélée à un mauvais pronostic des patients. Ils peuvent exercer diverses fonctions régulant la prolifération, la viabilité, la résistance aux traitements et/ou la transformation métastatique des cellules cancéreuses.

##### *a. Les lncARNs favorisant la prolifération des cellules cancéreuses*

Nous avons précédemment abordé les lncARNs ANRIL, THOR, MYMLR et PVT1 qui favorisent la prolifération de différents types de cellules cancéreuses incluant des cellules tumorales pulmonaires<sup>306,311,312,358</sup>. D'autres lncARNs promeuvent la prolifération de cellules cancéreuses pulmonaires. C'est le cas de divers lncARNs nucléaires comme LUADT1 (Lung Adenocarcinoma Associated Transcript 1), HOTAIR et CASC9.5 (Cancer Susceptibility 9) qui guident des complexes répresseurs PRC2 sur les promoteurs de gènes bloquant le cycle cellulaire (p15, p16, p27) à l'instar du mode d'action d'ANRIL et de PVT1<sup>307,321,358,359</sup>. D'autre part, l'invalidation du lnc-Sox2ot (SOX2 Overlapping Transcript) par siARN réprime la prolifération des cellules cancéreuses pulmonaires par l'induction de la transcription de la protéine EZH2 du complexe PRC2<sup>361,362</sup>.

Inversement, d'autres lncARNs nucléaires favorisent l'expression de gènes pro-prolifératifs en guidant des facteurs de transcription sur leurs promoteurs. c'est le cas du lncARN EPEL (E2F-mediated cell proliferation enhancing lncRNA) qui recrute E2F1 sur les promoteurs des gènes *CCNB1* et *FEN1* (Flap Structure-Specific Endonuclease 1)<sup>363</sup>. De même, LINC00473 guide NONO (Non-POU Domain Containing Octamer Binding) et CRTC (CREB Regulated Transcription Coactivator) sur les séquences de CRE (cAMP response element) de gènes de la prolifération dans des cancers pulmonaires ayant une mutation du gène suppresseur de tumeur LKB1 (Liver Kinase B1)<sup>363</sup>.

Au niveau cytoplasmique, le lncARN LCAT1 (Lung Cancer Associated Transcript 1) stimule la prolifération cellulaire en séquestrant le microARN miR-4715-5p, ce qui stabilise la protéine Rac1<sup>364</sup>.

#### b. Les lncARNs favorisant la viabilité et la résistance aux traitements

Les lncARNs oncogéniques peuvent également altérer des processus apoptotiques et jouer un rôle dans la résistance aux traitements. Nous avons déjà présenté TUG1 qui promeut la résistance aux traitements dans les CBPC en guidant le complexe PRC2 sur le promoteur du gène LIMK2B ou encore PANDA qui séquestre le facteur de transcription NF- $\kappa$ B dans les CBNPC <sup>225,326,365</sup>.

D'autres transcrits favorisent la survie des cellules cancéreuses pulmonaires tels que VELUCT (Viability Enhancing LUng Cancer Transcript) en s'associant à la chromatine ou PILAR1 (Prognostic in Lung Adenocarcinoma LncRNA-1) exprimé dans les cancers présentant une mutation du gène KEAP1 et qui induit des mécanismes de réparation de l'ADN <sup>366,367</sup>.

Récemment, l'implication du lncARN ALAL-1 (Amplified LncRNA Associated with Lung cancer - 1) dans la prolifération et l'échappement à l'immunosurveillance a été mise en évidence. ALAL-1 est un lncARN cytosolique induit par TNF $\alpha$  qui en interagissant avec la protéine SART3 (Squamous cell carcinoma antigen recognized by T cells 3) permet la translocation nucléaire de USP4 (Ubiquitine specific protéase 4). La translocation de USP4 entraîne la répression des voies Wnt/ $\beta$ -catenin, p53, TGF $\beta$  et NF- $\kappa$ B et ainsi l'échappement à l'immunosurveillance <sup>369</sup>.

#### c. Les lncARNs favorisant la formation de métastases et l'invasion

L'un des aspects majeurs de la progression de l'agressivité des tumeurs concerne leur capacité à former des métastases et à envahir d'autres niches. Précédemment, nous avons discuté du mode d'actions des lncARNs MALAT1, H19, MetaLnc9/LINC00963 et lncRNA-HIT dans la formation de métastases dans différents cancers incluant les cancers broncho-pulmonaires <sup>254,320,337,369</sup>. D'autres lncARNs nucléaires régulent des effecteurs de la TEM. Ainsi, ELIT-1/LINC00842 (EMT associated lncRNA induced by TGF $\beta$ 1), induit par le TGF $\beta$ , guide SMAD3 sur le promoteur de *SNAIL* <sup>371</sup>. De même, NSCLCAT1 (Non Small Cell Lung Cancer Associated Transcript-1) active la voie YAP/TAZ, inhibant la cadhérine E et modifiant la survie et les capacités invasives et prolifératives des cellules cancéreuses par un mécanisme encore indéterminé <sup>372</sup>.

Certains lncARNs cytoplasmiques favorisent également l'invasion métastatique par différents mécanismes. Ainsi, les surexpressions de LIN28B-AS1 (Lin-28 Homolog B Antisens 1) et MUC5B-AS1 (Mucin 5B, Oligomeric Mucus/Gel-Forming Antisens 1) qui stabilisent respectivement les ARNm LIN28 et MUC5B stimulent la formation de métastases dans des modèles de xénogreffes chez la souris. Un autre exemple concerne RGMB-AS1 (Repulsive Guidance Molecule BMP Co-Receptor B antisens 1) qui favorise la TEM en déstabilisant la protéine RGMB répresseur de la voie BMP, et dont l'inactivation diminue la formation de métastases chez la souris <sup>372,373,374</sup>. Parmi les nombreux lncARNs décrits pour

servir d'éponges à microARNs, on peut citer PNUMS qui séquestre miR-205 et dont l'inhibition altère la formation de métastases *in vivo*. Enfin, des approches de gain et de perte de fonction *in vivo* ont révélé le rôle de LINC00673-v4 dans la formation de métastases via la formation du complexe DDX3-CK1 $\epsilon$  et la transduction des signaux de la voie Wnt/ $\beta$ -caténine <sup>375,376</sup>.

## ii. Les lncARNs pulmonaires suppresseur de tumeurs

D'autres lncARNs réprimés dans les cancers broncho-pulmonaires jouent le rôle de suppresseurs de tumeurs. Leur fort niveau d'expression est associé à un bon pronostic vital. Ils exercent diverses fonctions comme la répression de la prolifération, de la survie ou encore de la TEM et de la formation de métastases.

### a. Les lncARNs inhibiteurs de la prolifération

Certains lncARNs inhibent des voies de signalisation pro-tumorales comme NKILA qui réprime la voie NF- $\kappa$ B via un mécanisme que nous avons précédemment détaillé (Fig. 31.D) <sup>257,378</sup>. D'autres voies peuvent être impactées, telles que la voie Notch1 inhibée par NBR2 (Neighbor Of BRCA1 lncRNA 2) ou la voie JAK/STAT3 bloquée par PICART1 (P53 Inducible Cancer Associated RNA Transcript 1) par des mécanismes non encore détaillés <sup>378,379,380</sup>. Le variant nucléaire de PVT1, PVT1b, induit par p53, bloque l'expression de Myc <sup>382</sup>. TARID (TCF21 Antisense RNA Inducing Promoter Demethylation) est un autre exemple de lncARN nucléaire réprimant la prolifération des cellules cancéreuses. Il guide le remodelleur chromatinien GADD45A (Growth Arrest And DNA Damage Inducible Alpha) sur le promoteur du suppresseur de tumeur TCF21 (Transcription Factor 21) afin de le démétyler et d'induire son expression <sup>383</sup>.

De nombreuses études fonctionnelles mettent en évidence des interactions entre les lncARNs et les microARNs. Cependant, ces études sont sujettes à controverse à cause des différences stoéchiométriques existant entre les forts niveaux de microARNs et les lncARNs, faiblement exprimés et présentant peu de sites de fixations de miARNs <sup>384</sup>. De plus, certaines études ne tiennent pas compte non plus des localisations subcellulaires différentes des différents partenaires <sup>385</sup>. Dans ces conditions, il est peu probable que ces interactions conduisent à l'inhibition fonctionnelle des miARNs. Un exemple de la littérature concerne le lncARN cytoplasmique SPRY4-IT1 (Sprouty RTK Signaling Antagonist 4 Intronic Transcript 1) réprimé par le complexe EZH2 dans les CBNPC. Ce lncARN est décrit pour séquestrer miR-101-3p et ainsi réprimer la prolifération tumorale <sup>386,387</sup>.

### b. Les lncARNs pro-apoptotiques

Des transcrits comme MEG3, GAS5 et NKX2-1-AS, réprimés dans les cancers broncho-pulmonaires et dont nous avons déjà détaillé le mode d'action, favorisent la mort cellulaire en induisant

la voie p53 ou différentes protéines pro-apoptotiques <sup>331,350,387</sup>. Ainsi, P53RRA (P53 Related LncRNA) séquestre la protéine G3BP1 répresseur de la voie p53 dans le cytoplasme, tandis que MIR22HG stabilise le facteur de transcription YBX1 (Y-Box Binding Protein 1) dans le noyau permettant la transcription de p21, effecteur de la voie p53 <sup>389,390</sup>. Par ailleurs, LINC00961 favorise l'expression de la protéine Bax <sup>390</sup>.

### c. Les lncARNs anti-métastatiques

Certains lncARNs répriment la formation de métastases dans les cancers broncho-pulmonaires en inhibant les processus de la TEM comme BANCR que nous avons précédemment décrit <sup>316</sup>.

De même, LET, réprimé par l'hypoxie, séquestre dans le noyau la protéine ILF3/NF-90, impliquée dans la formation des granules de stress, des p-bodies et dans la translocation nucléocytoplasmique, et participe à l'inhibition des métastases <sup>227</sup>.

D'autres lncARNs répriment la TEM en séquestrant des protéines. C'est le cas de GIAT4RA (LncRNA between GINS4 and GPAT4) qui stabilise la protéine LSH (Lymphoid-Specific Helicase) dans le noyau et empêche la fixation de l'ubiquitine thiolesterase Uchl3 (Ubiquitin C-Terminal Hydrolase L3). Quant à NORAD, il séquestre la protéine S100P dans le cytoplasme <sup>392,393</sup>.

Enfin, la répression de lnc-LINP1 (LncRNA In Non-Homologous End Joining Pathway 1) favorise la formation de métastases via un mécanisme qui impliquerait sa capacité à former un échafaudage de protéines de la machinerie de réparation de l'ADN (DNA PKs et Ku80) <sup>394</sup>.

## LNCARNS ONCOGENES

### Prolifération

ANRIL (Nucl)	Guide PCR2 sur promoteur p15
CASC9.5 (Nucl)	Guide DNMT1
EPEL (Nucl)	Guide E2F1 sur promoteurs de CCNB1 et FEN1
HOTAIR (Nucl)	Guide PCR2
LCAT1 (Cyt)	Eponge à miR-4715-5p qui régule Rac1
LIN28B-AS1 (Cyt)	Stabilise LIN28 via la protéine IGF2BP1
LINC00473 (Nucl)	Guide NONO et CRTC
Lnc-Sox2ot (Nucl)	Favorise l'expression EZH2 (PCR2)
LUADT1 (Nucl)	Guide PCR2 sur le promoteur de p27
MYMLR (Nucl)	Repliement chromatidien favorisant l'expression de Myc
PILAR1 (Nucl)	Réparation de l'ADN dans les tumeurs KEAP1-/-
NSCLCAT1	Régulation de la voie YAP/TAZ et Cadherine E
PVT1 (Nucl)	Guide PCR2 sur promoteurs p15/p16
RGMB-AS1 (Cyt)	Déstabilise la protéine RGMB et voie BMP
THOR (Cyt)	Stabilise la protéine IGF2BP1

### Migration / Invasion

CASC9.5 (Nucl)	Guide DNMT1
ELIT-1/LINC00842 (Nucl)	Guide SMAD3 sur promoteur SNAIL
H19 (Cyt)	Stabilise l'ARNm b-caténine
HOTAIR (Nucl)	Guide PCR2
LIN28B-AS1 (Cyt)	Stabilise LIN28 via la protéine IGF2BP1
LINC00673-v4 (Cyt)	Transduit signal de la voie Wnt/b catenine
PNUTS (Nucl + Cyt)	Eponge à miR-205
LncRNA HIT (Nucl)	Stabilise le facteur de transcription ZEB1
MALAT1 (Nucl)	Modulation de l'épissage
MetaLnc9/LINC00963 (Cyt)	Stabilise PGK1
MUC5B-AS1 (Cyt)	Stabilise l'ARNm MUC5B
NSCLCAT1	Régulation voie YAP/TAZ et Cadherine E
RGMB-AS1 (Cyt)	Déstabilise la protéine RGMB et voie BMP

### Viabilité / Résistance aux traitements

ALAL-1	Favorise translocation nucléaire de USP4
NSCLCAT1	Régulation voie YAP/TAZ et Cadherine E
PANDA (Nucl)	Séquestration du facteur de transcription NF-YA
PILAR1 (Nucl)	Réparations de l'ADN dans les tumeurs KEAP1 -/-
TUG1 (Nucl)	Guide PCR2 sur le promoteur LIMK2B
VELUCT (Nucl)	Association à la chromatine

## LNCARNS SUPPRESSEURS DE TUMEURS

### Prolifération

NBR2	Inhibe la voie Notch
NKILA (Cyt)	Bloque la voie NF-kB et STAT3
P53RRA (Cyt)	Séquestre G3BP1, activation p53
PICART1	Inhibe la voie JAK2/STAT3 et MAPK
PVT1b (Nucl)	Inhibe la transcription de Myc
SPRY4-IT1 (Cyt)	Eponge à miR-101-3p
TARID (Nucl)	Active TCF21 en déméthylant le promoteur

### Migration / Invasion

BANCR (Nucl)	Induit gènes épithéliaux
GIAT4RA (Nucl)	Stabilise la protéine LSH
Inc-LINP1 (Nucl)	Echaffaudage des protéines DNA-PK et Ku-80
LncRNA-LET (Nucl)	Séquestre et dégrade ILF3/NF-90
NBR2	Inhibe la voie Notch
NKILA (Cyt)	Bloque la voie NF-kB et STAT3
NORAD (Cyt)	Séquestre la protéine S100P
SPRY4-IT1 (Cyt)	Eponge à miR-101-3p

### Viabilité / Sensibilité aux traitements

GAS5 (Nucl)	Guide PCR2 sur promoteur E2D1
LINC00961	Induit l'expression de Bax
MEG3 (Cyt)	Déstabilise MDM2 répresseur de p53
MIR22HG (Nucl)	Stabilise YBX1 et active p21
NKX2-1-AS (Nucl)	Séquestre NKX2 et expression de PD-L1
P53RRA (Cyt)	Séquestre G3BP1, activation p53

**Figure 36:** Les lncARNS dans les cancers broncho-pulmonaires.

Les cases oranges du tableau représentent les lncARNS oncogéniques et en bleu, les lncARNS suppresseurs de tumeurs. (Abréviations: Nucl. : Nucléaire ; Cyt. : Cytoplasmique)

## **IV. RESULTATS**



## Avant-propos

Les cancers broncho-pulmonaires constituent la première cause de mortalité par cancer en France et dans le monde. Ces tumeurs agressives se caractérisent par un mauvais pronostic associé à un important taux de récurrence. En effet, malgré des prises en charges précoces par exérèse curative, près de 50% des cancers pulmonaires récidivent dans les 5 ans. La découverte de nouveaux marqueurs pronostiques et de nouvelles cibles thérapeutiques constituent des enjeux majeurs de santé publique et ont été au cœur de mon projet de thèse.

Dans cette optique, nous nous sommes intéressés à l'environnement hypoxique et à la famille émergente des longs ARNs non codants. L'agressivité des cancers broncho-pulmonaires est en partie liée au microenvironnement tumoral et notamment à son appauvrissement en oxygène conséquent à la rapidité de la croissance tumorale. Ce contexte hypoxique induit une reprogrammation métabolique des cellules tumorales et favorise l'angiogenèse, l'expansion métastatique et la résistance aux traitements via une importante régulation de l'expression génique par les voies HIF, NF- $\kappa$ B et NRF2. Ces modifications sont liées à l'expression de gènes codants et de gènes non codants. Notre équipe avait précédemment mis en évidence le rôle de l'hypoxamiR, miR-210, dans le switch métabolique et dans la radiorésistance des CBNPC (Figure 13) <sup>154,155</sup>. Nous nous sommes également intéressés à la famille émergente des longs ARNs non codants qui régulent l'expression génique au travers de complexes ribonucléoprotéiques et qui apparaissent comme des marqueurs pronostiques d'intérêts et de potentielles cibles thérapeutiques.

Des cribles transcriptomiques et des analyses bioinformatiques antérieures à mon arrivée au laboratoire ont permis d'identifier une signature génique de lncARNs régulés par l'hypoxie dans les adénocarcinomes pulmonaires (ADC) et/ou associés à un mauvais pronostic des patients.

Parmi ces transcrits, LUCAT1 était le plus fortement induit par l'hypoxie et a été le premier candidat dont nous avons entrepris la caractérisation structurale et fonctionnelle. A mon arrivée au laboratoire, le Dr Laura Moreno Leon avait précisé les différentes isoformes du transcrit, sa régulation transcriptionnelle et caractérisé les effets de son invalidation par CRISPR/Cas9 sur le transcriptome, la prolifération cellulaire et la sensibilité au cisplatine. Mon premier travail a consisté à finaliser ces travaux en levant notamment des verrous technologiques afin d'appréhender i) la localisation subcellulaire et tissulaire de ces transcrits par imagerie à l'échelle de la molécule unique, et ii) leur mode d'action, par l'identification de leurs interactants moléculaires par des approches biochimiques de précipitations couplées à des analyses à haut débit (séquençage ADN, analyse protéomique). J'ai ainsi

précisé la fonction de l'isoforme majoritaire nucléaire NLUCAT1 et d'une isoforme cytosolique plus courte, dans la régulation i) de l'expression génique, ii) du stress oxydatif et iii) de l'invasion et la migration des cellules cancéreuses. Ce travail a été publié en Août 2019 dans la revue "Oncogene" et constitue la première partie des résultats de mon manuscrit.

Concomitamment, j'ai entrepris l'étude d'un autre lncARN candidat, LINC01116, induit par l'hypoxie dans les ADC et lié à une diminution de la survie globale des patients. J'ai réalisé la caractérisation structurale de ce transcrit et j'ai abordé son étude fonctionnelle par des approches de gain (transduction lentivirale) et de perte (siRNA, Gapmers, CRISPRi) de fonction dans des lignées de CBNPC (adénocarcinomes et carcinomes épidermoïdes) et dans des cellules endothéliales dans lesquelles ce lncARN est fortement exprimé. J'ai également identifié des partenaires moléculaires de ce transcrit afin de préciser son mode d'action. Les données obtenues sur ce lncARN sont présentées dans une seconde partie et font l'objet d'une seconde publication soumise.

L'ensemble des résultats obtenus et les différentes perspectives du projet seront ensuite discutés dans une conclusion générale.

## Partie 1 : NLUCAT1, un hypoxa-lnc nucléaire impliqué dans la régulation du stress oxydatif des adénocarcinomes pulmonaires.

Article

*“The nuclear hypoxia-regulated NLUCAT1 long non-coding RNA contributes to an aggressive phenotype in lung adenocarcinoma through regulation of oxidative stress.”*

Moreno Leon, L.\*, Gautier, M.\*, Allan, R., Ilié, M., Nottet, N., Pons, N., Paquet, A., Lebrigand, K., Truchi, M., Fassy, J., Magnone, V., Kinnebrew, G., Radovich, M., Check, M.H., Barbry, P., Vassaux, G., Marquette, C.-H., Ponzio, G., Ivan, M., Pottier, N., Hofman, P., Mari, B. and Rezzonico, R. (2019).  
(\* contributed equally) *Oncogene* 38, 7146–7165



# The nuclear hypoxia-regulated NLUCAT1 long non-coding RNA contributes to an aggressive phenotype in lung adenocarcinoma through regulation of oxidative stress

Laura Moreno Leon<sup>1,2</sup> · Marine Gautier<sup>1,2</sup> · Richard Allan<sup>1,2</sup> · Marius Ilié<sup>2,3,4</sup> · Nicolas Nottet<sup>1,2</sup> · Nicolas Pons<sup>1,2</sup> · Agnes Paquet<sup>1,2</sup> · Kévin Lebrigand<sup>1,2</sup> · Marin Truchi<sup>1,2</sup> · Julien Fassy<sup>1,2</sup> · Virginie Magnone<sup>1,2</sup> · Garrett Kinnebrew<sup>5</sup> · Milan Radovich<sup>5</sup> · Meyling Hua-Chen Cheok<sup>6</sup> · Pascal Barbry<sup>1,2</sup> · Georges Vassaux<sup>2,7</sup> · Charles-Hugo Marquette<sup>2,3,8</sup> · Gilles Ponzio<sup>1,2</sup> · Mircea Ivan<sup>9</sup> · Nicolas Pottier<sup>10</sup> · Paul Hofman<sup>2,3,4</sup> · Bernard Mari<sup>1,2</sup> · Roger Rezzonico<sup>2,7</sup>

Received: 15 August 2018 / Revised: 15 May 2019 / Accepted: 7 June 2019

© The Author(s), under exclusive licence to Springer Nature Limited 2019

## Abstract

Lung cancer is the leading cause of cancer death worldwide, with poor prognosis and a high rate of recurrence despite early surgical removal. Hypoxic regions within tumors represent sources of aggressiveness and resistance to therapy. Although long non-coding RNAs (lncRNAs) are increasingly recognized as major gene expression regulators, their regulation and function following hypoxic stress are still largely unexplored. Combining profiling studies on early-stage lung adenocarcinoma (LUAD) biopsies and on A549 LUAD cell lines cultured in normoxic or hypoxic conditions, we identified a subset of lncRNAs that are both correlated with the hypoxic status of tumors and regulated by hypoxia in vitro. We focused on a new transcript, NLUCAT1, which is strongly upregulated by hypoxia in vitro and correlated with hypoxic markers and poor prognosis in LUADs. Full molecular characterization showed that NLUCAT1 is a large nuclear transcript composed of six exons and mainly regulated by NF- $\kappa$ B and NRF2 transcription factors. CRISPR-Cas9-mediated invalidation of NLUCAT1 revealed a decrease in proliferative and invasive properties, an increase in oxidative stress and a higher sensitivity to cisplatin-induced apoptosis. Transcriptome analysis of NLUCAT1-deficient cells showed repressed genes within the antioxidant and/or cisplatin-response networks. We demonstrated that the concomitant knockdown of four of these genes products, GPX2, GLRX, ALDH3A1, and PDK4, significantly increased ROS-dependent caspase activation, thus partially mimicking the consequences of NLUCAT1 inactivation in LUAD cells. Overall, we demonstrate that NLUCAT1 contributes to an aggressive phenotype in early-stage hypoxic tumors, suggesting it may represent a new potential therapeutic target in LUADs.

These authors contributed equally: L. Moreno-Leon, M. Gautier

These authors contributed equally: B. Mari, R. Rezzonico

Supplementary information The online version of this article (<https://doi.org/10.1038/s41388-019-0935-y>) contains supplementary material, which is available to authorized users.

\* Bernard Mari  
mari@unice.fr

\* Roger Rezzonico  
rezzonico@ipmc.cnrs.fr

Extended author information available on the last page of the article  
Published online: 15 August 2019

## Introduction

Non-small cell lung cancer (NSCLC) accounts for at least 80% of lung cancer cases, with lung adenocarcinoma (LUAD) being the most common form (~50%). Although detection and further tumor surgical resection of early-stage NSCLC has significantly improved the survival of patients, ~50% of patients with stage I NSCLC die within 5–10 years [1, 2]. Hypoxia, a condition frequently encountered in solid tumors, activates the hypoxia-inducible factor (HIF) transcription complex that orchestrates the expression of a myriad of genes, whose functions are to ensure neovascularization, glucose metabolism, cell survival, tumor spread, and resistance to most anticancer drugs [3, 4].

surprisingly, the HIF1A transcript is part of a prognostic classifier signature of five genes in stage I LUAD [5] and high expression of its downstream targets such as CA9 (carbonic anhydrase 9) is a predictor of worse survival [6]. Extensive evidence demonstrates that microRNAs (miRNAs) are key components of the adaptive response to low oxygen in tumors [7, 8]. We have notably underlined the central role of the prototypical “hypoxamiR” miR-210 on tumor metabolic changes and on radio-resistance in lung cancer [9, 10].

Considerable attention has recently been focused on the emerging class of long non-coding RNAs (lncRNAs), with more than 10,000 transcripts validated to date. LncRNAs, defined as transcripts longer than 200 nucleotides that do not encode proteins, are usually transcribed at lower abundance than mRNAs in a tissue-specific manner and they usually exhibit a poor primary sequence conservation [11–13]. LncRNAs encompass pseudogenes, natural antisense transcripts (NATs), and the long intervening non-coding RNAs (lincRNAs). Many lncRNAs function as part of ribonucleoprotein complexes that regulate gene expression through diverse molecular mechanisms including recruitment of transcriptional activators or repressors, chromatin epigenetic modifications, or control of mRNA processing, translation, or stability [13].

Integrative genomic analyses have identified lncRNAs with a prognostic potential that may play a role in tumor progression in four types of cancers, including lung squamous cell carcinomas (LUSCs) [14]. Moreover, several lncRNAs map to regions associated with disease by genome-wide association studies and accumulating evidence has demonstrated that they control gene networks that are relevant to cancer development and metastasis [15]. Some lncRNAs regulated by hypoxia are notably involved in metabolic adaptation and tumor spreading [16–20].

However, only a few predicted human lncRNAs among thousands have been well characterized for their expression pattern, structure, precise function, and clinical significance, especially in the context of NSCLC and of their hypoxia-mediated aggressiveness [21–23].

In this context, we performed a comprehensive genome-wide profiling to define hypoxia-regulated lncRNAs in an early-stage LUAD cohort and in a LUAD cell line *in vitro*. Among several “hypoxia-lncRNA” candidates, we identified NLUCAT1, a large nuclear lncRNA, which was upregulated in hypoxic conditions in LUAD cell lines and overexpressed in early-stage LUADs according to their hypoxic and overall survival (OS) status. Through CRISPR-Cas9-mediated invalidation, we demonstrated the pro-tumoral properties of NLUCAT1 and we showed that this transcript is associated with the regulation of cell proliferation, cell invasion, and cell survival in response to

oxidative stress, and could, consequently, represent a new target for future therapies.

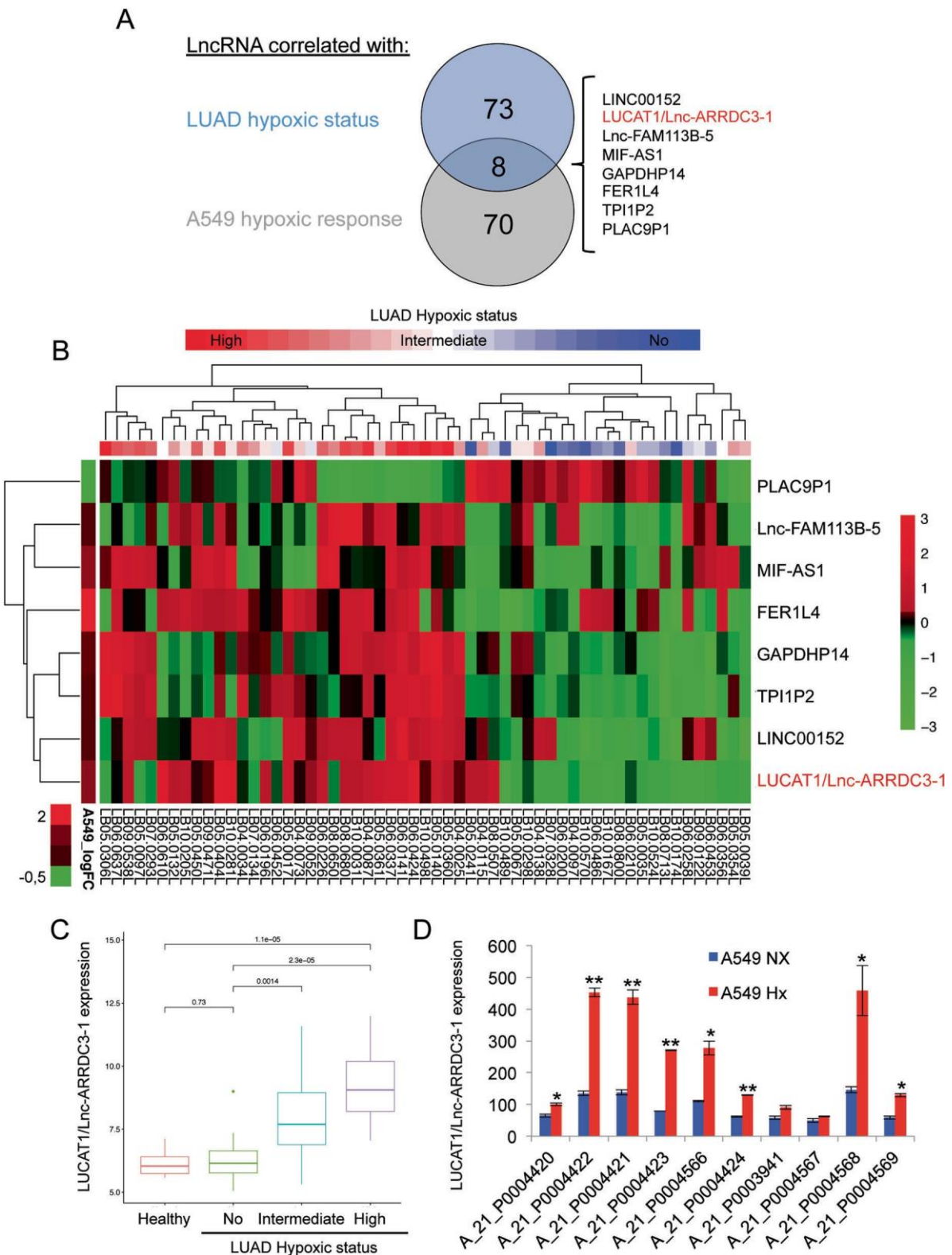
## Results

### Genome-wide profiling of hypoxia-regulated lncRNAs in LUAD

The identification of lncRNAs controlled by hypoxia in LUAD was performed through a combination of profiling approaches using microarrays enriched for lncRNA probes, on both clinical samples and cultured cell lines. First, we performed a transcriptomic profiling on lung samples from 57 early-stage LUAD patients (Supplementary Table S1 and Dataset 1) that were *a posteriori* classified in highly, moderately, and non-hypoxic groups using a metagene signature [24]. We overall found that 775 transcripts were differentially expressed between tumor and normal lung tissue, and positively correlated with the hypoxic status of tumors, whereas 1225 transcripts were negatively correlated with this hypoxic status. Among these genes, 81 potential lncRNAs were associated with the hypoxic status of the tumors (Fig. 1a and Supplementary Table S2). In parallel, we identified 78 predicted lncRNAs that are modulated in hypoxic condition in the A549 LUAD cell line (Fig. 1a, Supplementary Table S2, and Dataset 2). The comparative analysis of these data identified three lincRNAs (LINC00152, LUCAT1/Lnc-ARRDC3-1, Lnc-FAM113B-5), three pseudogenes (FER1L4, GAPDHP14, TLI1P2), and one NAT (MIF-AS1), which were upregulated by hypoxia both *in vivo* and *in vitro*, and conversely one pseudogene, PLAC9P1, which was repressed under oxygen deprivation (Fig. 1a, b).

### LUCAT1/Lnc-ARRDC3-1 is a new “hypoxia-lncRNA”

We focused our attention on the LUCAT1/Lnc-ARRDC3-1, the short isoform SCAL1 of which has previously been shown to be induced by cigarette smoke extracts in NSCLC cell lines [25]. We showed that LUCAT1/Lnc-ARRDC3-1 expression is directly correlated with the hypoxic status of tumors (Fig. 1c). In addition, microarray data showed that eight out of ten probes located within this locus detected a significant increase in LUCAT1/Lnc-ARRDC3-1 expression in A549 cells cultured under hypoxic conditions (Fig. 1d). RNA sequencing (RNA-seq) analysis of A549 cells cultured under normoxic or hypoxic conditions also showed reads aligned all along the genomic region of the *LUCAT1/Lnc-ARRDC3-1* locus (chr5: 90597137-90610208) and validated induction by hypoxia (Fig. 2a). In agreement, using distinct quantitative PCR (qPCR) assays distributed throughout the locus, we confirmed the induction of



LUCAT1/Lnc-ARRDC3-1 by hypoxia in two other LUAD cell lines (Supplementary Fig. S1).

We further investigated the expression of this transcript in RNA-seq datasets from The Cancer Genome Atlas

(TCGA) database (Supplementary Table S3). The analysis of 535 LUAD and 502 LUSC RNA-seq profiles confirmed that LUCAT1/Lnc-ARRDC3-1 is upregulated in relation to the hypoxic status of tumors in LUAD and, to a lesser



◀ Fig. 1 Identification of the LUAD hypoxia-regulated lncRNA signature. a Venn diagram summarizing lncRNA candidates that are modulated by hypoxia in LUAD biopsies and in the A549 LUAD cell line. lncRNA expression was first analyzed by microarray hybridization in 57 early-stage LUAD and 11 peritumoral healthy lung biopsies. Tumor samples were classified by function of their hypoxic status score and selection of hypoxia-associated lncRNAs was performed using  $\log_2(\text{intensity}) > 7$ , absolute  $\log_2(\text{LUAD/Healthy}) > 0.7$ ,  $\text{FDR} < 0.05$  and adj.  $P$ -value hypoxic status  $\text{pca} < 0.05$ . lncRNA signature of A549 LUAD cells cultured in normoxic or hypoxic conditions for 24 h was done using  $\log_2(\text{intensity}) > 7$ , absolute  $\log_2(\text{Hx/Nx}) > 0.7$  and adj.  $P$ -value  $< 0.05$ . The eight overlapping “hypoxia-lncRNA” candidates that are both deregulated by hypoxia in vivo and in vitro are shown. b Heatplot showing the profile of expression of the 8 “hypoxia-lncRNA” candidates in the 57 LUAD samples depending on their hypoxic status (High in red, No in blue). Expression corresponds to  $\log_2$  intensity followed by median centering; clustering was performed using a Manhattan distance metric and average linkage. Differential expression between normoxic and hypoxic A549 cells is displayed as  $\log_2$  fold change on the left. c Boxplot representation of LUCAT1/Lnc-ARRDC3-1 expression in peritumoral healthy lungs ( $n = 11$ ) and in LUAD sample subsets according to their hypoxic status (No  $n = 12$ , intermediate  $n = 25$ , High  $n = 20$ ). The results are displayed on a  $\log_2$  scale. Statistical relevance of the comparisons was determined with the Wilcoxon’s test. d Level of expression of LUCAT1/Lnc-ARRDC3-1 transcripts detected by the ten microarray probes spanning the LUCAT1/Lnc-ARRDC3-1 locus in A549 cells in normoxia or hypoxia. Data are means  $\pm$  SD from two replicates. Statistically significant differences are indicated (\* $p < 0.05$ , \*\* $p < 0.01$ )

extent, in LUSC samples (Fig. 2b). This increase in expression takes place in early-stage LUAD and LUSC with no supplemental exacerbation in more advanced stages (Fig. 2c). We then evaluated the impact of LUCAT1/Lnc-ARRDC3-1 expression on patient prognosis. LUAD patients assigned to the LUCAT1/Lnc-ARRDC3-1 upper- quartile group exhibited a statistically shorter OS time than those in the lower-quartile group (log-rank test  $p = 0.0012$ ) (Fig. 2d).

### Characterization of NLUCAT1, a large nuclear variant of LUCAT1/Lnc-ARRDC3-1

Currently, only two short LUCAT1/Lnc-ARRDC3-1 transcripts of 890 nt (LUCAT1:4) and 582 nt (LUCAT1:5), respectively, are listed in the Refseq database (Fig. 2a). However, as indicated, more than 20 distinct LUCAT1 predictive transcripts ranging from 212 nt to 9585 nt are annotated in the LNCipedia database, which regroups RNA-seq information from various sources (Broad Institute, Gencode, NONCODE, Refseq, Ensembl...) [26]. Based on this complex situation and on the poor definition of the transcript boundaries frequently associated with lncRNA sequences, we performed additional characterization of the main LUCAT1/Lnc-ARRDC3-1 species transcribed from this locus. Northern blot hybridizations with probes designed in putative exons 1, 5, and 6 of LNCipedia

LUCAT1/Lnc-ARRDC3-1 annotated transcripts detected a major  $\sim 10$  kb band markedly stimulated by hypoxia and an upper 12–13 kb isoform that likely corresponds to the unspliced LUCAT1/Lnc-ARRDC3-1 pre-lncRNA (Fig. 3a). Other, less abundant transcript variants were visualized including a faint transcript at around 900 nt that matched with the LUCAT1:4 isoform. To precisely define the exon composition of the main 10 kb isoform, we combined RNA-seq and reverse transcriptase-PCR (RT-PCR) approaches (Fig. 3b, c). We concluded that this main LUCAT1 variant is a 9807 nt transcript constituted of six exons (Supplementary Information). Moreover, we demonstrated that this transcript is almost exclusively localized in the nuclei of LUAD cells, with a nucleocytoplasmic ratio similar to that of NEAT1 or MALAT1 (Fig. 4). Based on this specific attribute, we named it NLUCAT1, for nuclear-LUCAT1.

Recent reports have shown that a large proportion of lncRNAs are unstable [27, 28]. Hypoxia/reoxygenation experiments and transcription inhibition with actinomycin D demonstrated that NLUCAT1 is indeed a short-lived non-coding transcript (half-life  $< 1$  h), suggesting that it could exert regulatory rather than housekeeping functions (Supplementary Fig. S2).

Hypoxia mainly controls gene expression through HIF, but also through activation of nuclear factor- $\kappa$ B (NF- $\kappa$ B) and/or NRF2, which are involved in hypoxia-induced inflammatory and antioxidant responses, respectively [29–31]. To better decipher NLUCAT1 regulation after hypoxic stress, we assessed the role of these transcription factors, as HRE (hypoxia response elements) and also NF- $\kappa$ B and NRF2 (ARE, antioxidant responsive element) consensus binding sites were found in the promoter region of the transcript. As expected, HIF1A inactivation by RNA interference (RNAi) blocked induction of the hypoxic marker CA9 but had no significant effect on NLUCAT1 expression, whereas HIF2A knockdown reduced NLUCAT1 induction by 30% (Supplementary Fig. S3A). Pharmacological inhibition of NF- $\kappa$ B-dependent transcription with BMS-345541 [32] totally abolished hypoxia-dependent NLUCAT1 upregulation (Fig. S3B). Accordingly, NLUCAT1 was also induced by the NF- $\kappa$ B-activating proinflammatory cytokine interleukin (IL)-1 $\beta$  in three LUAD cell lines (Supplementary Fig. S3C). As expected, this induction, along with that of the *PTGS2* inflammatory gene, was inhibited by BMS-345541 treatment. Finally, small interfering RNA (siRNA)-mediated depletion of NRF2 decreased both the basal and hypoxia-induced expression level of NLUCAT1 by 50% (Supplementary Fig. S3D). Chromatin immunoprecipitation (ChIP) coupled to qPCR was then performed to assess the direct interaction of these transcription factors with the promoter region of *NLUCAT1*. For this purpose, we used data from the ENCODE Consortium for several available human cell lines

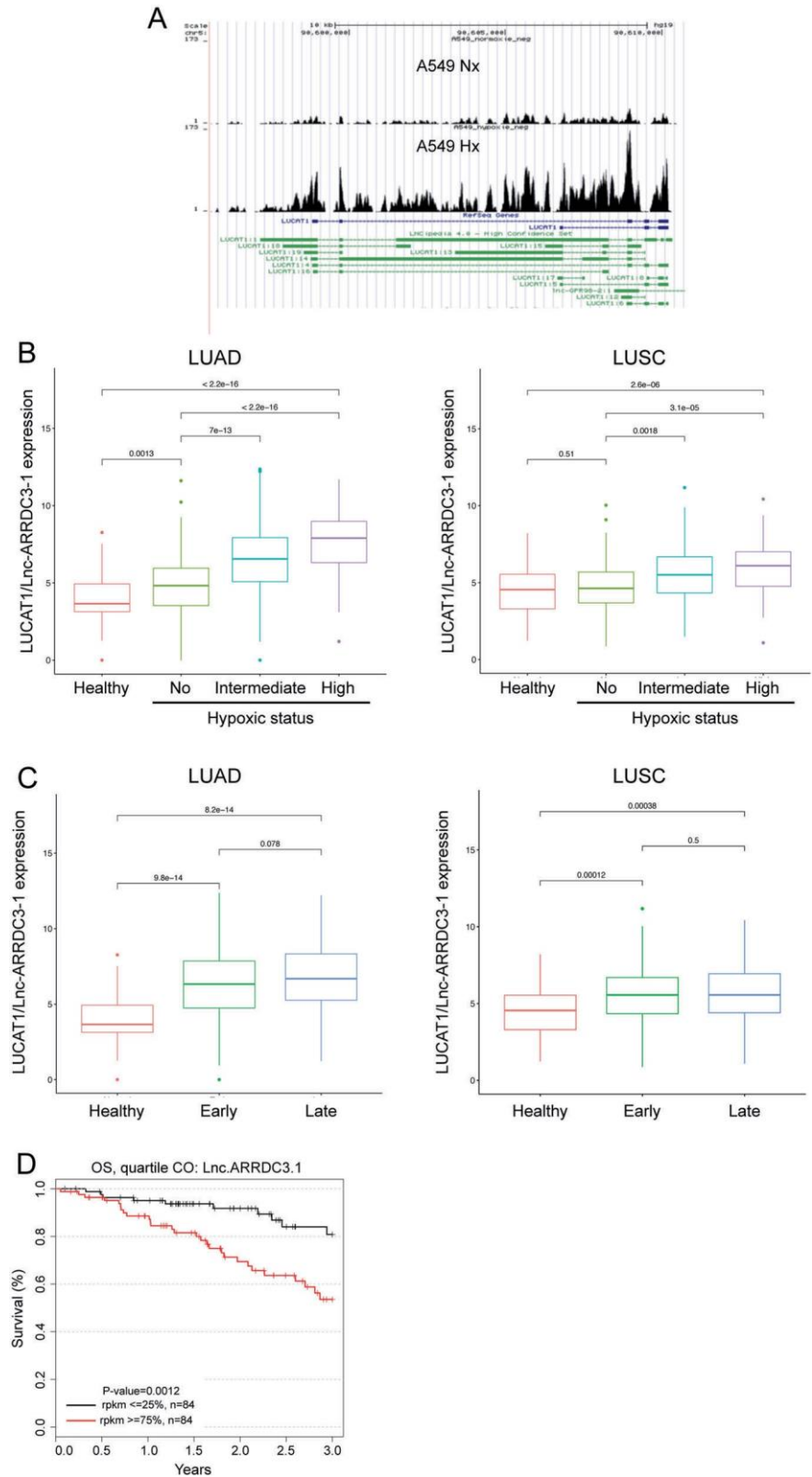
Fig. 2 LUCAT1/Lnc-ARRDC3-

1 is upregulated by hypoxia in vivo and in vitro. a

Screenshot from UCSC browser displaying *LUCAT1/Lnc-ARRDC3-1* genomic locus on chromosome 5, including the Refseq and LNCipedia database transcript annotations as well as raw RNA-seq reads plotted for A549 cells cultured for 24 h in normoxic or hypoxic conditions. b Distribution of LUCAT1/Lnc-ARRDC3-1 expression in LUAD and LUSC cohorts from TCGA according to their hypoxic status (LUAD: Healthy  $n = 59$ , No  $n = 121$ ,

Intermediate  $n = 311$ , High  $n =$

103) (LUSC: Healthy  $n = 49$ , No  $n = 52$ , Intermediate  $n = 346$ , High  $n = 104$ ). RNA-seq data were normalized as described in the Materials and Methods. The results are displayed on a  $\log_2$  scale. Statistical differences between samples were analyzed with the Wilcoxon's signed-rank test. c Boxplots showing the distribution of LUCAT1/Lnc-ARRDC3-1 expression in LUAD and LUSC cohorts from TCGA according to tumor stage (LUAD: Healthy  $n = 59$ , Early  $n = 403$ , Late  $n = 111$ ) (LUSC: Healthy  $n = 49$ , Early  $n = 408$ , Late  $n = 92$ ). Statistical relevance was determined with the Wilcoxon's test. d Kaplan-Meier analyses of correlations between the LUCAT1/Lnc-ARRDC3-1 expression level and overall survival of TCGA LUAD patients by quartiles.  $P$ -values were calculated using the log-rank test



and identified several putative binding sites for these factors in the close vicinity of histone 3 lysine 27 acetylation (H3K27Ac) and H3K4 mono- and tri-methylation

(H3K4Me1 and H3K4Me3) marks often found near active promoters or regulatory elements (Supplementary Fig. S4A). In A549 cells cultured under hypoxia, we found a



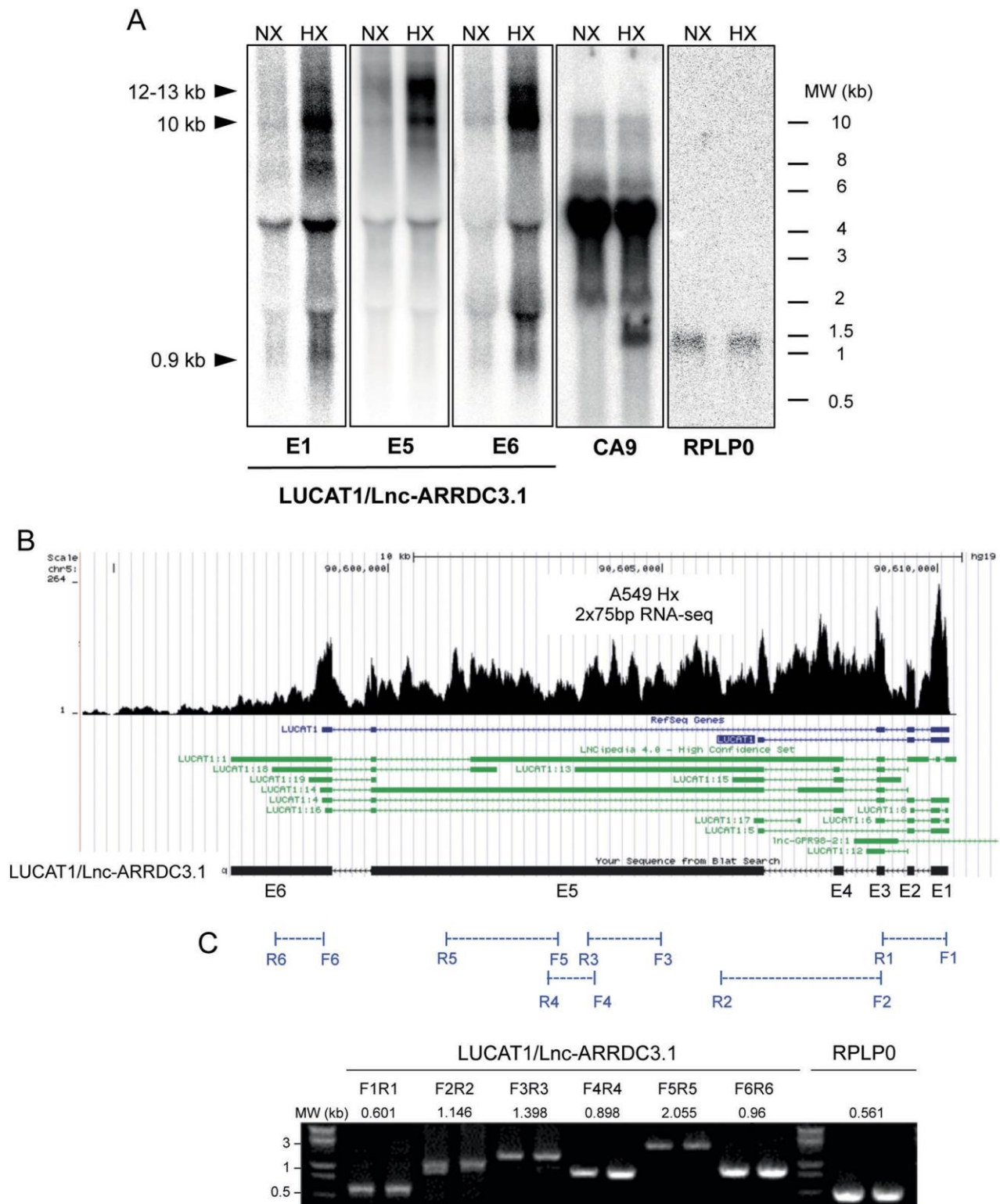


Fig. 3 Structural characterization of LUCAT1/Lnc-ARRDC3-1 transcripts in normoxic and hypoxic LUAD cell lines. a Northern blot analysis of LUCAT1/Lnc-ARRDC3-1 transcripts expression in A549 cells cultured for 24 h in normoxic or hypoxic conditions. Various [<sup>32</sup>P]-labeled cDNA probes corresponding to putative exon 1, exon 5, and exon 6 of LNCipedia annotated LUCAT1 transcript variants were successively hybridized on blotted RNA. RPLP0 (1229 nt) and CA9 (1561 nt) hybridizations were used as loading and hypoxic response

controls, respectively. b Screenshot from UCSC browser displaying raw RNA-seq reads for hypoxic A549 cells (NextSeq500 Illumina 75 bp paired-end reads) plotted on LUCAT1/Lnc-ARRDC3-1 genomic region and representation of de novo assembled ~ 10 kb LUCAT1/Lnc-ARRDC3-1 transcript. c RT-PCR analyses of LUCAT1/Lnc-ARRDC3-1 amplicons performed on two DNase-treated RNA preparations from A549 cells cultured for 24 h under hypoxia

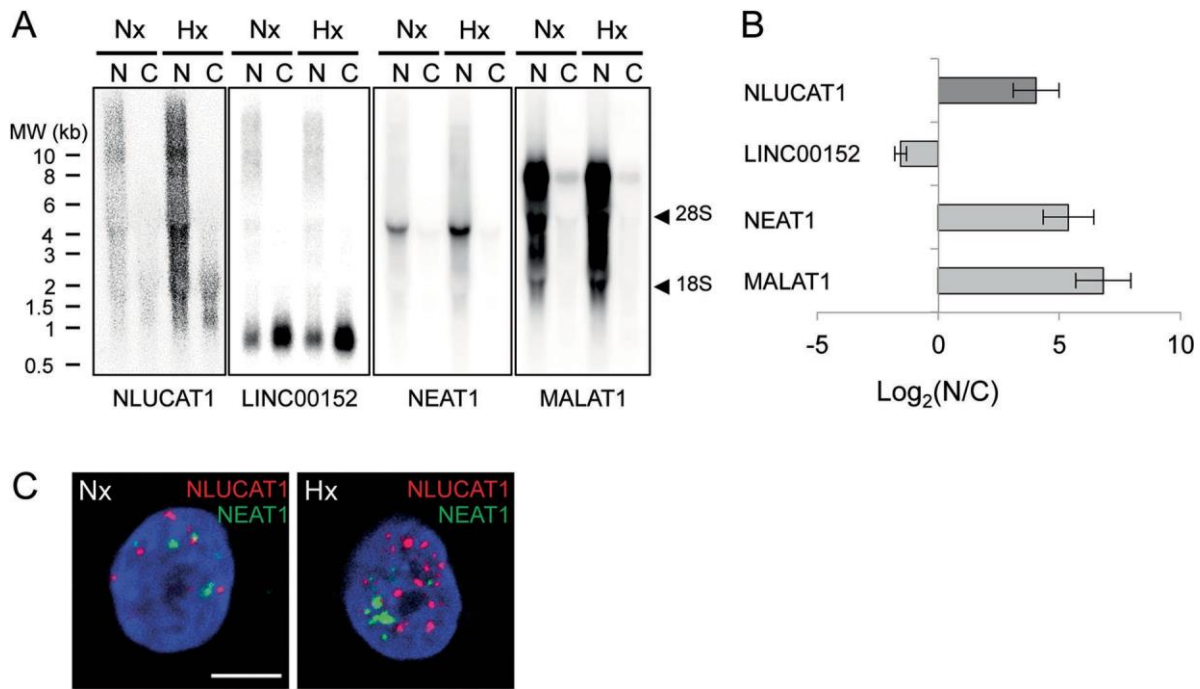


Fig. 4 NLUCAT1, a large nuclear variant of LUCAT1/Lnc-ARRDC3-

1. Subcellular localization of the 10 kb LUCAT1 variant was determined on A549 cells cultured in normoxic or hypoxic conditions for 24 h using: a Northern blotting: total RNAs were extracted from the cytoplasmic (C) and nuclear (N) fractions, then separated, blotted, and hybridized with [<sup>32</sup>P]-labeled cDNA probes for NLUCAT1 (exon 6), LINC00152, NEAT1 (short 3756 nt variant), and MALAT1. b qPCR: Horizontal bars represent the ratio of the expression of the indicated

lncRNAs between the nucleus and the cytosol. NEAT1 and MALAT1 are specific markers of the nuclear fraction, whereas LINC00152 is exclusively cytoplasmic. Data are means  $\pm$  SD from five independent experiments. c Single-molecule RNA-FISH (fluorescent in situ hybridization): cells were fixed and hybridized with NLUCAT1 and NEAT1 ACD® probes. Nuclei were counterstained with DAPI and cells were visualized by confocal microscopy. Scale bar corresponds to 5  $\mu$ m.

significant enrichment of the more distal HRE sequence (HRE3) in anti-HIF2A ChIP, as compared with control IgG (Fig. S4B). A strong binding of NRF2 was also detected in the ARE located 171 bp upstream from the transcription start site (TSS) (Supplementary Fig. S4C). Finally, we validated the interaction of three NF- $\kappa$ B family members, RelA, c-Rel, and p50, to the more distal NF- $\kappa$ B-binding site (NF- $\kappa$ B2) (Fig. S4D). This site was also validated for binding RelA in cells treated with IL-1 $\beta$  (Fig. S4E). Thus, we confirmed a direct interaction of the transcription factors HIF2A, NRF2, and NF- $\kappa$ B with the *NLUCAT1* promoter in LUAD cells under hypoxic environment.

Taken together, these data demonstrate that NLUCAT1 is a large, unstable, hypoxia- and inflammation-sensitive nuclear lncRNA regulated by HIF2A, NF- $\kappa$ B, and NRF2 transcription factors.

Targeted CRISPR-Cas9-mediated invalidation of NLUCAT1 in LUAD cells attenuates their oncogenic potential

To decipher the functional role of NLUCAT1, we developed a loss-of-function approach using CRISPR-Cas9-mediated deletion of the first three exons in the A549 cell line (Fig. 5a). After selection, individual clones were

screened for the complete deletion of the three alleles (Fig. 5b), which totally abolished *NLUCAT1* transcription, as shown by RNA-seq, northern blotting, and RNA-fluorescent in situ hybridization (FISH) (Fig. 5a, c, d). The consequences of NLUCAT1 knockout were then analyzed on distinct cellular processes using several independent clones. Cell growth was slightly but significantly (20%) decreased in NLUCAT1-deleted clones (Fig. 6a, b). We also found an increase in the basal production of reactive oxygen species (ROS) in NLUCAT1-deleted cells (Fig. 6c). Cisplatin is part of the standard first-line therapy for lung cancer and mediates DNA damage associated with mitochondrial oxidative stress [33, 34]. In this context, drug treatment produced 1.5-fold more ROS in NLUCAT1-invalidated cells than in their wild-type (WT) counterparts (Fig. 6c). Accordingly, mitochondrial superoxide production was potentiated in deleted clones in the presence of cisplatin (Supplementary Fig. S5A). Consequently, loss of NLUCAT1 expression strongly sensitized A549 cells to cisplatin-induced apoptosis visualized by caspase-3 enzymatic assay (Fig. 6d). This was reverted by an antioxidant treatment with *N*-acetyl

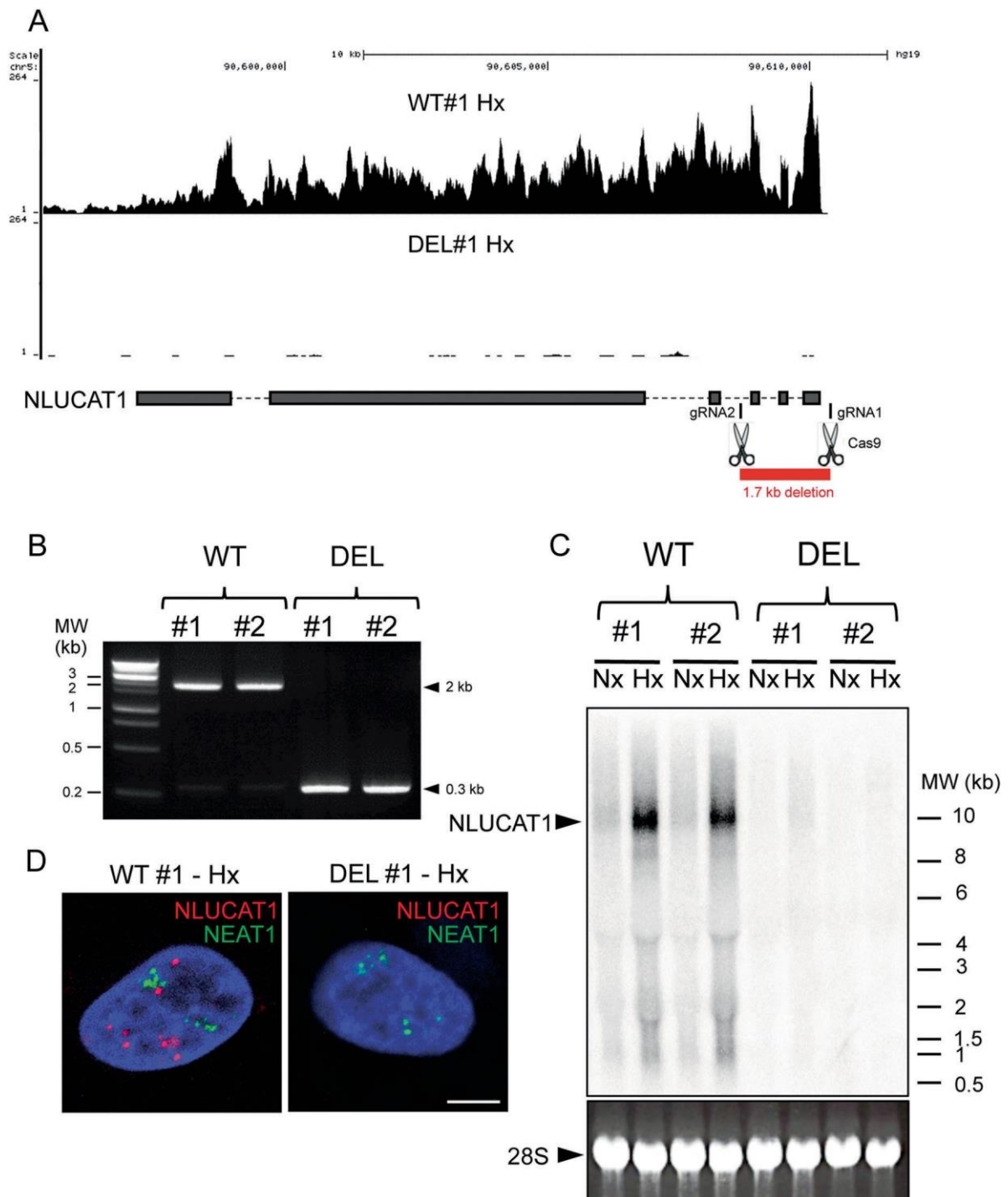
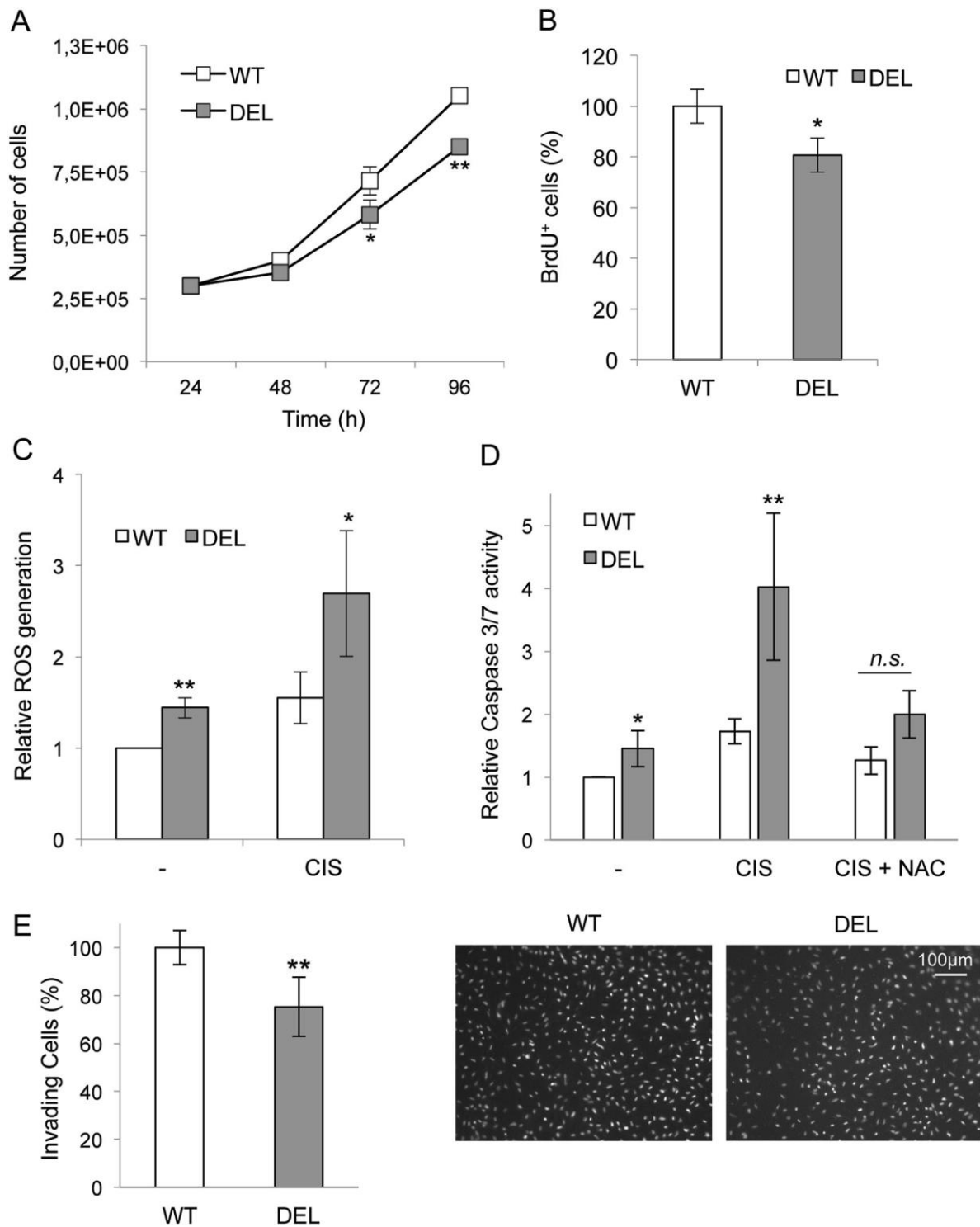


Fig. 5 CRISPR-Cas9 invalidation of the *LUCAT1/Lnc-ARRDC3-1* locus. a RNA-seq profile of reads mapped on the *LUCAT1/Lnc-ARRDC3-1* locus in wild-type (WT) and deleted (DEL) clones of A549 cells under hypoxia. An illustration of the strategy used for CRISPR-Cas9-mediated genomic deletion of the exon 1–3 region of the *LUCAT1/Lnc-ARRDC3-1* locus is shown (gRNA: Cas9 RNA guide). b PCR analysis of the 1.7 kb genomic deletion in two WT and

two DEL clones of A549 cells. c Northern blotting analysis of NLUCAT1 10 kb transcript expression in two WT and two DEL clones of A549 cells under normoxic or hypoxic conditions using a [<sup>32</sup>P]-labeled exon 6 cDNA probe. d Single-molecule RNA-FISH of NLUCAT1 and NEAT1 in WT and DEL clones of A549 cells under hypoxic conditions. Scale bar corresponds to 5  $\mu$ m

cysteine (NAC) (Fig. 6d). Accordingly, NLUCAT1-deleted cells were more sensitive to ROS-dependent apoptosis induced by hydrogen peroxide or menadione (Supplementary Fig. S5C). In addition, NLUCAT1 invalidation slowed down

the migratory and invasive properties of A549 cells (Supplementary Fig. S5B and Fig. 6e).



In a recent report, Thai et al. [25] suggested that the siRNA-mediated knockdown of the short LUCAT1 isoform SCAL1/LUCAT1:4 alters cell viability in the presence of cigarette smoke extract. We showed here that this minor isoform is cytosolic (Supplementary Fig. S6A) and confirmed that it is sensitive to two distinct siRNAs, whereas the major large nuclear NLUCAT1 transcript is not affected

by this treatment (Supplementary Fig. S6B). Moreover, in our hands, siRNA-mediated inhibition of SCAL1/LUCAT1:4 had no impact on cell proliferation as well as on cisplatin-induced oxidative stress and apoptosis (Supplementary Fig. S6C, D, E). Furthermore, the stable re-expression of LUCAT1:4 in CRISPR-Cas9-deleted A549 clones showed no significant effect on cell proliferation,



◀ Fig. 6 NLUCAT1 knockout affects proliferation, ROS production, cisplatin sensitivity, and invasion of LUAD cells. Effect of NLUCAT1 invalidation on: a cell proliferation: cells were plated (250,000/well) in triplicate and counted at the indicated times. Data are means  $\pm$  SD from two individual WT and two individual CRISPR-Cas9 DEL clones of A549 cells, and are representative of three independent experiments. b Percentage of cells in S phase: cells were plated (250,000/well), incubated for 48 h to reach exponential growth and pulsed for 4 h with BrdU. Cells were labeled with anti-BrdU and nuclei were counterstained with DAPI. Data are expressed as percentage of BrdU-positive cells determined in four independent microscope fields. Data are means  $\pm$  SD from five WT and five DEL clones of A549 cells, and are representative of three independent experiments. c ROS production: 48 h after plating, cells (200,000) were treated or not with 20  $\mu$ M cisplatin for 24 h, then labeled with fluorescent CM-H2DCF-DA probe and analyzed by cytometry for intracellular ROS production as indicated in the Materials and Methods. Data are means  $\pm$  SE from four independent experiments performed on three WT and three DEL A549 clones. d Caspase activity: 48 h after plating, cells (200,000) were treated or not with 20  $\mu$ M cisplatin for 48 h in the presence or absence of NAC (1 mM), then lysed and analyzed for caspase-3/7 activity as indicated in the Materials and Methods. Data are means  $\pm$  SE from three independent experiments performed on three WT and three DEL A549 clones. e Cell Invasion: cells (50,000) from WT or DEL clones were seeded in serum-free medium in the upper chambers of Matrigel-coated inserts and attracted in the lower chambers with medium containing 10% FCS and 5 ng/ml EGF. After 24 h, invading cells were scored. Data are means  $\pm$  SD from two independent experiments performed on three WT and three DEL A549 clones. Statistically significant differences are indicated (\* $p$  < 0.05, \*\* $p$  < 0.01)

ROS production, and caspase-3 activity (Supplementary Fig. S7).

Taken together, these data clearly indicate that the major large nuclear NLUCAT1 isoform promoted an aggressive phenotype in LUAD, affecting the proliferative, cell survival, and invasive properties of tumor cells.

### Impact of NLUCAT1 invalidation on the transcriptome of A549 cells

To gain further insight in the molecular mechanisms controlled by NLUCAT1, we performed transcriptomic profiling and we identified 354 genes for which the expression was significantly altered in NLUCAT1-deleted cells compared with WT clones (Fig. 7a and Supplementary Table S4). These genes were mainly (63%) repressed in NLUCAT1-deleted cells. According to our functional studies, gene annotation indicated that the modulated genes were significantly associated with Gene Ontology terms such as proliferation, invasion, apoptosis, inflammatory response, and oxidative stress (Fig. 7a, right panel).

We further validated a subset of 28 genes in independent experiments by qPCR (Supplementary Fig. S8A). Importantly, these genes were neither modulated by LUCAT1:4 RNAi in A549 cells nor by its stable re-expression in CRISPR-Cas9 *LUCAT1/Lnc-ARRDC3-1* locus-deleted clones (Supplementary Fig. S8B, C), demonstrating that

the loss of short LUCAT1:4 variant expression was not involved in the gene signature alteration observed in the *NLUCAT1* locus-deleted clones.

Among the 354 genes altered by NLUCAT1 deletion, 187 were dysregulated both in normoxic and hypoxic conditions (Fig. 7b). Although the global hypoxic response appeared largely unaffected by NLUCAT1 invalidation (Fig. 7b), the transcriptome of NLUCAT1-deficient cells revealed repressed gene networks predicted to be controlled by the HIF1A, NRF2, and NF- $\kappa$ B transcription factors (Fig. 7c), suggesting that NLUCAT1 could exert a positive feedback on these pathways. Accordingly, concerning HIF1A, we identified 52 hypoxia-regulated genes that were altered by NLUCAT1 invalidation (Supplementary Fig. S9A). Most of these genes were upregulated under hypoxic conditions in WT cells and markedly repressed in deleted cells (Supplementary Fig. S9B, C). Among these, we notably found the well-known hypoxia-regulated gene *CA9* and the glucose transporter *GLUT3 (SLC2A3)* (Supplementary Fig. S9C).

As NLUCAT1 knockout favored ROS production and cisplatin-induced ROS-dependent apoptosis (Fig. 6c, d and Supplementary Fig. S5A, C), we focused more specifically on NRF2, which is crucial for the maintenance of oxidative homeostasis through both hypoxia and redox signaling [31, 35]. Among the NRF2- and/or cisplatin-regulated gene networks, we identified four genes that were downregulated in NLUCAT1-knockout cells: *ALDH3A1* (aldehyde dehydrogenase 3A1), *GPX2* (glutathione peroxidase 2), *GLRX* (glutaredoxin), and *PDK4* (pyruvate dehydrogenase kinase 4) (Fig. 7d, Fig. S8A, Fig. 8a). Three of these genes products, GPX2, GLRX, and PDK4, were upregulated by hypoxia in WT cells (Fig. 8a). These enzymes either counterbalance at different levels ROS-induced damage to proteins, lipids, and DNA, or play a role in chemoresistance [36–40].

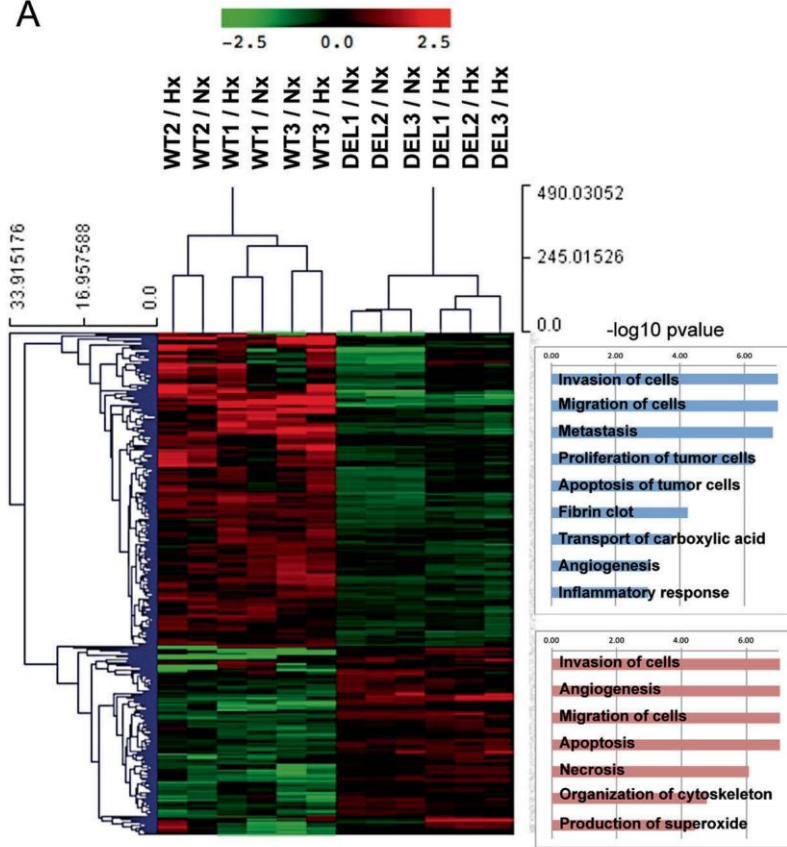
We demonstrated that the concomitant siRNA-mediated knockdown of *ALDH3A1*, *GPX2*, *GLRX*, and *PDK4* significantly increased the production of ROS in LUAD cells (Fig. 8b), and consequently the ROS-dependent caspase-3 activation induced by cisplatin, hydrogen peroxide, or menadione, thus partially mimicking the consequences of NLUCAT1 inactivation (Fig. 8c, d).

These data indicate that the NLUCAT1-dependent regulation of these genes in LUAD cells would most likely participate in redox signaling and maintenance of oxidative homeostasis.

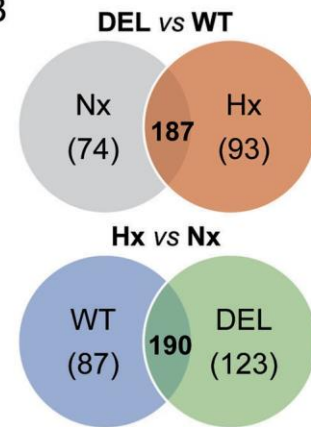
### Discussion

Many lncRNAs are aberrantly expressed in cancer and are involved as regulators at various stages of tumorigenesis [41]. Growing evidence supports their involvement in

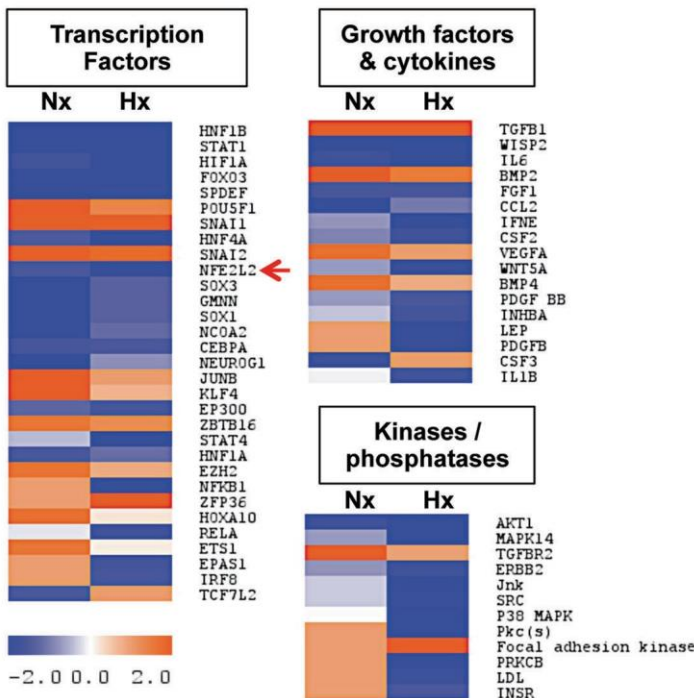
A



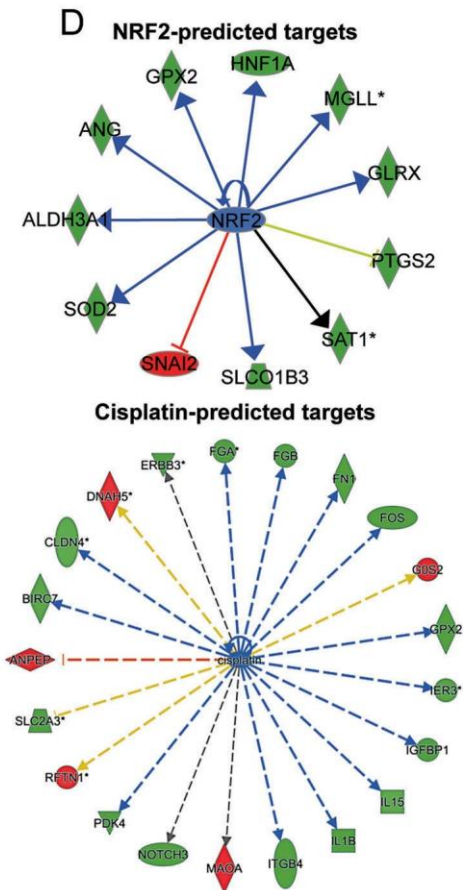
B



C



D



hypoxia-regulated tumor aggressiveness. Indeed, various lncRNAs, including NEAT1, H19, HOTAIR, and Linc-p21, are dysregulated in response to different cellular stresses, including hypoxia, and contribute to tumor hallmarks [20].

In the present study, we identified new “hypoxia-lncRNAs”, eight of which were correlated with the hypoxic status of LUADs and also modulated by hypoxia in vitro in LUAD cell lines. We focused our attention on transcripts

Fig. 7 NLUCAT1 invalidation alters A549 cells transcriptome in normoxic and hypoxic conditions. a Unsupervised hierarchical clustering of differentially expressed genes between three WT and three DEL clones of A549 cells. The gene expression profiles were determined on cells cultured for 24 h in normoxic or hypoxic conditions using pan genomic microarrays. The heat map shows the 354 genes for which expression is statistically modulated by NLUCAT1 invalidation based on a  $\log_2(\text{intensity}) > 6$ , an absolute  $\log_2(\text{fold change}) > 1$ , and an adj.  $P$ -value  $< 0.01$ . Expression corresponds to  $\log_2$  intensity followed by median centering. Clustering was performed using a Manhattan distance metric and average linkage. Gene annotation was done using the IPA software. b Venn diagrams comparing the distribution of genes that are (i) altered by NLUCAT1 deletion under normoxia or hypoxia ( $n = 354$ ), or (ii) regulated by hypoxia in WT and/or DEL cells ( $n = 400$ ,  $\log_2(\text{intensity}) > 6$ , an absolute  $\log_2(\text{fold change}) > 1$ , and an adj.  $P$ -value  $< 0.05$ ). c Heatmaps of predicted upstream regulators for the comparison DEL vs. WT in normoxic or hypoxic conditions. The red arrow indicates an inhibition of NRF2-regulated transcripts. d Two examples of the most significant IPA NRF2- and cisplatin-regulated gene networks altered by NLUCAT1 invalidation. Genes repressed or upregulated in the signature are indicated in green and red, respectively. Orange and blue indicate activation or inhibition of the pathway, respectively

from the *LUCAT1/Lnc-ARRDC3-1* locus, because (i) the signals detected by the multiple microarray probes indicated that this was the best candidate upregulated by hypoxia in vivo and in vitro, (ii) it was validated as a “hypoxia-lncRNA” in LUAD and LUSC TCGA cohorts, and (iii) its expression was associated with lower OS.

Short-read RNA-seq profiling combined with de novo transcript-reconstruction software has provided a first draft of the lncRNA world, but this characterization remains largely incomplete [42–45]. Despite curation of cDNAs and Expressed Sequence Tags (ESTs) mainly through GENCODE annotations [46, 47], information regarding transcript structure, in particular splice junctions and boundaries, lacks precision. The *LUCAT1* locus-associated transcripts represent a good example of the complexity of the lncRNA annotation situation. The LNCipedia database proposes 25 possible transcripts, annotated *LUCAT1:1–1:25* [48]. Only one transcript has been functionally annotated so far and corresponds to the 890 nt lncRNA named *LUCAT1:4/SCAL1* [25]. We first fully characterized the sequence, structure, and subcellular localization of the different transcript(s) produced from the *LUCAT1/Lnc-ARRDC3-1* locus.

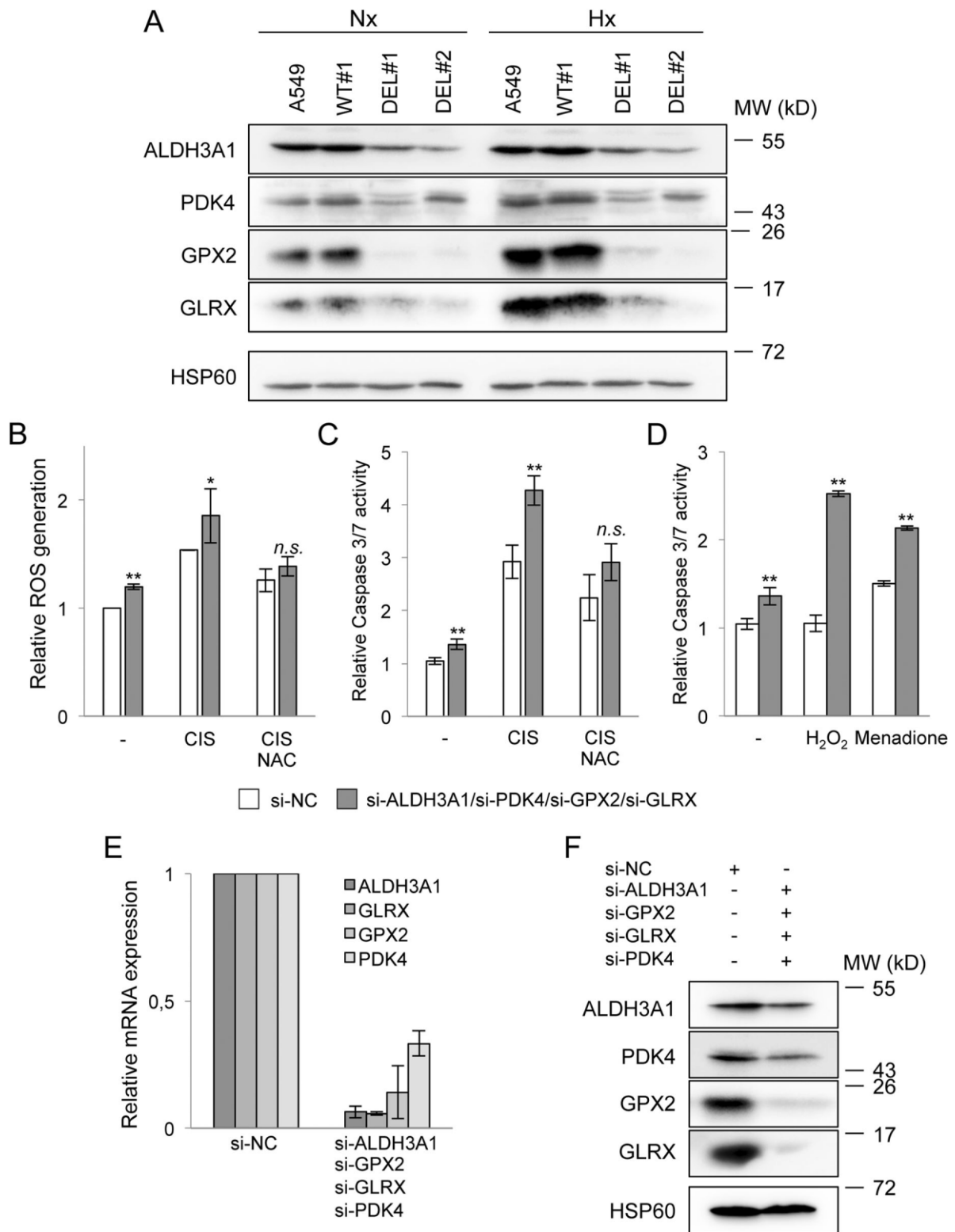
Our data confirmed the existence of the *LUCAT1:4/SCAL1* transcript in LUAD cells, but they clearly show that this represents a minor transcript product in terms of abundance, in both basal and hypoxic conditions. The major transcript corresponds to a large nuclear transcript of 9807 nt with a distinct exon structure and boundaries when compared with the 25 transcripts previously annotated so far. We thus propose the name NLUCAT1 (which may be also temporally annotated *LUCAT1:26*), but it appears to be

critical in the future to curate most of the transcripts annotated in the different databases.

Cancer development is a complex multistep process in which hypoxic, inflammatory, and oxidative stress responses are interconnected, notably because of significant crosstalk between the HIF, NF- $\kappa$ B, and NRF2 transcription factors [49, 50]. Inflamed tissues frequently present highly hypoxic regions, and vice versa, which results in ROS production [51]. HIF signaling is intricately linked to that of NF- $\kappa$ B during hypoxic inflammation and oxidative stress activates NRF2 and NF- $\kappa$ B, as well as HIFs, in non-hypoxic conditions [30, 49, 52–56]. In this context, NLUCAT1 appears to be at the crossroad of these pathways, both by its transcriptional regulation and by its functional role. Indeed, we showed that in addition to being a “hypoxia-linc,” NLUCAT1 is also induced by NF- $\kappa$ B-dependent IL-1 $\beta$  proinflammatory signal. Furthermore, we demonstrated that the hypoxia-dependent upregulation of NLUCAT1 is at least in part dependent on NRF2 transcription factor. These data are supported by a previous study showing that NRF2 mediates the cigarette smoke-induced expression of *LUCAT1:4/SCAL1* short variant in airway epithelial cells [25]. In addition, *LUCAT1:4/SCAL1* is repressed by RNAi-mediated silencing of NRF2 in A549 cells and the *LUCAT1* transcripts were found to be associated with the mutation status of KEAP1 and NRF2 in LUADs and LUSCs, respectively [57, 58].

Functional annotations of gene signature alterations associated with NLUCAT1 loss-of-function revealed repressed gene networks controlled by the HIF, NRF2, and NF- $\kappa$ B transcription factors, suggesting that the NLUCAT1 transcript may exert a positive feedback on these pathways by a mechanism that remains to be elucidated.

Platinum-based therapy is the first-line treatment for the majority of NSCLC patients. The cytotoxic effects of cisplatin are mainly mediated by DNA-damage response but also by mitochondrial oxidative stress [33, 34]. However, resistance to these chemotherapies is frequent and increasing evidence indicates that local tumor conditions, including a hypoxic microenvironment, may be involved [59]. We showed that NLUCAT1 expression is increased in hypoxic conditions. This observation is consistent with the increased expression of *LUCAT1* observed in a cisplatin-resistant human ovarian carcinoma cell line [60]. In addition, we showed that targeted deletion of NLUCAT1 decreased cell proliferation and invasion, while increasing oxidative stress-dependent sensitivity to cisplatin-induced apoptosis. By contrast, and contrary to a previously published report using an siRNA-based approach [61], we found that the short variant *LUCAT1:4/SCAL1* was unable to affect these parameters. Altogether, our dataset strongly suggests that NLUCAT1 is a member of the expanding family of



lncRNAs such as HOTAIR, UCA1, and H19 that are involved in platinum-based cancer drug resistance [62, 63].

We then focused on four genes of the NRF2- and/or cisplatin-regulated networks, all of which were

downregulated in NLUCAT1-deleted cells, and which participate in drug-resistance mechanisms or redox signaling and maintenance of oxidative homeostasis to protect cells from ROS-mediated damage. In this context, GPX2



Fig. 8 Concomitant knockdown of ALDH3A1, GPX2, GLRX, and PDK4 partially mimics NLUCAT1 invalidation. a Western blotting analysis of ALDH3A1, GPX2, GLRX, and PDK4 expression in A549, WT, and DEL clones in normoxia or hypoxia. HSP60 was used as a loading control. b, c Effect of the concomitant RNA interference of ALDH3A1, GPX2, GLRX, and PDK4 on cisplatin-induced ROS production and caspase activity. A549 cells were plated and transfected 24 h later with 5 nM each of ALDH3A1, PDK4, GPX2, and GLRX siRNA or 20 nM non-relevant si-NC. Forty-eight hours after transfection, cells were treated or not with 20  $\mu$ M cisplatin in the presence or absence of NAC (1 mM) for 24 h, after which ROS production and caspase-3/7 activity were assessed. Data are means  $\pm$  SD of duplicates from two (ROS) or six (caspase) independent experiments, respectively. d Effect of the concomitant RNA interference of ALDH3A1, GPX2, GLRX, and PDK4 on caspase activity induced by oxidative stress: 48 h after transfection, cells were treated or not with 1 mM H<sub>2</sub>O<sub>2</sub> or 100  $\mu$ M menadione for 24 h, then lysed and analyzed for caspase-3/7 activity. Data are means  $\pm$  SD of duplicates from three independent experiments. e, f Efficiency of concomitant RNA interference for ALDH3A1, PDK4, GPX2, and GLRX (5 nM of each) was analyzed for transcript and protein expression 48 h after transfection. Statistically significant differences are indicated (\* $p$  < 0.05, \*\* $p$  < 0.01)

catalyzes the reduction of hydrogen peroxide in the cytoplasm, while glutathione is oxidized to glutathione disulfide [36], thus protecting against ROS-dependent apoptosis [64, 65]. GLRX exerts its protective role against oxidative stress through the reduction of protein oxidation called deglutathionylation [37, 66]. GLRX is, notably, known to protect epithelial cells from oxidative damage via the prevention of AKT glutathionylation [37, 66]. Accordingly, GLRX downregulation could participate in the repression of an AKT-regulated pro-survival network in NLUCAT1-deleted cells (Fig. 7c). Our third candidate, ALDH3A1, neutralizes highly reactive lipid peroxidation-derived aldehydes and attenuates ROS-mediated protein modification and genotoxin-induced DNA damage and apoptosis [38, 67, 68]. Interestingly, high expression of ALDH3A1 in cancer has been associated with regulation of cell proliferation and drug resistance to ROS-dependent cyclophosphamide- and 5-fluorouracil-mediated cell death [68, 69]. Finally, PDK4, one of the four PDK family members, may favor the metabolic switch to aerobic glycolysis and chemoresistance in cancers via inhibition of the pyruvate dehydrogenase complex [39, 40]. Knockdown or inhibition of PDK4 reduced cell migration and invasion, while increasing sensitivity to 5-Fluorouracil or cisplatin in colon and/or bladder cancer cells, confirming this pro-tumorigenic function [39, 40, 70].

In agreement with these important cancer-associated functions, we demonstrated that the concomitant knockdown of these four genes in LUAD cells increased ROS-dependent apoptosis, thus partially mimicking the consequences of NLUCAT1 inactivation.

Taken together, these data suggest that dysregulation of redox signaling in NLUCAT1-deleted LUAD cells may be

likely to result from an adaptation to overcome the lack of expression of NLUCAT1, which could slow proliferative and invasive properties, and sensitize cancer cells to ROS-dependent apoptosis.

Further experiments are required to fully characterize the underlying molecular mode of action of NLUCAT1 in LUAD. LUCAT1:4 short isoform has been proposed to favor cell cycle progression through epigenetic silencing of p21 and p57 cyclin-dependent kinase inhibitors due to its association with the histone methyltransferase PRC2 (polycomb repressive complex 2) [61]. This putative function is not consistent with our data, since in our hands NLUCAT1 appeared rather as a negative regulator of PRC2 signaling, as indicated in our transcriptomic analyses with a significant positive  $z$ -score association for EZH2 (functional enzymatic component of PRC2) in NLUCAT1-deleted cells (Fig. 7c). Moreover, although a large set of LincRNAs has been proposed to guide chromatin-modifying complexes, notably PRC2, to specific genomic loci to regulate gene expression [71], recent evidence has questioned this link, notably through the demonstration of PRC2 promiscuous RNA binding in vitro [72, 73]. Of note, the transcriptional repression induced by HOTAIR is not dependent on PRC2 as previously postulated [74]. LincRNA loci can, however, also influence the expression of neighboring genes in *cis* and some can function as enhancers, termed eRNAs, stimulating gene transcription by serving as the *cis*-regulator [75, 76]. Histone marks obtained from ChIP-sequencing assays on A549 cells of the ENCODE project indicate that NLUCAT1 is transcribed from regulatory regions enriched for H3K4me1 (mono-methylation of lysine 4 of the H3 histone protein) and H3K27ac (acetylation of lysine 27 of the H3 histone protein) marks (not shown) that are predominant histone modifications found at nucleosomes around enhancer elements. We found that CRISPR-Cas9-mediated deletion of the NLUCAT1 TSS region slightly repressed the expression of its neighboring gene *ARRDC3* in normoxic condition but not under hypoxia. This does not support an eRNA-mediated mode of action but rather suggests that the deletion of a specific DNA regulatory element located in the vicinity of NLUCAT1 TSS influences the basal expression of *ARRDC3*.

In summary, we provide evidence that NLUCAT1 is a new hypoxia- and inflammation-driven large nuclear lincRNA in LUAD. Targeted deletion of NLUCAT1 using a CRISPR-CAS9 strategy in the A549 cell line revealed a decrease in proliferative and invasive properties, an increase in oxidative stress, and a higher sensitivity to cisplatin-induced apoptosis, demonstrating that NLUCAT1 represents a novel coordinator of the hypoxic, inflammatory, and antioxidant responses in LUADs.

## Material and methods

### Cell culture and siRNA transfection

The A549, H1975, and H2228 human LUAD cell lines were obtained from ATCC (Rockville, MD). Cells were grown in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum (Dominique Dutscher, Brumath, France). A triple gas incubator set at 1.2% oxygen, 94% nitrogen, and 5% carbon dioxide was used for hypoxic conditions. Cell lines were authenticated by PCR-single-locus technology (Eurofins, Ebersberg, Germany) and are routinely tested for mycoplasma contamination.

ON-Targetplus siRNAs (for NRF2, PDK4, GPX2, GLRX, ALDH3A1, HIF1A, HIF2A, LUCAT1:4, and negative control) were purchased from Dharmacon (Lafayette, CO). Two distinct siRNAs were used for each gene (Supplementary Table S5). Cells were plated and transfected 24 h later at 30–50% confluency, with siRNAs (5–10 nM), using Lipofectamin™ RNAiMAX reagent according to the manufacturer's procedure (Thermo Fisher Scientific, Courtaboeuf, France).

### Local LUAD patient cohort used for transcriptomic profiling

Fifty-seven LUAD patients hospitalized between 2004 and 2010 at the Pasteur Hospital (Departments of Pulmonary Medicine and Thoracic Surgery, CHU de Nice, France) were enrolled in this study. The diagnosis of LUAD was based on examination of all tumor specimens using the seventh pTNM classification and on the last histological classification of NSCLC. Written informed consent was obtained from participants after explaining the nature of the study, which was approved by the research ethics board of the Nice University hospital and was performed according to the guidelines of the Declaration of Helsinki. The main clinical and pathological data are summarized in Supplementary Table S1. Eleven normal lung tissue specimens were taken from areas at standard distance (3 cm) from the same cohort of patients.

### Microarray hybridization and statistical analyses

Transcriptomic analyses of LUAD samples and A549 cells were performed on SurePrint G3 Human Gene Expression v2 8 × 60 K microarrays from Agilent Technology (Loveland, CO, USA) as previously described [77].

Data analyses were performed using R. We used the Bioconductor package array QualityMetrics and custom R scripts for quality control. Additional analyses were done using Bioconductor package Limma. Data were normalized using the quantile method. A linear modeling approach was

used to calculate log ratios, moderated *t*-statistics, and *P*-values for each comparison of interest. *P*-values were adjusted for multiple testing using the Benjamini–Hochberg method that controls the false discovery rate.

First, we isolated lncRNAs that are differentially expressed between 57 LUAD samples and 11 peritumoral healthy lung tissues (Dataset 1). LUAD hypoxic status was determined a posteriori based on a selection of 27 genes extracted from a hypoxia-derived metagene signature [24] using a score corresponding to the number of genes with a fold change > 2 compared with the mean average expression in healthy tissues (from 0 to 27: 0–9: “no”; 10–18: “intermediate”; 19–27: “high” hypoxic). A principal component analysis strategy using this hypoxic score was then performed to calculate the correlation between hypoxia and the log-transformed gene expression values of each probe, as previously described [78]. Spearman's rank correlation was applied and probes were ranked according to *p*-values.

Two biological replicates were performed for comparison of A549 cells cultured in normoxic and hypoxic conditions (Dataset 2).

For NLUCAT1 WT and -deleted A549 clones, three distinct clones of each phenotype cultured in normoxic or hypoxic conditions were compared (Dataset 5). Enrichment in biological themes (Molecular function, Upstream regulators, and canonical pathways) and biological networks analysis were performed using Ingenuity Pathway Analysis software (<http://www.ingenuity.com/>). Hierarchical clusterings were performed with the MultiExperiment Viewer (MeV) program version 4.9, using a Manhattan distance metric and average linkage.

### TCGA dataset analyses

We used LUAD and LUSC datasets from TCGA. RNA-seq and survival data were available for 522 and 504 patients, respectively (Supplementary Table S3). RNA-seq data were normalized using the Bioconductor package DESeq2. The hypoxic status for each sample was determined using the same method as for the local cohort. Patients were separated into high-, intermediate-, and no hypoxic status groups. Survival curves were generated using the Kaplan–Meier method and significance was assessed using the log-rank test.

RNA extraction and gene expression analyses by qPCR, RT-PCR, and northern blotting

Total RNAs were isolated with TRIzol® reagent (Thermo Fisher Scientific) or RNAeasy kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Gene expression was measured by qPCR as previously described [79] (Supplementary Table S6). For standard

RT-PCR, genes were amplified by GoTaq DNA polymerase (Promega, Madison, WI, USA) using 100 ng cDNA and analyzed on agarose gels. All primer sequences corresponding to the different genes analyzed are in Supplementary Table S7. Northern blotting were done with 20 µg of total RNAs and quantified on a Fujifilm FLA-5100 phosphorimager.

#### Subcellular fractionation

Cytosolic and nuclear fractions were prepared as previously described [80], and cytosolic and nuclear RNAs were extracted with TRIzol® and TRIzol®LS reagents (Thermo Fisher Scientific), respectively.

#### Single-molecule RNA fluorescence in situ hybridization

LncRNA-FISH was performed using RNAscope Multiplex Fluorescent assays (Advanced Cell Diagnostics, Hayward, CA) according to the manufacturer's protocol. Two distinct sets of short oligonucleotide probes were designed for NLUCAT1: (i) a specific NLUCAT1 probe-set targeting the 3575–5860 sequence in the exon 5, and (ii) a probe-set targeting the full-length NLUCAT1 sequence (9807 bp) that detects NLUCAT1 and also the short LUCAT1:4 variant. Commercially available human *NEAT1* and bacterial *dapB* probe sets were used as positive and negative controls, respectively.

#### Chromatin immunoprecipitation

ChIP was performed using the iDeal ChIP-qPCR kit from Diagenode (Belgium) according to the manufacturer's protocol. Briefly, cells were exposed to hypoxia (1% O<sub>2</sub>) or IL-1β (20 ng/ml) and then cross-linked with ChIP cross-link Gold reagent (Diagenode) for 30 min and 1% formaldehyde for 15 min. Cross-linking was stopped with 0.125 M glycine. Cells were collected by scrapping, lysed at 4 °C, and resuspended in shearing buffer iS1b containing protease inhibitor cocktail. After sonication using the Bioruptor (Diagenode) to yield DNA fragments of 200–1000 bp, lysates were cleared by centrifugation at 16,000 × g for 10 min. For each immunoprecipitation assay, lysates corresponding to 4 × 10<sup>6</sup> cells were incubated with 5 µg of rabbit antibodies pre-adsorbed on Diamag protein A-coated magnetic beads and incubated overnight at 4 °C with rotation.

After washing and proteinase digestion, the immunoprecipitates were de-cross-linked 15 min at 100 °C and genomic DNA was purified. The primers used for the amplification of the *NLUCAT1* promoter are listed in Supplementary Table S8.

#### CRISPR-Cas9-mediated invalidation of NLUCAT1

The sequence targeted for CRISPR-Cas9 genome modification was identified using the online design tool (<http://crispr.mit.edu/>). Two single-guide RNAs (sgRNAs) were selected: gRNA1, located 172 bp upstream of the TSS of *NLUCAT1*, and gRNA2, located in intron 3, 1518 bp downstream of the TSS. Sense and antisense oligonucleotides were synthesized containing the gRNA sequence, an additional 5'-end guanine nucleotide and the *BbsI* overhangs (Supplementary Table S9). The annealed sgRNAs were ligated into the *BbsI*-digested pSpCas9(BB)-2A-Puro (pX459) V2.0 plasmid (Adgene, Cambridge, MA). Inserts were sequenced for final validation. A549 cells (1.8 × 10<sup>6</sup>) were reverse transfected with 7.5 µg each of pX459-gRNA1 and pX459-gRNA2 constructs using Lipofectamine LTX® reagent (Thermo Fisher Scientific), according to the manufacturers' procedures. Selection with puromycin (2 µg/ml) was performed 24 h after transfection for at least 48 h. Cells were then cloned by serial dilution and amplified individual clones were screened for complete tri-allelic deletion on 50 ng genomic DNA using three qPCR assays (Supplementary Table S10).

#### RNA sequencing

##### Solid protocol (Dataset 3)

Two micrograms of total RNAs from normoxic or hypoxic A549 cells were depleted from ribosomal RNA with the Ribo-Zero™ rRNA removal kit (Epicentre, Madison, WI) and libraries were processed following the NEBNext® mRNA Library Prep Set for SOLiD™ (New England Biolabs, Ipswich, MA), then amplified, quantified with the Bioanalyzer High Sensitivity DNA Kit (Agilent), and sequenced on SOLiD 5500XL (Thermo Fisher Scientific) with single-end 50 bp reads. Reads were aligned to the human genome release hg19 with the LifeScope software v2.5.1 (Thermo Fisher Scientific) using whole transcriptome pipeline for RNA-seq libraries with default parameters.

##### Illumina protocol (Dataset 4)

Libraries were generated from 500 ng of total RNA from WT and CRISPR-Cas9-deleted A549 clones in hypoxic conditions using the TruSeq Stranded Total RNA Library Prep kit (Illumina, San Diego, CA) according to the manufacturer's instructions. Libraries were then quantified with a KAPA library quantification kit (Roche, Basel, Switzerland) and pooled. Four nanomoles of this pool were loaded onto a high-output flowcell and sequenced on a NextSeq500 sequencer (Illumina) with 2 × 75 bp paired-end chemistry. Reads were aligned to the human genome release hg19 using STAR v2.4.0a with default parameters.

## Cloning procedures and stable lentiviral transduction

The short 890 nt LUCAT1:4 variant (NR\_103548) was synthesized as a gBlock® double-stranded DNA fragment (Integrated DNA Technologies, Leuven, Belgium) and ligated into the pLJM1 vector (Addgene) between AgeI and EcoRI sites.

Lentiviral particles were produced in 293TN cells using the pPACKH1 HIV Lentivector Packaging Kit (System Biosciences, Palo Alto, CA) with Lipofectamine LTX Plus reagent according to manufacturers' procedures. The lentiviral particles were then collected, mixed with PEG-it virus precipitation kit (System Biosciences), and concentrated 100-fold. Titration of lentiviral particles was performed on HELA cells using the Global UltraRapid™ Lentiviral Titer Kit (System Biosciences).

A549 cells ( $8 \times 10^5$ ) were transduced with pLJM1-LUCAT1:4 lentiviral particles (multiplicity of infection 5) in DMEM containing 8 µg/ml polybrene. Cells were kept under puromycin selection (2 µg/ml) 48 h after transduction until only transduced cells remained.

## ROS and superoxide production

ROS and superoxide were measured by flow cytometry with the cell permeant 2',7'-dichlorodihydrofluorescein diacetate (CM-H2DCF-DA) and mitoSOX™ red fluorescent probes, respectively (Thermo Fisher Scientific). Cells were incubated in the dark with 5 µM CM-H2DCF-DA or mitoSOX™ red in Hank's buffered salt solution for 45 min at 37 °C, trypsinized, washed twice, and detection of oxidized highly fluorescent 2',7'-dichlorofluorescein or mitoSOX™ red was performed on a flow cytometer (LSR II Fortessa, BD Biosciences) using excitation wavelengths of 488 nm and 510 nm, respectively (emission 525 nm and 580 nm, respectively).

## Caspase assays

The activation of executioner caspase-3 in A549 cells was determined with the Caspase-Glo 3/7 Assay kit (Promega) according to the manufacturer's instructions. Cells were lysed in passive lysis buffer and the protein concentration determined using the Pierce™ BCA Protein Assay kit (Thermo Fisher Scientific). Cell lysates (5 µg of proteins) were incubated with the caspase substrates and luminescence was quantified after 30 min and normalized to the amount of proteins.

## Migration and invasion assays

Boyden's chamber assays were performed as described previously [81]. Invasion was assessed using a commercially available 24-well BioCoat Matrigel Invasion Chamber (BD Biosciences).

## Statistical analyses

All experimental samples were included in the final analyses. Unless indicated otherwise, all data were representative of at least two independent experiments and expressed as means  $\pm$  SD. Before statistical analysis, data were checked for normal distribution (Shapiro–Wilk's test) and comparable variance (F-test for equality of variances for data with normal distribution). Comparisons between two groups of normally distributed data with equal variances were performed using the unpaired two-sided Student's *t*-test. Differences were considered statistically significant at  $p < 0.05$ . Sample size (*n*), statistical tests, and *p*-values are indicated in the figure legends.

## Antibodies

For ChIP experiments, rabbit anti-HIF-2 $\alpha$  (ab199) was from Abcam (Cambridge, UK), normal rabbit IgG and rabbit anti-NRF2 (D179C-XP) were from Cell Signaling Technology (Danvers, MA, USA), and rabbit polyclonal antibodies to RelA (sc-372), NF- $\kappa$ B p50 (sc7178x), and c-Rel (sc-71  $\times$ ) were from Santa Cruz Biotechnology (Heidelberg, Germany). For western blotting analyses, goat polyclonal antibody to Hsp60 (sc-1052) and mouse monoclonal antibodies to GLRX (sc-293250) and ALDH3A1 (sc-376089) were purchased from Santa Cruz Biotechnology. Anti-PDK4 (AP7041B) was from Abgent (San Diego, CA, USA) and anti-GPX2 (GTX100292) from GeneTex (Irvine, CA, USA). Secondary horseradish peroxidase-conjugated antibodies were from Dako (Copenhagen, Denmark).

## Data availability

The experimental data from microarray and RNA sequencing have been deposited in the [NCBI Gene Expression Omnibus \(GEO\)](#) database under SuperSerie GSE117049 containing 5 distinct datasets.

**Acknowledgements** We gratefully acknowledge the outstanding technical support of the UCA GenomiX platform and MICA imaging facility of the University Côte d'Azur, and the staff from the CHU Nice Biobank. This work was supported by Plan Cancer 2018 « ARN non-codants en cancérologie: du fondamental au translationnel » (number 18CN045), Cancéropole PACA, Fondation ARC pour la Recherche sur le Cancer, Fondation Unice (AIR project), Ligue contre le cancer (comité départemental du Nord), and the French Government (Agence Nationale de Recherche, ANR) through the Investments for the Future LABEX SIGNALIFE (ANR-11-LABX-0028-01) and FRANCE GENOMIQUE (ANR-10-INBS-09-03 and ANR-10-INBS-09-02). LML was a recipient of the Fondation pour la Recherche Médicale.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest. **Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



## References

1. Chang MY, Mentzer SJ, Colson YL, Linden PA, Jaklitsch MT, Lipsitz SR, et al. Factors predicting poor survival after resection of stage IA non-small cell lung cancer. *J Thorac Cardiovasc Surg.* 2007;134:850–6.
2. Ou SH, Zell JA, Ziogas A, Anton-Culver H. Prognostic factors for survival of stage I nonsmall cell lung cancer patients: a population-based analysis of 19,702 stage I patients in the California Cancer Registry from 1989 to 2003. *Cancer.* 2007;110:1532–41.
3. Pouyssegur J, Dayan F, Mazure NM. Hypoxia signalling in cancer and approaches to enforce tumour regression. *Nature.* 2006;441:437–43.
4. Holohan C, Van Schaeybroeck S, Longley DB, Johnston PG. Cancer drug resistance: an evolving paradigm. *Nat Rev Cancer.* 2013;13:714–26.
5. Akagi I, Okayama H, Schetter AJ, Robles AI, Kohno T, Bowman ED, et al. Combination of protein coding and non-coding gene expression as a robust prognostic classifier in stage I lung adenocarcinoma. *Cancer Res.* 2013;73:3821–32.
6. Ilie M, Mazure NM, Hofman V, Ammadi RE, Ortholan C, Bonnetaud C, et al. High levels of carbonic anhydrase IX in tumour tissue and plasma are biomarkers of poor prognostic in patients with non-small cell lung cancer. *Br J Cancer.* 2010;102:1627–35.
7. Gee HE, Ivan C, Calin GA, Ivan M. HypoxamiRs and cancer: from biology to targeted therapy. *Antioxid Redox Signal.* 2014;21:1220–38.
8. Bertero T, Rezzonico R, Pottier N, Mari B. Impact of microRNAs in the cellular response to hypoxia. *Int Rev Cell Mol Biol.* 2017;333:91–158.
9. Puissegur MP, Mazure NM, Bertero T, Pradelli L, Grosso S, Robbe-Sermesant K, et al. miR-210 is overexpressed in late stages of lung cancer and mediates mitochondrial alterations associated with modulation of HIF-1 activity. *Cell Death Differ.* 2011;18:465–78.
10. Grosso S, Doyen J, Parks SK, Bertero T, Paye A, Cardinaud B, et al. MiR-210 promotes a hypoxic phenotype and increases radioresistance in human lung cancer cell lines. *Cell Death Dis.* 2013;4:e544.
11. Guttman M, Rinn JL. Modular regulatory principles of large non-coding RNAs. *Nature.* 2012;482:339–46.
12. Ulitsky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. *Cell.* 2013;154:26–46.
13. Quinn JJ, Chang HY. Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet.* 2016;17:47–62.
14. Du Z, Fei T, Verhaak RG, Su Z, Zhang Y, Brown M, et al. Integrative genomic analyses reveal clinically relevant long non-coding RNAs in human cancer. *Nat Struct Mol Biol.* 2013;20:908–13.
15. Schmitt AM, Chang HY. Long noncoding RNAs in cancer pathways. *Cancer Cell.* 2016;29:452–63.
16. Yang F, Huo XS, Yuan SX, Zhang L, Zhou WP, Wang F, et al. Repression of the long noncoding RNA-LET by histone deacetylase 3 contributes to hypoxia-mediated metastasis. *Mol Cell.* 2013;49:1083–96.
17. Ferdin J, Nishida N, Wu X, Nicoloso MS, Shah MY, Devlin C, et al. HINCUTs in cancer: hypoxia-induced noncoding ultra-conserved transcripts. *Cell Death Differ.* 2013;20:1675–87.
18. Yang F, Zhang H, Mei Y, Wu M. Reciprocal regulation of HIF-1 and lincRNA-p21 modulates the Warburg effect. *Mol Cell.* 2014;53:88–100.
19. Lin A, Li C, Xing Z, Hu Q, Liang K, Han L, et al. The LINK-A lincRNA activates normoxic HIF1alpha signalling in triple-negative breast cancer. *Nat Cell Biol.* 2016;18:213–24.
20. Chang YN, Zhang K, Hu ZM, Qi HX, Shi ZM, Han XH, et al. Hypoxia-regulated lincRNAs in cancer. *Gene.* 2016;575:1–8.
21. Schmidt LH, Spieker T, Koschmieder S, Schaffers S, Humberg J, Jungen D, et al. The long noncoding MALAT-1 RNA indicates a poor prognosis in non-small cell lung cancer and induces migration and tumor growth. *J Thorac Oncol.* 2011;6:1984–92.
22. Loewen G, Jayawickramarajah J, Zhuo Y, Shan B. Functions of lincRNA HOTAIR in lung cancer. *J Hematol Oncol.* 2014;7:90.
23. Dong J, Xu J, Wang X, Jin B. Influence of the interaction between long noncoding RNAs and hypoxia on tumorigenesis. *Tumour Biol.* 2016;37:1379–85.
24. Buffa FM, Harris AL, West CM, Miller CJ. Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene. *Br J Cancer.* 2010;102:428–35.
25. Thai P, Statt S, Chen CH, Liang E, Campbell C, Wu R. Characterization of a novel long non-coding RNA, SCAL1, induced by cigarette smoke and elevated in lung cancer cell lines. *Am J Respir Cell Mol Biol.* 2013;49:204–11.
26. Volders PJ, Hensens K, Wang X, Menten B, Martens L, Gevaert K, et al. LNCipedia: a database for annotated human lincRNA transcript sequences and structures. *Nucleic Acids Res.* 2013;41:D246–D251.
27. Clark MB, Johnston RL, Inostroza-Ponta M, Fox AH, Fortini E, Moscato P, et al. Genome-wide analysis of long noncoding RNA stability. *Genome Res.* 2012;22:885–98.
28. Tani H, Mizutani R, Salam KA, Tano K, Ijiri K, Wakamatsu A, et al. Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. *Genome Res.* 2012;22:947–56.
29. Fitzpatrick SF, Tambuwala MM, Bruning U, Schaible B, Scholz CC, Byrne A, et al. An intact canonical NF-kappaB pathway is required for inflammatory gene expression in response to hypoxia. *J Immunol.* 2011;186:1091–6.
30. Rius J, Guma M, Schachtrup C, Akassoglou K, Zinkernagel AS, Nizet V, et al. NF-kappaB links innate immunity to the hypoxic response through transcriptional regulation of HIF-1alpha. *Nature.* 2008;453:807–11.
31. Toth RK, Warfel NA. Strange bedfellows: nuclear factor, erythroid 2-like 2 (Nrf2) and hypoxia-inducible factor 1 (HIF-1) in tumor hypoxia. *Antioxidants (Basel).* 2017;6:E27.
32. Burke JR, Pattoli MA, Gregor KR, Brassil PJ, MacMaster JF, McIntyre KW, et al. BMS-345541 is a highly selective inhibitor of I kappa B kinase that binds at an allosteric site of the enzyme and blocks NF-kappa B-dependent transcription in mice. *J Biol Chem.* 2003;278:1450–6.
33. Choi YM, Kim HK, Shim W, Anwar MA, Kwon JW, Kwon HK, et al. Mechanism of cisplatin-induced cytotoxicity is correlated to impaired metabolism due to mitochondrial ROS generation. *PLoS ONE.* 2015;10:e0135083.
34. Marullo R, Werner E, Degtyareva N, Moore B, Altavilla G, Ramalingam SS, et al. Cisplatin induces a mitochondrial-ROS response that contributes to cytotoxicity depending on mitochondrial redox status and bioenergetic functions. *PLoS ONE.* 2013;8:e8116.
35. Shelton P, Jaiswal AK. The transcription factor NF-E2-related factor 2 (Nrf2): a protooncogene? *FASEB J.* 2013;27:414–23.
36. Singh A, Rangasamy T, Thimmulappa RK, Lee H, Osburn WO, Brigelius-Flohe R, et al. Glutathione peroxidase 2, the major cigarette smoke-inducible isoform of GPX in lungs, is regulated by Nrf2. *Am J Respir Cell Mol Biol.* 2006;35:639–50.
37. Liu X, Xavier C, Jann J, Wu H. Salvianolic acid B (Sal B) protects retinal pigment epithelial cells from oxidative stress-induced cell death by activating glutaredoxin 1 (Grx1). *Int J Mol Sci.* 2016;17:E1835.
38. Duong HQ, You KS, Oh S, Kwak SJ, Seong YS. Silencing of NRF2 reduces the expression of ALDH1A1 and ALDH3A1 and sensitizes to 5-FU in pancreatic cancer cells. *Antioxidants (Basel).* 2017;6:E52.

39. Woolbright BL, Choudhary D, Mikhalyuk A, Trammel C, Shanmugam S, Abbott E, et al. The role of pyruvate dehydrogenase kinase-4 (PDK4) in bladder cancer and chemoresistance. *Mol Cancer Ther*. 2018;17:2004–12.
40. Zhang Y, Zhang Y, Geng L, Yi H, Huo W, Talmon G, et al. Transforming growth factor beta mediates drug resistance by regulating the expression of pyruvate dehydrogenase kinase 4 in colorectal cancer. *J Biol Chem*. 2016;291:17405–16.
41. Wu X, Tudoran OM, Calin GA, Ivan M. The many faces of long noncoding RNAs in cancer. *Antioxid Redox Signal*. 2017;29:922–35.
42. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic non-coding RNAs reveals global properties and specific subclasses. *Genes Dev*. 2011;25:1915–27.
43. Hangauer MJ, Vaughn IW, McManus MT. Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. *PLoS Genet*. 2013;9:e1003569.
44. Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, et al. The landscape of long noncoding RNAs in the human transcriptome. *Nat Genet*. 2015;47:199–208.
45. Fang S, Zhang L, Guo J, Niu Y, Wu Y, Li H, et al. NON-CODEV5: a comprehensive annotation database for long non-coding RNAs. *Nucleic Acids Res*. 2018;46:D308–D314.
46. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res*. 2012;22:1760–74.
47. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res*. 2012;22:1775–89.
48. Volders PJ, Verheggen K, Menschaert G, Vandepoele K, Martens L, Vandesompele J, et al. An update on LNCipedia: a database for annotated human lncRNA sequences. *Nucleic Acids Res*. 2015;43:4363–4.
49. Reuter S, Gupta SC, Chaturvedi MM, Aggarwal BB. Oxidative stress, inflammation, and cancer: how are they linked? *Free Radic Biol Med*. 2010;49:1603–16.
50. Tafani M, Sansone L, Limana F, Arcangeli T, De Santis E, Polese M, et al. The interplay of reactive oxygen species, hypoxia, inflammation, and sirtuins in cancer initiation and progression. *Oxid Med Cell Longev*. 2016;2016:3907147.
51. Eltzschig HK, Carmeliet P. Hypoxia and inflammation. *N Engl J Med*. 2011;364:656–65.
52. Gorlach A, Bonello S. The cross-talk between NF-kappaB and HIF-1: further evidence for a significant liaison. *Biochem J*. 2008;412:e17–19.
53. Oliver KM, Taylor CT, Cummins EP. Hypoxia. Regulation of NFkappaB signalling during inflammation: the role of hydroxylases. *Arthritis Res Ther*. 2009;11:215.
54. Patten DA, Lafleur VN, Robitaille GA, Chan DA, Giaccia AJ, Richard DE. Hypoxia-inducible factor-1 activation in nonhypoxic conditions: the essential role of mitochondrial-derived reactive oxygen species. *Mol Biol Cell*. 2010;21:3247–57.
55. Nakajima S, Kitamura M. Bidirectional regulation of NF-kappaB by reactive oxygen species: a role of unfolded protein response. *Free Radic Biol Med*. 2013;65:162–74.
56. Buelna-Chontal M, Zazueta C. Redox activation of Nrf2 & NF-kappaB: a double end sword? *Cell Signal*. 2013;25:2548–57.
57. Ashouri A, Sayin VI, Van den Eynden J, Singh SX, Papa-panakopoulos T, Larsson E. Pan-cancer transcriptomic analysis associates long non-coding RNAs with key mutational driver events. *Nat Commun*. 2016;7:13197.
58. White NM, Cabanski CR, Silva-Fisher JM, Dang HX, Govindan R, Maher CA. Transcriptome sequencing reveals
59. Galluzzi L, Vitale I, Michels J, Brenner C, Szabadkai G, Harel-Bellan A, et al. Systems biology of cisplatin resistance: past, present and future. *Cell Death Dis*. 2014;5:e1257.
60. Zheng ZG, Xu H, Suo SS, Xu XL, Ni MW, Gu LH, et al. The essential role of H19 contributing to cisplatin resistance by regulating glutathione metabolism in high-grade serous ovarian cancer. *Sci Rep*. 2016;6:26093.
61. Sun Y, Jin SD, Zhu Q, Han L, Feng J, Lu XY, et al. Long non-coding RNA LUCAT1 is associated with poor prognosis in human non-small lung cancer and regulates cell proliferation via epigenetically repressing p21 and p57 expression. *Oncotarget*. 2017;8:28297–311.
62. Heery R, Finn SP, Cuffe S, Gray SG. Long non-coding RNAs: key regulators of epithelial-mesenchymal transition, tumour drug resistance and cancer stem cells. *Cancers (Basel)*. 2017;9:38.
63. Hu Y, Zhu QN, Deng JL, Li ZX, Wang G, Zhu YS. Emerging role of long non-coding RNAs in cisplatin resistance. *Onco Targets Ther*. 2018;11:3185–94.
64. Yan W, Chen X. GPX2, a direct target of p63, inhibits oxidative stress-induced apoptosis in a p53-dependent manner. *J Biol Chem*. 2006;281:7856–62.
65. Naiki T, Naiki-Ito A, Asamoto M, Kawai N, Tozawa K, Etani T, et al. GPX2 overexpression is involved in cell proliferation and prognosis of castration-resistant prostate cancer. *Carcinogenesis*. 2014;35:1962–7.
66. Liu X, Jann J, Xavier C, Wu H. Glutaredoxin 1 (Grx1) protects human retinal pigment epithelial cells from oxidative damage by preventing AKT glutathionylation. *Invest Ophthalmol Vis Sci*. 2015;56:2821–32.
67. Lassen N, Pappa A, Black WJ, Jester JV, Day BJ, Min E, et al. Antioxidant function of corneal ALDH3A1 in cultured stromal fibroblasts. *Free Radic Biol Med*. 2006;41:1459–69.
68. Muzio G, Maggiora M, Paiuzzi E, Oraldi M, Canuto RA. Aldehyde dehydrogenases and cell proliferation. *Free Radic Biol Med*. 2012;52:735–46.
69. Jang JH, Bruse S, Liu Y, Duffy V, Zhang C, Oyamada N, et al. Aldehyde dehydrogenase 3A1 protects airway epithelial cells from cigarette smoke-induced DNA damage and cytotoxicity. *Free Radic Biol Med*. 2014;68:80–6.
70. Leclerc D, Pham DN, Levesque N, Truongcao M, Foulkes WD, Sapienza C, et al. Oncogenic role of PDK4 in human colon cancer cells. *Br J Cancer*. 2017;116:930–6.
71. Khalil AM, Guttman M, Huarte M, Raj A, Rivea Morales D, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci USA*. 2009;106:11667–72.
72. Davidovich C, Zheng L, Goodrich KJ, Cech TR. Promiscuous RNA binding by Polycomb repressive complex 2. *Nat Struct Mol Biol*. 2013;20:1250–7.
73. Blanco MR, Guttman M. Re-evaluating the foundations of lncRNA-Polycomb function. *EMBO J*. 2017;36:964–6.
74. Portoso M, Ragazzini R, Brencic Z, Moiani A, Michaud A, Vassilev I, et al. PRC2 is dispensable for HOTAIR-mediated transcriptional repression. *EMBO J*. 2017;36:981–94.
75. Engreitz JM, Haines JE, Perez EM, Munson G, Chen J, Kane M, et al. Local regulation of gene expression by lncRNA promoters, transcription and splicing. *Nature*. 2016;539:452–5.
76. Li W, Notani D, Rosenfeld MG. Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nat Rev Genet*. 2016;17:207–23.
77. Lino Cardenas CL, Henaoui IS, Courcot E, Roderburg C, Cauffiez C, Aubert S, et al. miR-199a-5p is upregulated during fibrogenic response to tissue injury and mediates TGFbeta-induced lung fibroblast activation by targeting caveolin-1. *PLoS Genet*. 2013;9: altered long intergenic non-coding RNAs in lung cancer. *Genome Biol*. 2014;15:429.
- motherapy crossresistance and treatment response in childhood acute lymphoblastic leukemia. *Cancer Cell*. 2005;7:375–86.

79. Bertero T, Gastaldi C, Bourget-Ponzio I, Mari B, Meneguzzi G, Barbry P, et al. CDC25A targeting by miR-483-3p decreases CCND-CDK4/6 assembly and contributes to cell cycle arrest. *Cell Death Differ.* 2013;20:800–11.
80. Wang Y, Zhu W, Levy DE. Nuclear and cytoplasmic mRNA quantification by SYBR green based real-time RT-PCR. *Methods.* 2006;39:356–62.
81. Gastaldi C, Bertero T, Xu N, Bourget-Ponzio I, Lebrigand K, Fourre S, et al. miR-193b/365a cluster controls progression of epidermal squamous cell carcinoma. *Carcinogenesis.* 2014;35:1110–20.

## Affiliations

Laura Moreno<sup>1,2</sup> · Leon<sup>1,2</sup> · Marine Gautier<sup>1,2</sup> · Richard Allan<sup>1,2</sup> · Marius Ilié<sup>2,3,4</sup> · Nicolas Nottet<sup>1,2</sup> · Nicolas Pons<sup>1,2</sup> · Agnes Paquet<sup>1,2</sup> · Kévin Lebrigand<sup>1,2</sup> · Marin Truchi<sup>1,2</sup> · Julien Fassy<sup>1,2</sup> · Virginie Magnone<sup>1,2</sup> · Garrett Kinnebrew<sup>5</sup> · Milan Radovich<sup>5</sup> · Meyling Hua-Chen Cheok<sup>6</sup> · Pascal Barbry<sup>1,2</sup> · Georges Vassaux<sup>2,7</sup> · Charles-Hugo Marquette<sup>2,3,8</sup> · Gilles Ponzio<sup>1,2</sup> · Mircea Ivan<sup>9</sup> · Nicolas Pottier<sup>10</sup> · Paul Hofman<sup>2,3,4</sup> · Bernard Mari<sup>1,2</sup> · Roger Rezzonico<sup>2,7</sup>

<sup>1</sup> Université Côte d'Azur, CNRS UMR7275, IPMC, Valbonne, France

<sup>2</sup> FHU-OncoAge, Nice, France

<sup>3</sup> Université Côte d'Azur, CNRS, INSERM, IRCAN, Nice, France

<sup>4</sup> Hospital-Related Biobank (BB-0033-00025), Pasteur Hospital, Nice, France

<sup>5</sup> Department of Surgery, Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN 46202, USA

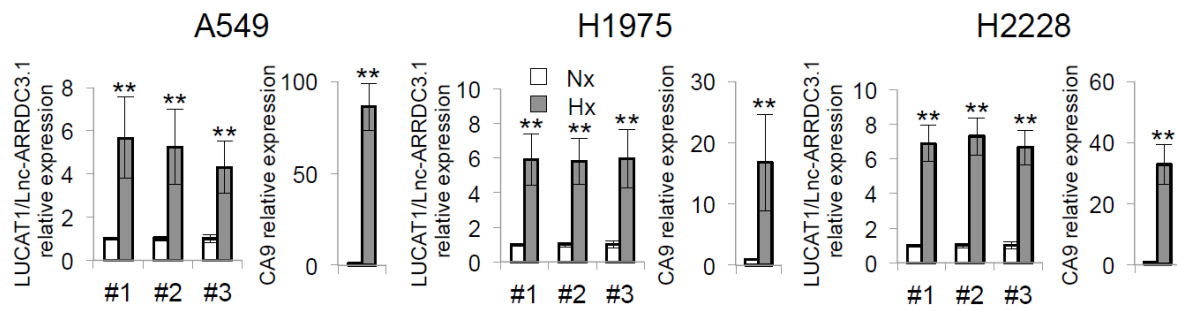
<sup>6</sup> INSERM UMR-S1172, Institute for Cancer Research of Lille, Factors of Leukemia Cell Persistence, Lille, Cedex, France

<sup>7</sup> Université Côte d'Azur, INSERM, CNRS UMR7275, IPMC, Valbonne, France

<sup>8</sup> Department of Pneumology, CHU-Nice, Nice, France

<sup>9</sup> Department of Medicine and Department of Microbiology and Immunology, Indiana University School of Medicine, Indianapolis, IN 46202, USA

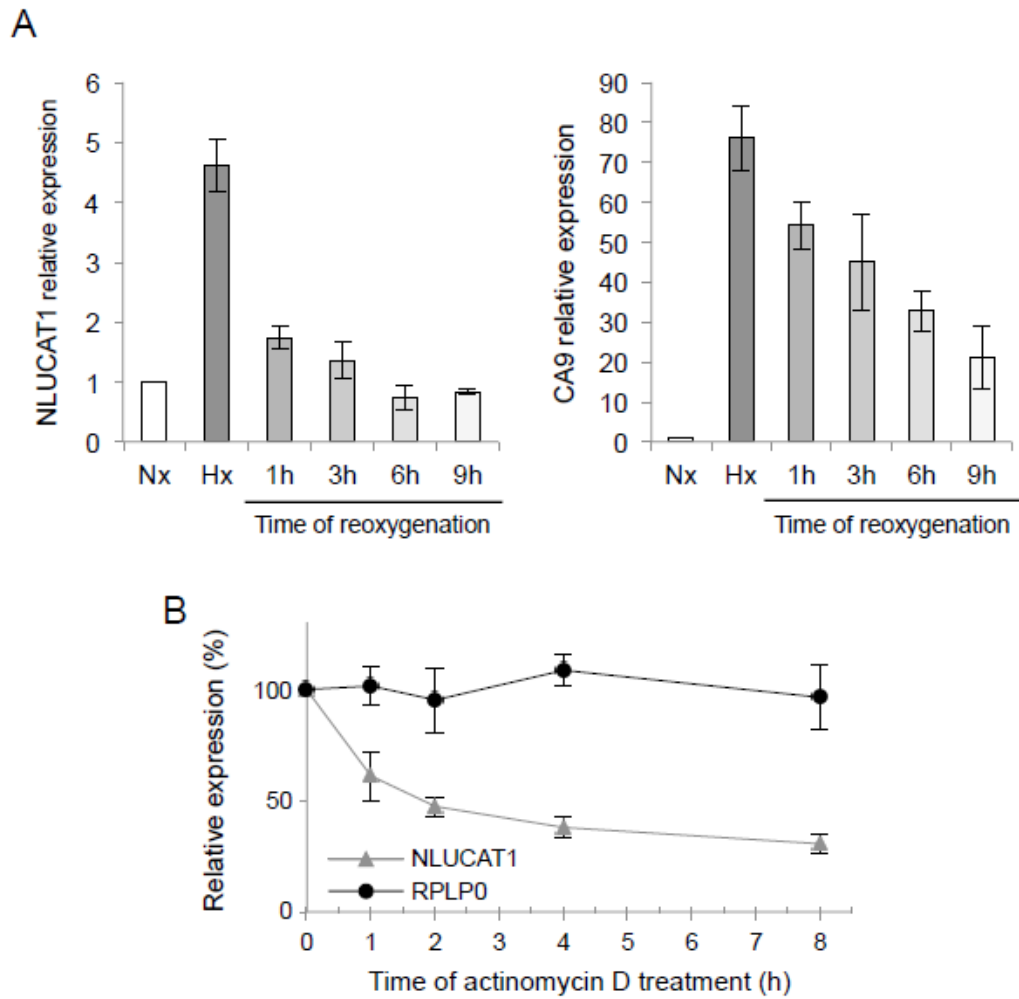
<sup>10</sup> EA4483, Faculté de Médecine de Lille, Pole Recherche, Lille, France



**Supplementary Figure S1: Hypoxia induces the expression of LUCAT1/Lnc-ARRDC3-1 in LUAD cell lines.**

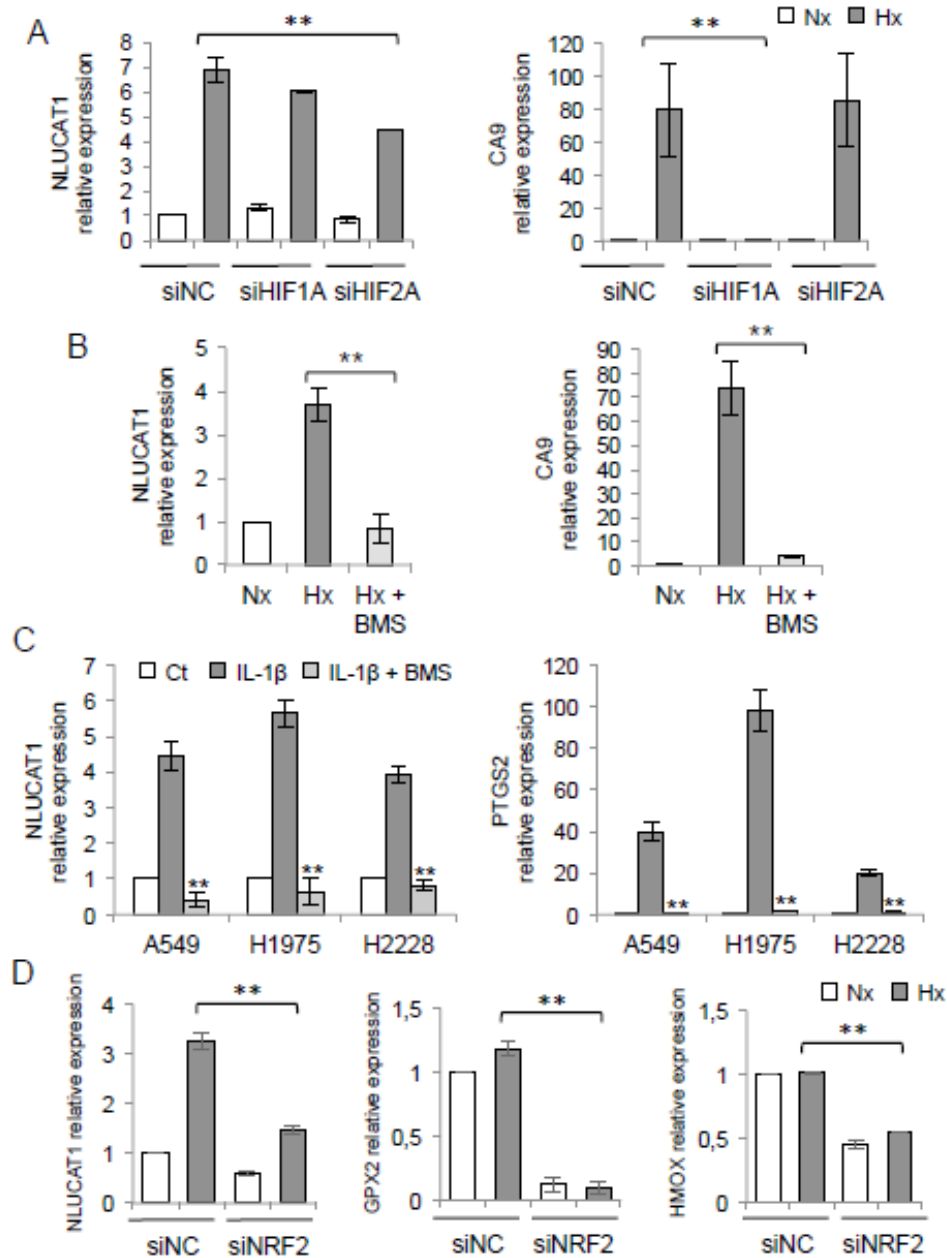
LUCAT1/Lnc-ARRDC3-1 expression was analyzed using 3 distinct qPCR assays in A549, H1975 and H2228 LUAD cell lines cultured for 24 h in normoxic or hypoxic conditions. CA9 gene expression was analyzed as a marker of the hypoxic response. Data are means  $\pm$  SD from 3 independent experiments. Statistically significant differences are indicated (\*\* $p < 0.01$ ).





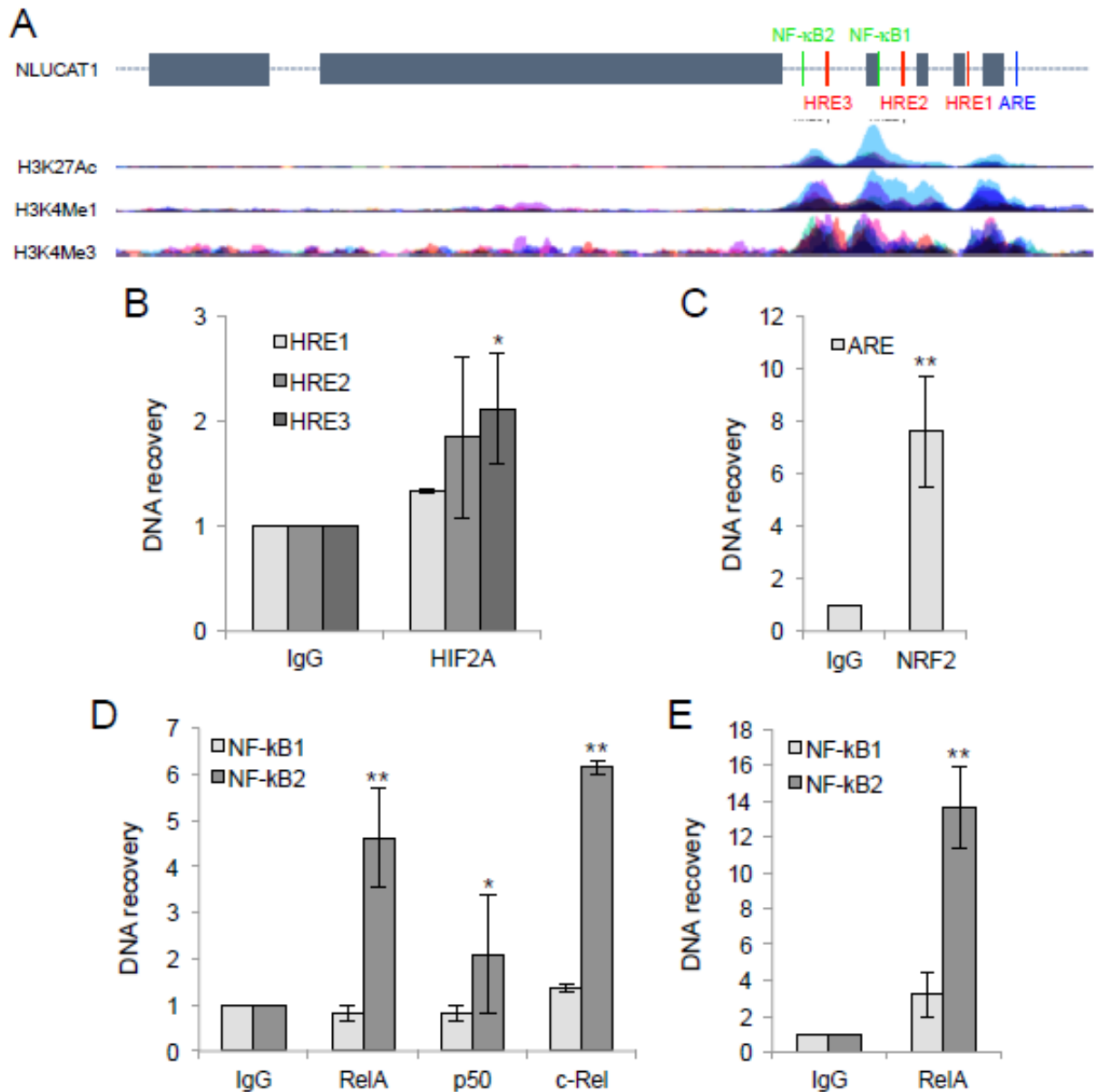
**Supplementary Figure S2: NLUCAT1 is an unstable transcript.**

**A.** A549 cells were cultured in normoxic or hypoxic conditions for 24 h and then reexposed to normoxic conditions for the indicated periods of time. RNA was extracted from the different experimental conditions and analyzed for NLUCAT1 and CA9 expression by RT-qPCR. Data are means  $\pm$  SD of triplicates and are representative of 2 experiments. **B.** Exponentially growing A549 cells were treated or not with actinomycin D (10  $\mu$ g/ml) for the indicated times, RNA extracted and analyzed for NLUCAT1 and RPLP0 expression by RT-qPCR. Data are means  $\pm$  SD from 3 independent experiments.



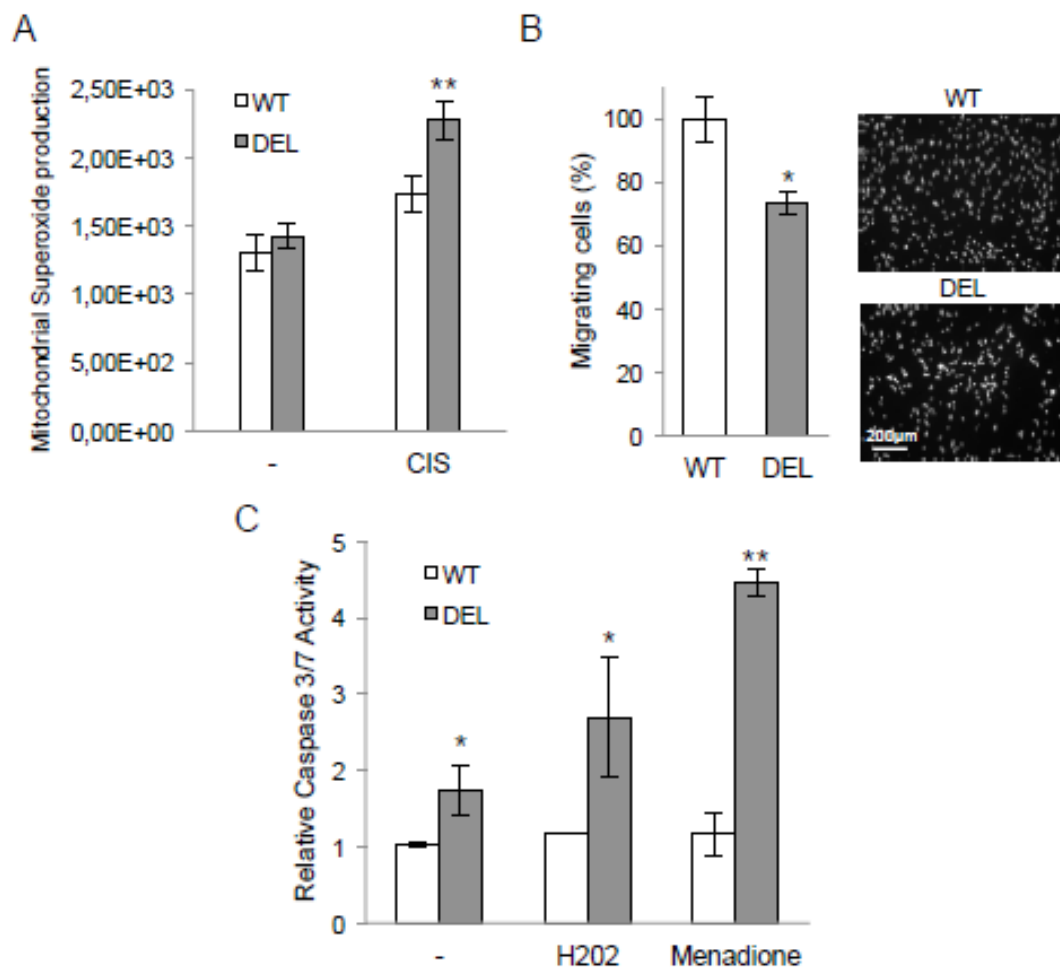
**Supplementary Figure S3: Transcriptional regulation of hypoxia-induced NLUCCAT1.**

**A.** Effect of HIF invalidation on hypoxia-induced NLUCCAT1 expression. A549 cells were transfected with 25 nM of HIF1A, HIF2A or non-relevant NC siRNAs. 24 h after transfection, cells were cultured in normoxia or hypoxia for 24 h. RNA was extracted and assessed for NLUCCAT1 and CA9 expression. Data are means  $\pm$  SD of 2 experiments. **B.** A549 cells were cultured under normoxia or hypoxia for 24 h in the presence or absence of BMS345541 (5  $\mu$ M). Data are means  $\pm$  SD of triplicates and are representative of 2 experiments. **C.** A549, H1975 and H2228 LUAD cell lines were stimulated with IL-1 $\beta$  (20 ng/ml) for 3 h in the presence or absence of BMS345541. PTGS2 gene expression was analyzed as a marker of the inflammatory response. Data are means  $\pm$  SD from 3 independent experiments. **D.** Effect of NRF2 invalidation on hypoxia-induced NLUCCAT1 expression. A549 cells were transfected with 10 nM of two distinct NRF2 or NC siRNAs. 24 h after transfection, cells were cultured in normoxia or hypoxia for 24 h. RNA was extracted and assessed for NLUCCAT1, GPX2 and HMOX expression. Data are means  $\pm$  SD from 2 independent experiments. Statistically significant differences are indicated (\*\* $p < 0.01$ ).



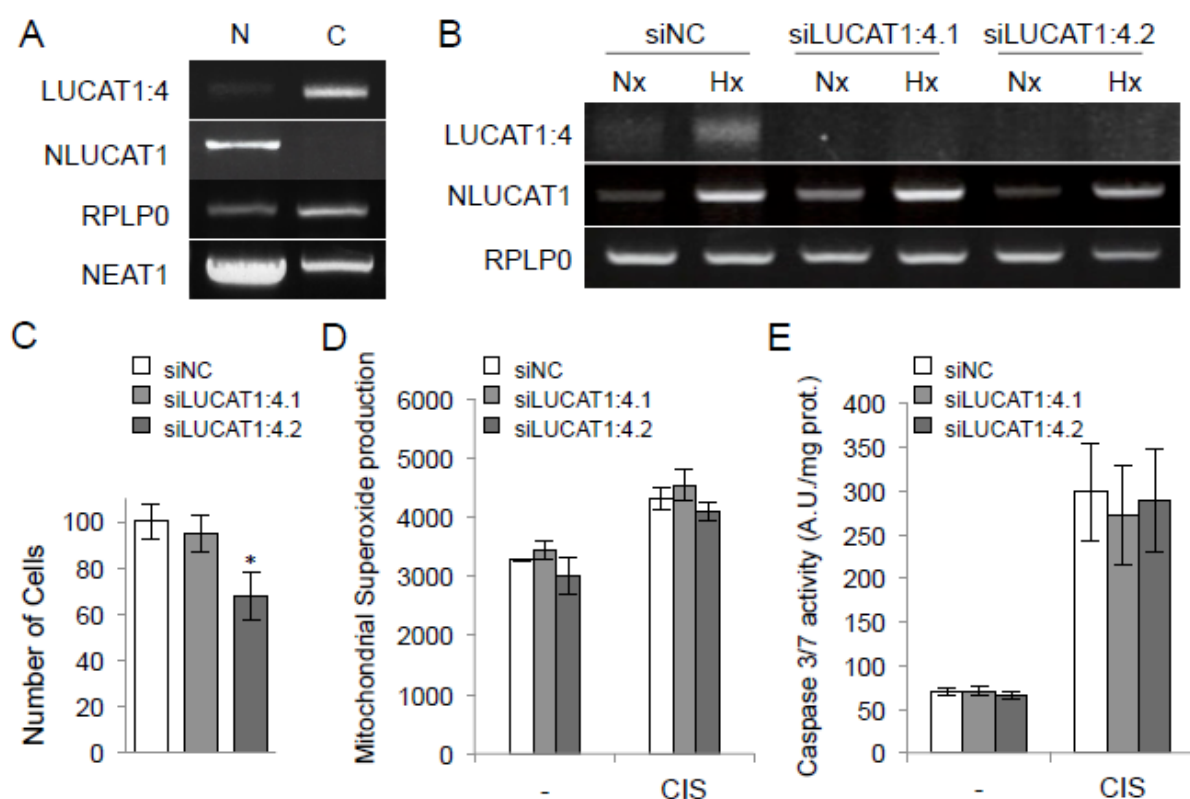
**Supplementary Figure S4: Confirmation of HIF-2 $\alpha$ , NRF2 and NF- $\kappa$ B binding to the promoter region of NLUCAT1 in hypoxic LUAD cells.**

**A.** Schematic representation of HIF, NRF2 and NF- $\kappa$ B putative binding sites in the promoter region of *NLUCAT1*. H3K27Ac, H3K4Me1 and H3K4Me3 marks on 7 cell lines from the ENCODE project are shown. **B-D.** A549 cells were cultured under hypoxia for 24 h (**BC**) or for 4 h (**D**) and sonicated chromatin was immunoprecipitated with anti-HIF-2 $\alpha$  (**B**), anti-NRF2 (**C**), or anti-NF- $\kappa$ B family members p65 RelA, p50 and c-Rel (**D**) antibodies and negative control IgG. **E.** A549 cells were stimulated with IL-1 $\beta$  (20 ng/ml) for 1 h and assayed for anti-p65 RelA ChIP. Real-time qPCR analyses of *NLUCAT1* promoter sequences purified from the different ChIP experiments are compared with control IgG. Data are means  $\pm$  SD from 2 independent experiments. Statistically significant differences are indicated (\* $p$ <0.01, \*\* $p$ <0.001).



**Supplementary Figure S5: Effects of NLUCAT1 silencing in A549 cells.**

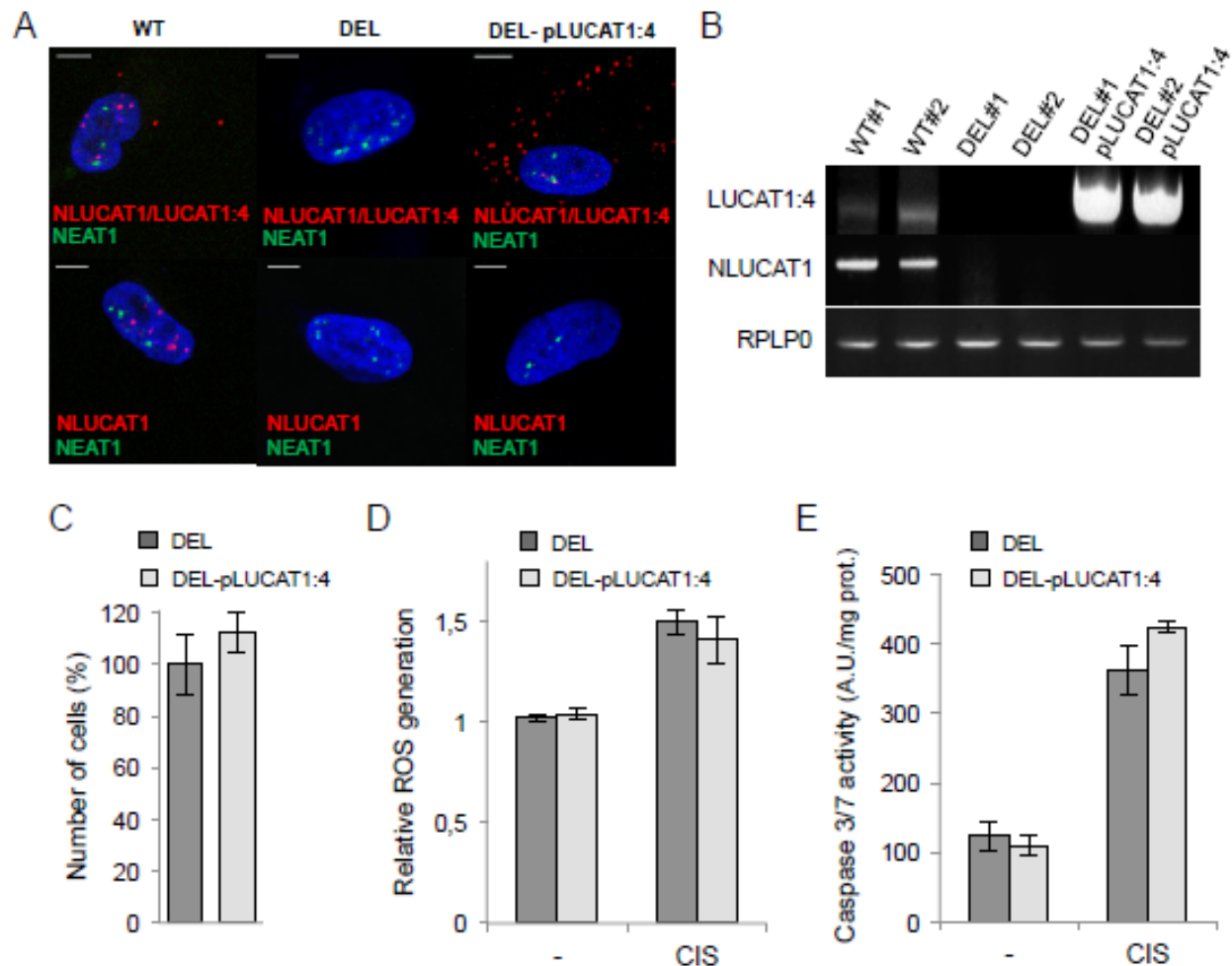
Effect of NLUCAT1 invalidation on: **A.** Mitochondrial superoxide production: 48 h after plating, cells were treated or not with 40  $\mu$ M Cisplatin for 24 h, then labelled with fluorescent mitoSOX™ red probe and analyzed by cytometry for mitochondrial superoxide production as indicated in the Materials and methods. Data are means  $\pm$  SD from 3 WT and 3 DEL A549 clones. **B.** Cell migration: Cells (50,000) from WT or DEL clones were seeded in serum-free medium in the upper chambers of inserts and attracted in the lower chambers with medium containing 10% FCS. After 24 h, migrating cells were scored. Data are means  $\pm$  SD from 2 independent experiments performed on 3 WT and 3 DEL A549 clones. **C.** Caspase activity induced by oxidative stress: 48 h after plating, cells were treated or not with 1mM H<sub>2</sub>O<sub>2</sub> or 100  $\mu$ M Menadione for 24h, then lysed and analyzed for caspase 3/7 activity. Data are means  $\pm$  SE from 2 independent experiments performed on 2 WT and 2 DEL A549 clones. Statistically significant differences are indicated (\* $p$ <0.05, \*\* $p$ <0.01).



**Supplementary Figure S6: Knock down of LUCAT1:4 short variant does not alter ROS production or cell survival.**

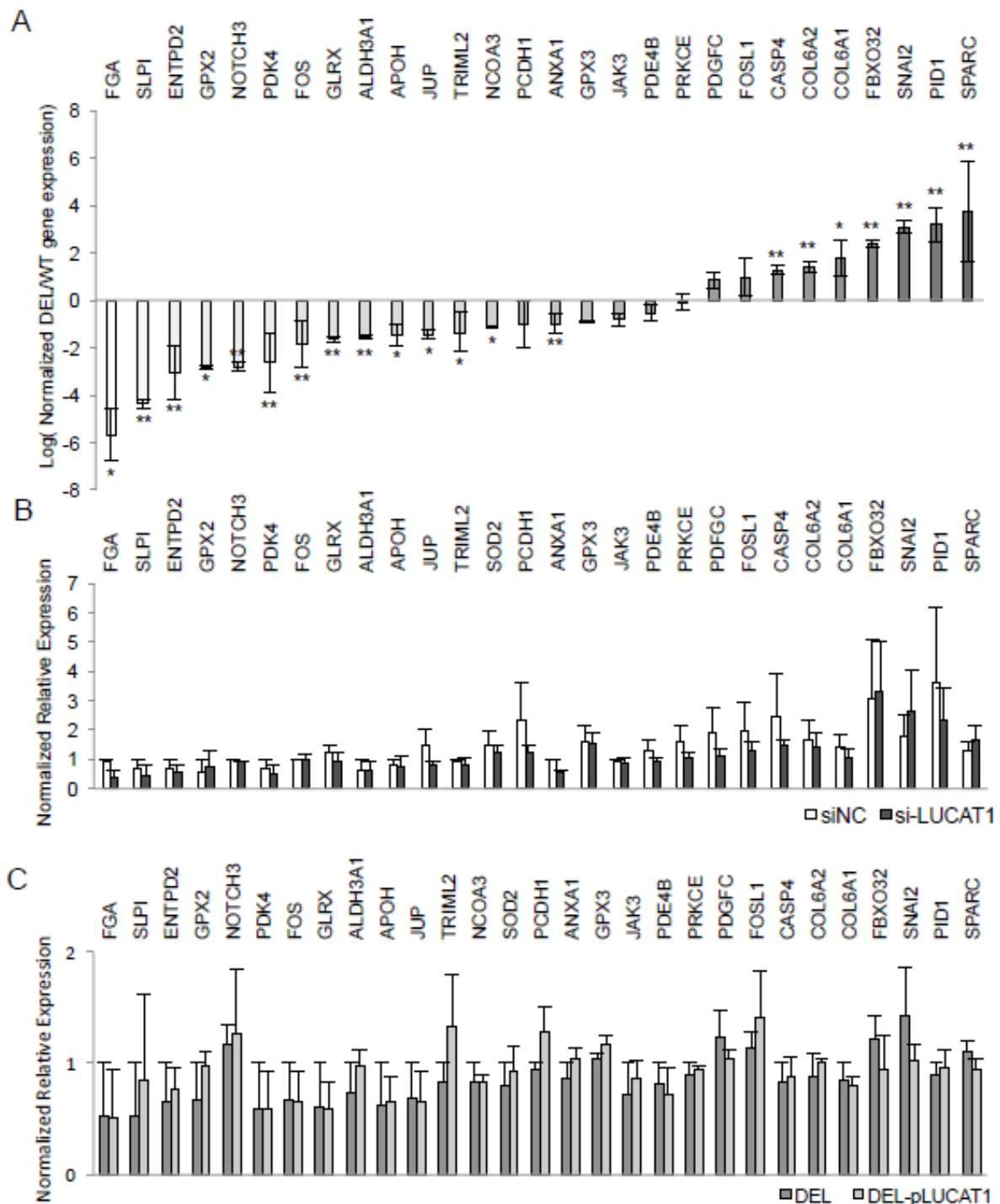
**A.** Subcellular localization of LUCAT1:4 and NLUCAT1 transcripts. RNA extracted from cytosolic and nuclear fractions of A549 cells were analyzed for LUCAT1:4 and NLUCAT1 expression. Data are representative of 2 experiments. **B.** Effects of LUCAT1:4 siRNAs on hypoxia-induced NLUCAT1 expression. A549 cells were plated and transfected 24 h later with 10 nM of 2 distinct LUCAT1:4 or non relevant NC siRNAs. 24 h after transfection, cells were cultured in normoxia or hypoxia for 24 h, and total RNA was extracted and assessed by RT-PCR for LUCAT1:4 and NLUCAT1 expression. Data are representative of 2 independent experiments. **C.** Effects of LUCAT1:4 siRNAs on cell proliferation. Cells were transfected as in B and counted 72 h post-transfection. Data are means  $\pm$  SD of triplicates from 2 independent experiments. **D.** Effects of LUCAT1:4 siRNAs on mitochondrial superoxide production. 24 h after transfection, cells were treated or not with 40  $\mu$ M Cisplatin for 24 h, then labelled with fluorescent mitoSOX™ red probe and analyzed by cytometry. **E.** Effects of LUCAT1:4 siRNAs on caspase activity. 24 h after transfection, cells were treated or not with 40  $\mu$ M Cisplatin for 48 h, then lysed and analyzed for caspase 3/7 activity. Data are means  $\pm$  SD from 2 independent experiments. Statistically significant differences are indicated (\* $p$ <0.05).





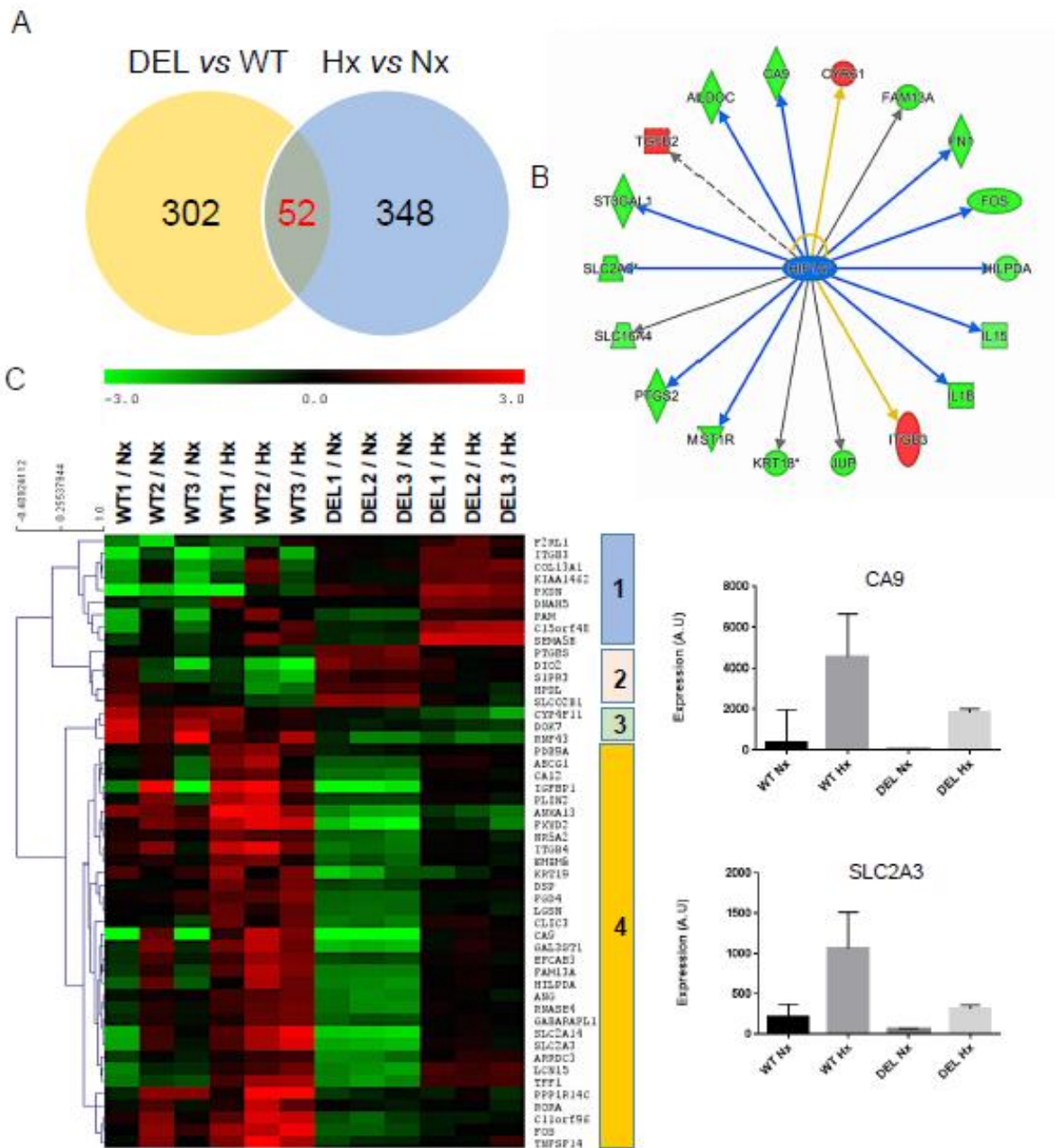
**Supplementary Figure S7: Effects of LUCAT1:4 short variant stable expression in LUCAT1/Lnc-ARRDC3-1 locus invalidated clones.**

LUCAT1/Lnc-ARRDC3-1 locus invalidated clones of A549 cells were reverted for short LUCAT1:4 variant expression by lentiviral vector infection. **A.** Single molecule RNA FISH of NLUCAT1, NLUCAT1/LUCAT1:4 and NEAT1 in WT, DEL and DEL-pLUCAT1:4 revertant clones of A549 cells. Scale bar corresponds to 5  $\mu$ m. **B.** Analysis of LUCAT1:4 and NLUCAT1 expression in 2 WT, 2 DEL and 2 DEL-pLUCAT1:4 revertant clones. RNA was extracted and assessed by RT-PCR for LUCAT1:4 and NLUCAT1 expression. Data are representative of 2 independent experiments. **C.** Cell proliferation: Cells were plated (250,000/well) in triplicates and counted 96 h post-plating. Data are means  $\pm$  SD from 2 DEL and 2 DEL-pLUCAT1:4 reverted A549 clones and are representative of 2 independent experiments. **D.** ROS production: 48 h post-plating, cells were treated or not with 40  $\mu$ M Cisplatin for 24 h, then labelled with fluorescent CM-H2DCF-DA probe and analyzed by cytometry for intracellular ROS production. Data are means  $\pm$  SD from 2 DEL and 2 DEL-pLUCAT1:4 reverted A549 clones and are representative of 2 independent experiments. **E.** Caspase activity: 48 h after plating, cells were treated or not with 40  $\mu$ M Cisplatin for 48 h, then lysed and analyzed for caspase 3/7 activity. Data are means  $\pm$  SD from 2 DEL and 2 DEL-pLUCAT1:4 reverted A549 clones and are representative of 2 independent experiments.



**Supplementary Figure S8: Validation of gene expression dysregulation in LUCAT1/Lnc-ARRDC3-1 locus-invalidated clones.**

**A.** RT-qPCR analysis of the expression of 28 genes, differentially expressed between WT and LUCAT1/Lnc-ARRDC3-1 locus-silenced cells (DEL). Data are means  $\pm$  SD from 2 independent experiments performed on 3 WT and 3 DEL A549 clones. Statistically significant differences are indicated (\* $p < 0.05$ , \*\* $p < 0.01$ ). **B-C.** RNA interference of LUCAT1:4 short variant (**B**) or reversion of LUCAT1:4 expression in LUCAT1/Lnc-ARRDC3-1 locus-invalidated clones (**C**) did not significantly alter the expression of the selected panel of genes. Data are means  $\pm$  SD from 2 independent experiments performed either on A549 cells with 2 distincts LUCAT1:4 siRNAs (**B**), or on 2 DEL and 2 DEL-pLUCAT1:4 reverted clones (**C**).



**Supplementary Figure S9: The hypoxic response is partially repressed in NLUCAT1-deleted cells.**

**A.** Venn diagrams showing the 52 genes that are commonly altered i) by NLUCAT1-deletion under normoxia or hypoxia ( $n=354$ ,  $\log_2(\text{intensity}) > 6$ , absolute  $\log_2(\text{fold change}) > 1$  and *adj. P* value  $< 0.01$ ), and ii) by hypoxia in WT and/or deleted-cells ( $n=400$ ,  $\log_2(\text{intensity}) > 6$ , absolute  $\log_2(\text{fold change}) > 1$  and *adj. P* value  $< 0.05$ ). **B.** Schematic of the most significant IPA HIF1A-regulated genes altered by NLUCAT1-invalidation. Genes repressed or upregulated in the signature are in green or red, respectively. **C.** Unsupervised hierarchical clustering of the 52 hypoxia-regulated genes that are differentially expressed between WT and NLUCAT1-DEL clones of A549 cells as shown in the Venn diagram in **A**. Genes were classified in 4 distinct clusters. Histograms on the right illustrate 2 important genes of cluster 4, namely CA9 and the glucose transporter GLUT3 (SLC2A3) that are induced by hypoxia in WT cells and repressed in DEL cells.



## ***Discussion et perspectives***

### **Identification de NLUCAT1, un hypoxa-lnc biomarqueur de l'agressivité des adénocarcinomes pulmonaires (ADC).**

De nombreux exemples démontrent d'ores et déjà le lien entre la dérégulation des lncARNs et différents processus de tumorigenèse. Les différentes analyses transcriptomiques de biopsies provenant de la cohorte locale du CHU de Nice et de cellules d'ADC cultivées au laboratoire ont mis en évidence une signature de lncARNs dérégulés dans les cancers. Parmi ces lncARNs, certains sont particulièrement corrélés à l'hypoxie, un marqueur de l'agressivité des CBNPC *in vivo* et *in vitro* et font de ces molécules des marqueurs pronostics ou de nouvelles cibles thérapeutiques d'intérêt.

Parmi ces lncARNs candidats, nous nous sommes focalisés sur les transcrits du locus *LUCAT1* qui est fortement corrélé à l'hypoxie et à un mauvais pronostic dans les adénocarcinomes suggérant son rôle de biomarqueur des ADC agressifs.

L'induction des lncARNs du gène *LUCAT1* par l'hypoxie est liée à différents facteurs de transcription activés par l'hypoxie, HIF2, NRF2 et NF-κB. Ces facteurs de transcription sont impliqués dans la progression tumorale et font l'objet de nombreuses recherches cliniques dans le traitement du cancer<sup>395-397</sup>. Les lncARNs *LUCAT1*, en aval de ces facteurs de transcription, semblent ainsi être des cibles thérapeutiques intéressantes et leurs caractérisations fonctionnelles et mécanistiques suscitent un grand intérêt.

Les progrès des outils de prédiction de séquences et les techniques de séquençage à haut débit a permis de déterminer l'existence de 25 isoformes du transcrit *LUCAT1* répertoriés dans la base de données LNCipédia (<https://lncipedia.org/>). La caractérisation des différentes isoformes du locus par séquençage d'ARNs a été confrontée à différentes limitations technologiques. Les lncARNs sont des transcrits riches en séquences répétées et les techniques de séquençage d'ARN de type « short read » séquentent uniquement des fragments de 75pb qui s'alignent difficilement sur le génome de référence dans le cas des lncARNs. Actuellement des techniques de « long reads » permettraient de lever une partie des biais technologiques en séquençant des fragments plus longs voir le transcrit en entier<sup>398</sup>. Nous avons toutefois réalisé la caractérisation structurale des isoformes du locus *LUCAT1* par des approches de RT-PCR et Northern Blot. Nous avons discriminé deux isoformes, NLUCAT1 « Nuclear *LUCAT1* » une isoforme majoritaire de 10kb et SCAL1/*LUCAT1* une isoforme cytosolique minoritaire de 0,9kb. SCAL1 (Smoke and Cancer Associated lncRNA 1) fait déjà l'objet de différentes publications et est induit par des extraits

de cigarettes dans différents modèles de lignées cellulaires de CBNPC <sup>398</sup>. Une étude récente publiée après nos travaux associe l'isoforme SCAL1 avec la réduction du stress lié à l'oxyde nitrique (NO) dans les cellules pulmonaires <sup>399</sup>. L'implication de la petite isoforme du gène *LUCAT1* et l'induction des transcrits par l'hypoxie laisse supposer un rôle fonctionnel dans la régulation du stress oxydatif.

### **Implication du transcrit NLUCAT1 dans la régulation du stress oxydatif des cancers broncho-pulmonaires**

La caractérisation de la localisation subcellulaire des isoformes nous a permis de différencier une isoforme majoritaire nucléaire, NLUCAT1 et une isoforme minoritaire cytosolique SCAL1. Afin d'appréhender la fonction de ce gène, nous avons invalidé son expression à l'aide du système de « genome editing » CRISPR/Cas9. Le système CRISPR (Clustered Regularly Interspaced Palindromic Repeats) /Cas9 (CRISPR Associated protein 9) est un moyen de défense bactérien contre l'attaque de phage et l'intégration d'ADN viral <sup>401</sup>. Ce système d'édition du génome utilisable dans les cellules eucaryotes est devenu en quelques années un outil fondamental de génie génétique et utilisé dans le traitement de nombreuses pathologies dont les cancers <sup>402</sup>. Dans le contexte de l'étude des lncARNs, l'utilisation de cet outil reste limité dû notamment à la place des gènes de lncARNs au sein des séquences de gènes codants (cf : III. 3. D). Toutefois, dans le cas de l'étude du locus LUCAT1, il s'agit d'un lncARN intergénique isolé des gènes à proximité. Afin d'invalider l'ensemble des isoformes du transcrit, nous avons excisé les séquences promotrices du transcrit ainsi que les 3 premiers exons correspondant à une région d'environ 2 kb et permettant une efficacité maximale de coupure par l'enzyme <sup>403</sup>.

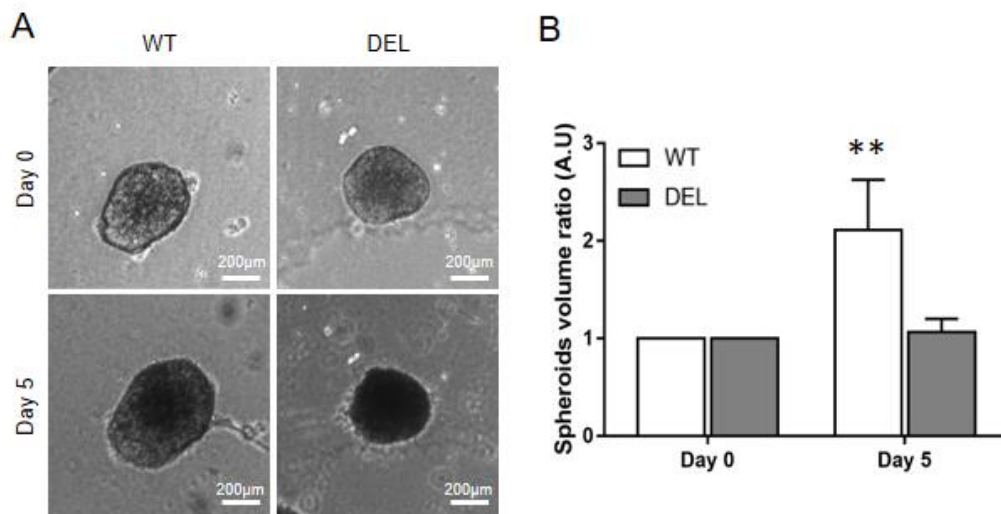
Les cellules d'ADC A549 invalidées pour le locus LUCAT1 ont un défaut de prolifération, de migration, d'invasion, une production accrue de radicaux libres oxygénés et une sensibilité plus importante au cisplatine utilisé comme traitements des ADC agressifs. Ces résultats confirment le rôle pro-tumoral des transcrits du locus LUCAT1 dans les ADC.

Afin de mieux comprendre la fonction des différentes isoformes, nous avons soit réexprimé l'isoforme SCAL1 par transduction virale dans les cellules invalidées par CRISPR/Cas9, soit réprimé uniquement la petite isoforme cytosolique par transfection de siRNA ne pouvant cibler que les transcrits cytosoliques. La réexpression de SCAL1 ou son invalidation par siARNs ne diminue pas la prolifération des ADC, ni n'augmente la sensibilité au cisplatine ou la production de radicaux libres oxygénés. Actuellement, deux études ont été publiées et font état du rôle de LUCAT1 dans la prolifération cellulaire des CBNPC après invalidation par siARNs <sup>403,404</sup>. Bien que ces résultats confirment les nôtres, l'implication de

l'isoforme responsable de ces phénotypes n'a pas été clairement identifiée dans ces modèles d'études. La littérature a démontré une efficacité moindre des siRNAs sur les lncARNs nucléaires. En effet, le variant RISC-AGO2 impliqué dans les mécanismes d'interférence des siRNAs sont majoritairement localisés et actifs au sein du cytoplasme <sup>406-409</sup>. Bien que la protéine AGO2 ait été détectée dans le noyau <sup>410-412</sup> et que la dégradation de l'ARN nucléaire 7SK par des approches de siRNAs ait été rapportée <sup>413</sup>, l'efficacité des siRNAs sur des régions introniques des pré-ARNs ou des ARNm dont la localisation est strictement nucléaire reste limitée <sup>408,409</sup>. Un consensus concernant la localisation subcellulaire du complexe RISC-AGO2 estime que la fraction nucléaire est variable en fonction des tissus de même que son activité et son accessibilité. Toutefois, de nombreuses études utilisent des siRNAs pour réprimer des lncARNs strictement nucléaires tel que les lncARNs NEAT1 et MALAT1 par des processus moléculaires inconnus <sup>413,414</sup>. Dans nos mains, l'utilisation de siRNAs réprime quasi-totalement l'expression de SCAL1 mais n'altère en aucune façon l'expression de l'isoforme majoritaire nucléaire NLUCAT1.

En revanche, compte tenu de la taille du variant NLUCAT1 (10kb) et de la présence de nombreuses séquences répétées nous n'avons pas réussi à générer un amplicon complet afin d'aborder sa caractérisation fonctionnelle par une approche de gain de fonction par expression ectopique. De nouvelles approches telles que le CRISPRa (activation) pourraient résoudre ces difficultés techniques en utilisant un ARN guide pour recruter des activateurs transcriptionnels au niveau des régions promotrices du gène et induire son expression <sup>416</sup>. Cette approche permettrait d'induire son expression dans un épithélium respiratoire sain, ou il n'est pas exprimé, et de déterminer son impact sur l'homéostasie tissulaire et éventuellement sur des mécanismes précoces de la tumorigenèse <sup>417</sup>. Malheureusement, étant donné la faible conservation inter-espèces des lncARNs, nous n'avons pas pu identifier d'orthologue murin de NLUCAT1 <sup>163</sup> qui nous aurait permis de mieux appréhender sa fonction *in vivo* dans des modèles murins d'ADC induits par l'expression de la forme mutée du gène *KRAS*.

Par ailleurs, nous avons développé une approche de culture tridimensionnelle de sphéroïdes afin de se rapprocher des conditions de croissance tumorale *in vivo* <sup>418</sup>. Nous avons confirmé nos résultats obtenus en culture 2D, puisque la croissance des sphéroïdes formés à partir de cellules invalidées pour NLUCAT1 est plus faible que celle des sphéroïdes issus de cellules non invalidées (Figure 37).



**Figure 37:** NLUCAT1 favorise la croissance de sphéroïdes tumoraux. L'invalidation du transcrit NLUCAT1 limite la croissance cellulaire dans un modèle de culture tridimensionnelle de sphéroïdes. A. prises de vues en contraste de phase au Cytation 5 (Bioteck). B. Mesure du volume des sphéroïdes à 0 et 5 jours pour 10 sphéroïdes par conditions.

### Identification de cibles moléculaires potentielles de NLUCAT1

Afin de mieux comprendre les mécanismes moléculaires mis en jeu lors de l'invalidation de NLUCAT1, nous avons analysé le transcriptome de cellules invalidées ou non pour le transcrit puis cultivées en normoxie ou en hypoxie. Des analyses d'expression différentielle ont mis en évidence 354 gènes impactés par l'invalidation du transcrit. Ces gènes participent à différents réseaux d'expression génique notamment les réseaux de la réponse au cisplatine et de la réponse anti-oxydante contrôlée par NRF2. Parmi ces gènes, nous nous sommes focalisés sur 4 gènes : GLRX, GPX2, ALDH3A1 et PDK4 dont l'invalidation concomitante par siARNs mime en partie les effets de l'invalidation du transcrit NLUCAT1 sur la régulation du stress oxydatif. Ces résultats confirment le rôle de ces gènes en tant qu'intermédiaires de la réponse médiée par NLUCAT1.

Cependant, l'étude du phénotype de cellules invalidées pour NLUCAT1 générées par délétion en utilisant le système CRISPR/Cas9 peut présenter des failles. En effet, cette approche nécessite de sélectionner des clones de cellules dans lesquelles les 3 allèles du gène ont été excisés (la lignée A549 est tri-allélique). Cette sélection prend du temps et peut introduire des biais comme la sélection de cellules s'adaptant à l'absence du transcrit NLUCAT1. Ces cellules pourraient être résistantes à l'altération de certaines voies de signalisation et ne permettraient pas d'appréhender les cibles directes du transcrit NLUCAT1. Une nouvelle approche appelée CRISPR interférence (CRISPRi) est développée

au laboratoire pour préciser les conséquences à court terme de l'inactivation de NLUCAT1. Ce système CRISPRi consiste à recruter, à l'aide de séquences d'ARN guides, un complexe répresseur de la transcription sur le promoteur du gène cible. Ce complexe est constitué de la fusion d'une forme inactive de l'enzyme Cas9 (dCas9) et des répresseurs transcriptionnels KRAB (Krüppel Associated Box) qui méthyle les histones du promoteur cible, et MeCP2 (Methyl-CpG-Binding Protein 2) qui se lie aux dinucléotides CpG méthylés et recrute des histones déacétylases. Après avoir établi des lignées exprimant de façon stable le complexe répresseur dCas9-KRAB-MeCP2, les cellules sont transduites pour exprimer des ARN guides et les conséquences de la répression du gène cible sont analysées à court terme par des analyses transcriptomiques <sup>419</sup>.

Par ailleurs, la détermination des cibles potentielles et intermédiaires moléculaires du transcrit NLUCAT1 ne renseigne pas sur le mécanisme d'action du transcrit dans le noyau. La caractérisation du mode d'action est donc liée à la recherche des partenaires moléculaires par des expériences de précipitations et de co-marquage des différents partenaires.

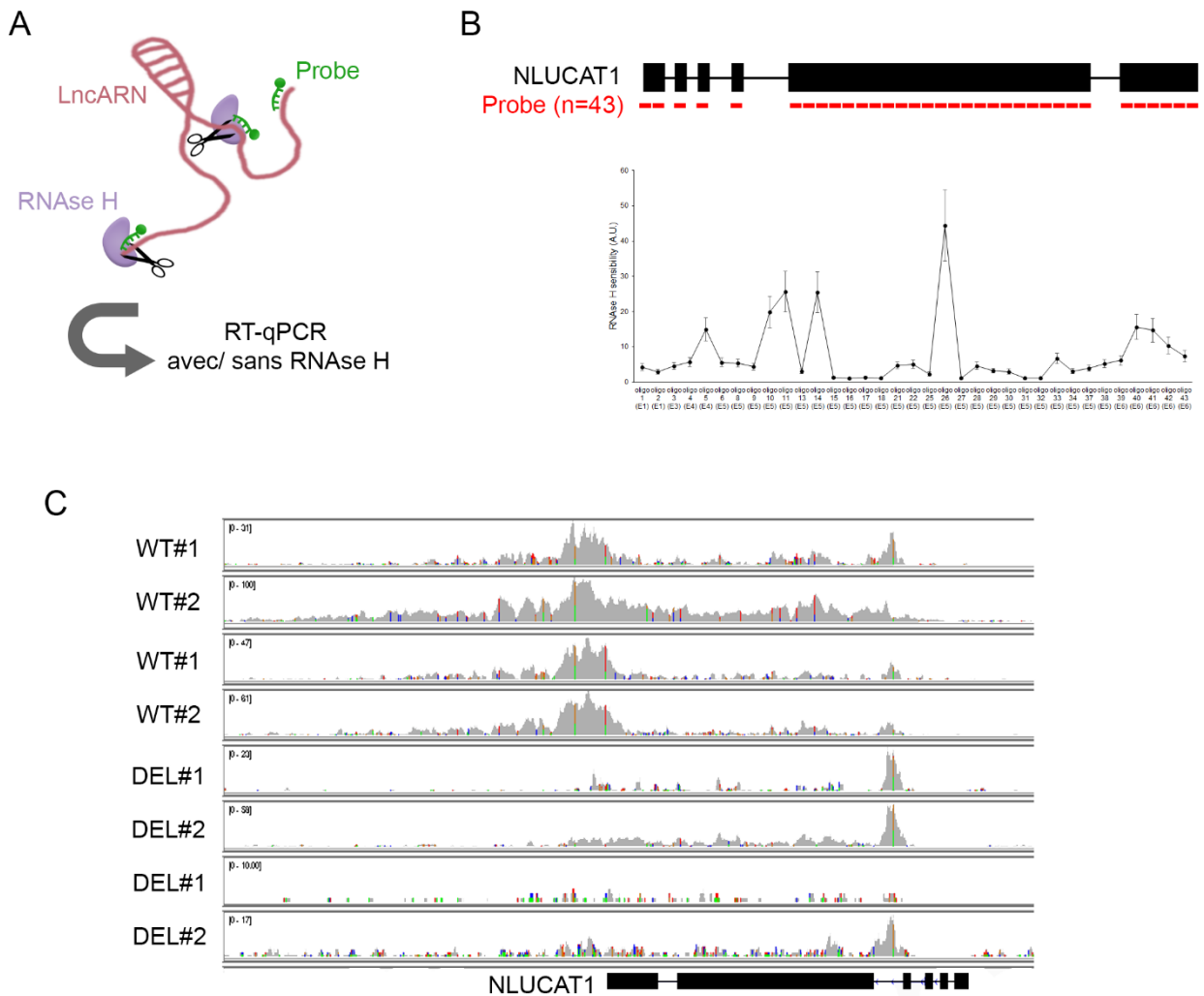
### **Recherche des partenaires moléculaires de NLUCAT1**

Considérant la localisation nucléaire de NLUCAT1, nous nous sommes focalisés sur la recherche de ses partenaires moléculaires dans le noyau. J'ai mis en place au laboratoire la technique de CHART-DNAseq afin d'identifier des sites d'interactions génomiques potentiels du transcrit NLUCAT1 à partir d'extraits nucléaires de cellules invalidées ou non pour le transcrit NLUCAT1. Nous avons utilisé 16 sondes antisens biotinylées complémentaires du transcrit NLUCAT1 et sélectionnées par cartographie à la RNase H (Figure 38AB). Les séquences génomiques co-précipitées avec le transcrit ont été identifiées par séquençage ADN puis alignement sur le génome de référence. Les résultats ont mis en évidence un enrichissement en séquences génomiques situées en aval du locus *LUCAT1* (Figure 38C). D'autres exemples de la littérature utilisant cette technique rapportent des enrichissements de séquences correspondant au gène du transcrit ciblé <sup>420</sup>. Cette interaction du transcrit avec son locus pourrait être liée au processus de maturation co-transcriptionnel moins efficace dans le cas des lncARNs, soit lié au mode d'action de NLUCAT1 qui reste fixé à cette séquence génomique et séquestre sur d'autres zones de sa séquence des protéines partenaires encore inconnues.

Afin de déterminer les protéines partenaires du transcrit NLUCAT1, nous avons réalisé des expériences de CHART-MS consistant à analyser les protéines co-précipitées par spectrométrie de masse. Cependant, le faible niveau d'expression du transcrit ne nous a

pas permis de précipiter suffisamment de complexes macromoléculaires pour identifier les protéines co-précipitées.

Nous avons alors réalisé des expériences de RNA pulldown en utilisant une séquence transcrite *in vitro* et biotinylée correspondant à un domaine de NLUCAT1 situé dans l'exon 5 et dont la structure secondaire semble conservée dans différentes analyses *in silico* de modélisation structurale <sup>421,422</sup>. De plus, ce domaine riche en structures tiges-boucles pourrait présenter des sites de liaison pour diverses ribonucléoprotéines. Des protéines co-précipitées avec ce domaine ont été identifiées par spectrométrie de masse puis validées par immunoprécipitation d'ARN (RIP). Des études fonctionnelles sur ces protéines partenaires restent à réaliser afin d'appréhender le mode d'action du transcrit NLUCAT1 ou le rôle de certaines de ces associations sur le métabolisme du transcrit.



**Figure 38:** Recherche des partenaires moléculaires du transcript NLUCAT1 par CHART-DNaseq.

**A.** Représentation schématique du protocole de cartographie à la RNase H de l'accessibilité de sondes antisens biotinylées sur un transcript. **B.** Graphe représentant la sensibilité à la RNase H des sites d'interactions des oligonucléotides. **C.** Séquençage des séquences génomiques co-précipitées avec le transcript NLUCAT1 par CHART-seq dans des cellules invalidées (DEL) ou non (WT) pour le transcript NLUCAT1.



**Partie 2 : LINC01116, un hypoxa-lnc cytosolique corrélé  
à l'agressivité des adénocarcinomes pulmonaires et  
impliqué dans la régulation du stroma tumoral**

## ***Introduction***

Précédemment, nous avons abordé l'importance des études portant sur les lncARNs dans les cancers et à leur importance en tant que marqueurs pronostics et/ou cibles thérapeutiques (cf. III. 5. B.). De même, l'étude que nous avons réalisée sur NLUCAT1 nous a permis de mettre en évidence l'implication d'un lncARN dans des processus pro-tumoraux et d'identifier une nouvelle molécule d'intérêt dans le traitement des CBNPC.

Au cours de ma thèse, j'ai également abordé l'étude d'un second lncARN candidat, le transcrit LINC01116. Lors de l'initiation de ce second projet, aucuns travaux n'étaient publiés sur ce transcrit. Seules quelques informations structurelles étaient répertoriées au sein de la base de données LNCipedia et rapportait l'existence d'une quinzaine d'isoformes potentielles allant de 300 nt à 13 000 nt.

Ainsi, afin de définir si LINC01116 est un lncARN candidat d'intérêt, nous avons recherché ses potentielles corrélations d'expression avec différents facteurs d'agressivité des CBNPC puis nous avons poursuivi notre étude par la caractérisation de sa structure, de sa fonction et de son mode d'action.

## **Matériels et Méthodes**

### **Culture cellulaire et transfection d'ARNs interférents et d'oligonucléotides antisens (ASOs)/Gampers**

Les lignées cellulaires d'adénocarcinomes pulmonaires A549, H1975, et de carcinome épidermoïde pulmonaire CALU-1, proviennent de l'ATCC (Rockville, MD) et sont cultivées dans du milieu DMEM contenant 10% de sérum de veau fœtal (Dominique Dutscher, Brumath, France). Les cellules endothéliales HUVEC et HMEC1 sont cultivées en milieu "Endothelial Serum-Free Medium" (Fisher) supplémenté avec 20% de sérum de veau fœtal, 20ng/mL de FGFb et 10ng/mL d'EGF. Pour l'hypoxie, les cellules sont cultivées pendant 24 h dans une station InvivoO<sub>2</sub> (Ruskin Technology Ltd, UK) réglée à 1% d'oxygène, 94% d'azote et 5% de dioxyde de carbone.

Les siARNs ON-Targetplus (pour LINC01116, HIF1A, HIF2A, ILF3, PABPC1) ont été achetés chez Dharmacon (Lafayette, CO). Deux siARNs distincts ont été utilisés par gène cible (Tableau S3). Les cellules sont ensemencées 24h avant la transfection à 50% de confluence puis transfectées avec 10nM de siARNs à l'aide de l'agent de transfection lipofectamine™ RNAiMax selon le protocole du fournisseur (Thermo Fisher Scientific, Courtaboeuf, France). Pour les expérimentations de migration transendothéliale des cellules cancéreuses, les cellules endothéliales ont été transfectés le jour de l'ensemencement. Les Gampers ont été achetés chez Exiqon et sont répertoriés dans le tableau S4.

### **Description de la cohorte locale utilisée pour l'analyse transcriptomique.**

Cinquante-sept tumeurs d'ADC de patients hospitalisés entre 2004 et 2010 à l'Hôpital Pasteur (Département de pneumologie et de chirurgie thoracique, CHU de Nice, France) ont été utilisées dans ce projet. Cette cohorte et les informations clinico-pathologiques correspondantes ont été détaillées dans une précédente étude <sup>423</sup>.

### **Hybridation de puces à ADN et analyses statistiques**

Les ARNs totaux des biopsies d'ADC, des biopsies péri-tumorales, des cellules A549 cultivées en normoxie ou en hypoxie ont été extrait, rétrotranscrit puis marqué en 3' à l'aide du marqueur fluorescent Cy3. Les ADC complémentaires marqués ont ensuite été hybridé sur puce à ADN « Agilent » puis la fluorescence a été quantifiée. L'analyse statistique des données est ensuite effectuée et permet d'obtenir une liste de gènes différentiellement exprimés à l'aide d'un score statistique dérivé d'un modèle linéaire utilisant une approche empirique bayésienne. Une P-Value ajustée (correction de Benjamini et Hochberg) < 0.05 a été utilisée comme seuil de significativité.

Dans un premier temps, nous avons isolé les lncARNs différentiellement exprimés entre les 57 biopsies d'ADC et les 11 biopsies péri-tumorales (Dataset 1). Le score hypoxique des tumeurs fut déterminé a posteriori à partir d'une sélection de 27 gènes faisant parti d'une signature métagène <sup>424</sup> correspondant au nombre de gènes dont le niveau d'induction par l'hypoxie est supérieur à 2. Les tumeurs et les tissus sains sont considérés comme non hypoxique si le score se situe entre 0-9 gènes, comme intermédiaire si le score se situe entre 10-18 gènes et comme fortement hypoxique si le score se situe entre 19-27 gènes. Cette étude a pris en compte uniquement les lncARNs dont l'expression est significative et correspond à une intensité supérieure à 6. Les lncARNs dont l'expression est corrélée à la signature métagène de l'hypoxie dans la cohorte locale ont un seuil de significativité <0,001. D'autre part, nous avons identifié les lncARNs dont l'expression est corrélée à la survie des patients à l'aide d'une régression de Cox à partir des données cliniques des patients. Le seuil de significativité a été fixée à <0,05. Enfin, la corrélation d'expression des lncARNs avec la signature métagène *in vitro* a été établie à l'aide de deux répliques biologiques de cellules A549 cultivés en normoxie ou en hypoxie (Dataset 2). Le seuil de significativité a été fixé à <0,001. Nous avons également réalisé l'analyse comparative du transcriptome de cellules HMEC1 traitées par des siARNs ciblant LINC01116 (siLINC01116#1 et siLINC01116#2) ou par un siRNA contrôle (siNC#1) (Dataset 3). La classification hiérarchique des gènes différenciellement exprimés (HeatMap) a été réalisée à l'aide du logiciel MultiExperiment Viewer (MeV) version 4.9 en utilisant la distance métrique de Manhattan et concerne les gènes dont le fold change est >0,5, la p-value est <0,05 et l'intensité est >6. L'annotation fonctionnelle des processus biologiques, des réseaux géniques et des voies de signalisation impactés par la diminution d'expression de LINC01116 a été réalisée avec le logiciel Ingenuity Pathway Analysis (<http://www.ingenuity.com/>) qui permet de croiser des données transcriptomiques et les informations bibliographiques.

### **Séquençage d'ARN**

Les banques ont été générées à partir de 500ng d'ARN totaux de cellules HUVEC en utilisant le kit TruSeq Stranded Total RNA Library (Illumina, San Diego, CA) et en suivant le protocole du fournisseur, puis quantifiées à l'aide du kit KAPA Library quantification kit (Roche, Basel, Switzerland) et mélangées de manière équimolaire (4 nmoles). Les librairies sont ensuite séquencées sur NextSeq500 (Illumina, San Diego, CA) afin de générer des reads pairés de 75pb. Les séquences obtenues ont ensuite été alignées sur le génome humain Hg19 avec le logiciel STAR v2.4.0.

### **Analyses de données publiques de la cohorte adénocarcinomes pulmonaires du TCGA**

Nous avons utilisé les données cliniques (tableau S2) et de séquençage ARN de biopsies d'ADC provenant de 522 patients. Les données de séquençage ARN ont été normalisées à l'aide du package « R » Bioconductor DESeq2. Le score hypoxique des tumeurs a été déterminé de la même façon que pour la cohorte locale. Les courbes de survie ont été générées en utilisant la méthode de Kaplan-Meier et la significativité déterminée à l'aide du test de log-rank.

### **Extraction d'ARN totaux et analyses d'expression des gènes par RT-qPCR et RT-PCR**

Les ARNs totaux ont été isolés avec du réactif TRIzol® (Thermo Fisher Scientific, Courtaboeuf, France) ou le kit RNeasy (Qiagen, Hilden, Germany) en suivant le protocole du fournisseur. L'analyse d'expression des gènes par qPCR a été réalisée comme précédemment décrit (ref Moreno et al.). L'analyse RT-PCR a été réalisée à l'aide de GoTaq DNA polymerase (Promega, Madison, WI, USA) à partir de 100ng d'ADN complémentaire puis analysée sur gel d'agarose. Les séquences des amorces utilisées sont présentées dans les tableaux S5 et S6.

### **Fractionnement subcellulaire**

Les fractions cytosoliques, chromatinienne et organites ont été préparé comme décrits précédemment<sup>205</sup> et les ARN ont été extraits par TRIzol® (Thermo Fisher Scientific, Courtaboeuf, France) en respectant les instructions du fournisseur.

### **Analyse par Hybridation *in situ* d'ARN en fluorescence à l'échelle de la molécule unique (smRNA FISH)**

Le marquage fluorescent des lncARNs à l'échelle de la molécule unique a été réalisé en utilisant la technologie RNAscope (Fluorescent Multiplex Assay, Advanced Cell Diagnostics, Hayward, CA) en suivant le protocole du fournisseur sur des cellules en culture (lignées, cellules endothéliales) et sur des coupes de tumeurs fixées et cryogénisées. Un lot de sondes oligonucléotidiques antisens fluorescentes spécifiques du transcrit LINC01116 a été conçu afin de recouvrir l'ensemble de sa séquence (1-1121pb). Des sondes commerciales ciblant le gène bactérien *dapB* ont servi de contrôle négatif.

### **Clonage du transcrit LINC01116**

Le transcrit LINC01116 a été amplifié entièrement par RT-qPCR à partir d'ARN totaux provenant de cellules endothéliales. L'amplicon de 1121pb a été inséré dans le vecteur d'expression lentiviral pLJM1 (Addgene, Cambridge, MA) entre les sites de digestion Age1 et EcoR1, mais également dans le vecteur pGemt Easy (Promega, Charbonnières les bains, France) afin de réaliser des expériences de transcription *in vitro*. Les particules lentivirales ont été produites par la Plateforme de Virologie de Montpellier (PVM).

### **Analyse de l'adhésion cellulaire**

Des puits de plaque 96 puits ont été coatés pendant 2h à 37°C avec différentes protéines matricielles: fibronectine (20µg/mL), collagène I (20µg/mL), collagène IV (20µg/mL), gélatine 0,1% ou sérum de veau foetal 10%. Les cellules ont été trypsinées, lavées, comptées puisensemencées à 20 000 cellules/puits. Après un temps d'adhésion allant de 30min à 2h, les puits sont lavés avec du PBS puis incubés 15 min dans une solution de crystal violet (0,5% crystal violet, 20% Methanol). Après lavage avec de l'eau, les puits sont photographiés au Cytation 5 (BioTek, Colmar, France), puis incubés durant 10min dans du Triton X-100 (Sigma, Saint Quentin, France) 0,1% avant quantification par lecture de l'absorbance à 650nm.

### **Anticorps**

Les anticorps primaires utilisés pour les expériences de Western Blot sont anti-pFAK-Tyr397 (Cell Signaling, #3283S), anti-FAK (Cell Signaling, D1A8), anti-HEY2 (Abcam, ab32231), anti-HSP60

(Santa Cruz, sc-1052), anti-PABPC1 (ProteinTech, 66809-1-Ig) et anti-ILF3 (Sigma, HPA001897). Les anticorps secondaires conjugués à la peroxidase proviennent de Dako (Copenhagen, Denmark). Les anticorps précipitants utilisés sont anti-IgG rabbit (Cell Signaling, #2729S), anti-IgG1 mouse (Cell signaling, #5415), anti-ILF3 (Abcam, ab225626), anti-PABPC1 (ProteinTech, 10970-1-AP), anti-PKR (Santa Cruz, sc-377103) et anti-Tia1 (Santa Cruz, sc-166247). Pour les expériences d'immunofluorescences, les anticorps utilisés sont anti-paxilin (BD Bioscience, #610052), anti-VWF (Abcam, ab6994) et anti-cadherin E (Cell Signaling, #24E10).

### **Marquages en Immunofluorescence**

Les cellules endothéliales sont trypsinées puisensemencées sur lamelles de verre préalablement coatées avec 20µg/mL de fibronectine. Après différents temps d'adhésion, les cellules sont lavées en PBS puis fixées dans une solution de paraformaldéhyde 4% durant 15min à température ambiante. Les lamelles sont ensuite incubées 1h à température ambiante dans du PBS contenant 1% de sérum de chèvre et 0,1% de Triton X-100. Les cellules sont ensuite incubées avec un anticorps primaire sur la nuit à 4°C dilué dans du PBS contenant 1% d'albumine sérique bovine, et 0,1% de Triton X-100 (anti-paxiline 1/200, anti-VWF 1/400). Les lamelles sont ensuite lavées puis incubées 1h à avec un anticorps secondaire anti-souris (1/500) couplé à un fluorochrome-647nm et de la phalloïdine-488nm (1/300). Les lamelles sont ensuite lavées puis montées sur lame à l'aide de Prolong Gold (Thermo Fisher, Courtaboeuf, France).

### **Analyse de la perméabilité de l'endothélium**

Les cellules endothéliales sontensemencées à confluence (10<sup>5</sup> cellules/cm<sup>2</sup>) dans des inserts de polycarbonate contenant des pores de 8µm. Un jour plus tard, la chambre inférieure est remplie avec 1mL de milieu de culture et la chambre supérieure avec 300µl de milieu contenant 12,5µg/mL de FITC-dextran. La diffusion du dextran est analysée toutes les 5min pendant 30min, en prélevant 50µl de milieu dans la chambre inférieure et en les disposant dans une plaque 96 puits à fond noir pour mesure de la fluorescence à 488nm.

### **Analyse de la migration transendothéliale de cellules cancéreuses**

Les cellules endothéliales sont transfectées avec des siRNAs. 48h après transfection, elles sontensemencées à 40 000 cellules par insert perméable (8µm de diamètre) pendant 24h. Les cellules cancéreuses A549 sont marquées pendant 1h à 37°C à l'aide d'un marqueur vital Bodipy® émettant à 505nm (Thermo Fisher, Courtaboeuf, France) puisensemencées à 50 000 cellules par insert. Après 24h, les inserts sont lavés, les cellules fixées au paraformaldéhyde, perméabilisées au triton X-100 pour marquer les noyaux au Dapi. Les cellules présentes dans la chambre supérieure sont éliminées par lavage à l'aide d'un coton tige, et les inserts sont découpés et montés entre lame et lamelle. Les cellules ayant migré au travers de l'endothélium sont comptabilisées sur champs capturés au microscope à épifluorescence.

### **Analyse de la migration et de l'invasion cellulaires**

Les expériences de migration ont été réalisées dans des chambres de Boyden comme précédemment décrit <sup>425</sup>. Les expériences d'invasion ont été effectuées à l'aide du kit commercial 24-well BioCoat Matrigel Invasion Chamber (BD Biosciences, San Jose, CA).

La migration a également été analysée en "scratch-assay". Les cellules sontensemencées à confluence dans une plaque 6 puits. Une blessure est réalisée 24 h plus tard au centre du puits à l'aide d'une pointe de pipette. Des prises de vues en contraste de phase sont effectuées toutes les 30 min sur différents sites le long de la blessure durant 24h à l'aide du Cytation 5 (BioTek, Colmar, France). Les images sont ensuite analysées avec le logiciel ImageJ afin de quantifier la colonisation de la surface blessée au cours du temps.

### **Analyse de la tubulogénèse in vitro**

Les puits d'une plaque 96 puits sont coatés avec 50 µl de matrigel (Corning, Amsterdam, Pays Bas) et incubé 1h à 37°C pour solidification. Les cellules endothéliales HUVEC ou HMEC1 (4000/puit) sontensemencées et des prises de vues sont réalisées en contraste de phase à 24h à l'aide du Cytation 5 (BioTek, Colmar, France). Les images sont analysées à l'aide du plugin "Angiogenesis Analyzer" pour le logiciel ImageJ afin de quantifier le nombre de noeuds, de mailles et de segments du réseau tubulaire.

### **Analyses de la prolifération cellulaire**

Les cellules sontensemencées en plaques de 96 puits (5000 cellules/puit) et le nombre de cellules déterminé au cours du temps par ajout de réactif MTT (0,5 mg/ml, Sigma) incubé durant 1h à 37°C. Les puits sont ensuite lavés puis incubés avec 100 µl de DMSO, et la DO mesurée à 550 nm. La capacité proliférative des cellules est également évaluée en déterminant le pourcentage de cellules en phase S par incorporation de BrdU en suivant les recommandations du kit commercial 5-Bromo-2'-deoxy-uridine Labeling and Detection Kit (Roche, Meylan, France).

### **Culture tridimensionnelle de sphéroïdes**

Les cellules d'ADC sontensemencées dans des plaques de culture 96 puits à très faible adhérence et à fond rond (10 000 cellules) (Fisher) pendant 24h à 37°C. Les sphéroïdes formés sont photographiés puis récupérés pour inclusion dans un gel composé de matrigel et de collagène I. Un gel acellulaire constitué de 10% FBS, 10% DMEM 5X, 20% de Matrigel, 2 mg/ml de collagène I et 20% de milieu de culture est déposé au fond de plaque MatTek (6 puits). Après solidification du gel acellulaire durant 10 min à 37°C, les sphéroïdes sont repris dans un mix de gel similaire et déposés sur le gel acellulaire. Les puits sont ensuite remplis de milieu de culture.

### **Immunoprécipitation d'ARN (RIP)**

Des extraits cellulaires ont été préparés en tampon RIPA (Fisher) puis dilués dans une solution RIP contenant du Tris-HCl 25mM pH 7.0, du NaCl 150mM, du NP40 0,5%, du DTT 0,5mM, de la RNAsin 100U/mL (Promega) et des inhibiteurs de protéases (Thermo Fisher, Courtaboeuf, France) afin d'obtenir une concentration protéique à 500 µg/mL. Un aliquot de lysat cellulaire (250 µg de protéine) est incubé toute la nuit à 4°C avec 1 µg d'anticorps primaire précipitant.

Le lendemain, les complexes immuns sont précipités durant 1h à 4°C par ajout de 30 µl de billes magnétiques, couplées avec de la protéine A pour les anticorps primaires de lapins ou couplées avec de la protéine G pour les anticorps primaires murins, préalablement lavées avec la solution RIP. Les billes sont ensuite lavées 5 fois avec de la solution RIP supplémentée avec 1% de Tween-20 sur un portoir magnétique. Une fraction des billes (10%) est récupérée puis reprise en tampon Laemmli 3% SDS + β mercapto-éthanol afin d'analyser l'efficacité de précipitation des anticorps primaires. Les ARNs fixés sur les billes restantes sont extraits par TRIzol® et précipités à l'aide d'un carrier (GlycoBlue, Thermo Fisher). La révélation des ARNs co-précipités est réalisée par qPCR.

### **Identification de protéines partenaires par "pulldown" d'ARN**

#### Transcription in vitro et biotinylation des ARNs néosynthétisés

LINC01116 est transcrit in vitro à partir du vecteur pGemt Easy à l'aide du kit T7 RiboMax™ Express Large Scale RNA (Promega, Charbonnières les bains, France). Les ARNs sont extraits à l'aide du réactif TRIzol® (Thermo Fisher, Courtaboeuf, France) en suivant le protocole du fournisseur puis ils sont biotinylés en 3' en utilisant le kit commercial RNA 3' End Biotinylation Kit (Pierce™, Thermo Fisher, Courtaboeuf, France) et en respectant les conditions du fournisseur pour la ligation et la purification des ARNs. Les ARNs sont ensuite dénaturés et quantifiés.

#### Précipitation des complexes macromoléculaires

Les ARNs biotinylés (10 µg) sont incubés 30 min à température ambiante avec 20µl de billes magnétiques Dynabeads™ MyOne™ Streptavidin C1 (Thermo Fisher, Courtaboeuf, France) dans une solution de capture des ARNs contenant 20mM Tris-HCl pH 7.0, 1M NaCl, 1mM EDTA et 100U/mL de RNAsin (Promega). Les complexes billes-ARNs sont ensuite lavés plusieurs fois dans une solution de Tris-HCl 20mM pH 7.0, puis les billes sont resuspendues dans 300 µl de lysats cellulaires contenant 1mg de protéines et incubées 4h à 4°C sous agitation. Les billes magnétiques ayant fixé les complexes macromoléculaires sont ensuite capturées à l'aide d'un portoir magnétique, lavées plusieurs fois dans du tampon 20mM Tris-HCl pH 7.0, 0,5M NaCl, 0,2% Tween-20 puis resuspendues dans du tampon Laemmli 3% SDS contenant du β mercapto-éthanol et chauffées 10 min à 95°C.

#### Analyse des échantillons par spectrométrie de masse

Les protéines précipitées sont séparées par électrophorèse sur un gel d'acrylamide en gradient 4-12% puis révélées par coloration à l'argent à l'aide du kit Silver Stain Coloration (Pierce™, Thermo Fisher, Courtaboeuf, France) en suivant les recommandations du fournisseur.

Les bandes différenciellement exprimées entre les conditions de précipitation en présence ou en absence du transcrit LINC01116 sont découpées puis digérées à la trypsine. Brièvement, les bandes

de gel sont déshydratées à l'acétonitrile puis incubées 1h à 56°C dans une solution de DTT 10mM. Les échantillons sont ensuite réduits dans une solution de iodoacétamide 100mM et de bicarbonate d'ammonium 25mM durant 30min à l'obscurité. Les morceaux de gel sont lavés à l'eau puis déshydratés à l'acétonitrile durant 10min. Les bandes de gels sont ensuite réhydratées dans du bicarbonate d'ammonium 50mM durant 45min à 4°C et les protéines digérées à 37°C sur la nuit avec 0,8µg de trypsine. Les peptides sont ensuite extraits à l'aide de différents lavages contenant 5% d'acide formique et 50% d'acétonitrile puis lyophilisés au SpeedVac avant d'être transmis pour analyse à la plateforme de spectrométrie de masse de l'IRCAN à Nice.

### **Expériences de purification d'ARN par antisens couplées à de la qPCR (RAP-qPCR)**

#### Fixation et lyse des cellules

Les cellules HUVEC sont cultivées dans le milieu précédemment défini et trypsinées à confluence. 10 millions de cellules par condition sont utilisées et après lavages, sont fixées dans une solution 3% Formaldéhyde. La fixation est arrêtée par ajout de 0,125M Glycine et les cellules sont lysées à l'aide d'un tampon 1% NP-40 (150mM NaCl, 0,8mM MgCl<sub>2</sub>, 5mM EDTA, 50mM Hepes, protease inhibitor, 1% NP-40, 100U/mL RNasin), soniquées et centrifugées.

#### Hybridation des sondes biotinylées et précipitations des complexes macromoléculaires

Les sondes biotinylées de 20pb sont désignées le long de la séquence de LINC01116 (n=8) et biotinylées en 3'. 50nm de sondes biotinylées sont hybridées avec 1mL de lysat cellulaire et 2mL de tampon d'hybridation (500mM NaCl, 1% Tween-20, 50mM Tris-HCl, 1mM EDTA, 15% Formamide, protéase inhibitor, 100U/mL RNasin) durant 4h à 37°C. Les complexes macromoléculaires sont précipités à l'aide de 50ul de billes magnétiques Dynabeads C1 (Thermo Fisher, Courtaboeuf, France) prélavées 3x dans du Tris-HCl 20mM et incubées 30min à 37°C. Les complexes macromoléculaires fixés au billes magnétiques sont lavés 5X à l'aide de Wash Buffer (2XSSC, 0,5% SDS).

#### Traitements et extractions des ARNs

Les billes magnétiques lavées sont ensuite incubées 1h à 55°C, 30min à 65°C puis 10min à 90°C dans du tampon Protéinase K (100mM NaCl, 10mM Tris-HCl pH :7, 1mM EDTA, 0,5% SDS, 1mg/mL Protéinase K) puis les ARNs sont extraits à l'aide de Trizol© (Thermo Fisher, Courtaboeuf, France) à l'aide d'un carrier (Glycoblue, ThermoFisher, Courtaboeuf, France) et selon la procédure du fournisseur. Les ARNs extraits et purifiés sont ensuite rétrotranscrit à l'aide du kit High Capacity (Thermo Fisher, Courtaboeuf, France) et l'expression des transcrits est analysés par qPCR.



## Résultats

### Identification des lncARNs régulés par l'hypoxie dans les adénocarcinomes pulmonaires (ADC).

Afin d'identifier une signature génique de lncARNs associés à l'agressivité des adénocarcinomes pulmonaires nous avons réalisé des analyses bioinformatiques de données de transcriptomiques et de données cliniques de patients en considérant 4 paramètres: i) la dérégulation d'expression entre tissu sain et tissu tumoral, ii) la corrélation au statut hypoxique des tumeurs, iii) la régulation par l'hypoxie dans nos modèles d'étude *in vitro* et, iv) la corrélation avec la survie globale des patients. Dans ce but, les données de transcriptomique précédemment générées, par hybridation de puces à ADN, à partir, i) de tumeurs d'adénocarcinomes pulmonaires issues de la biobanque du CHU de Nice, et ii) de la lignée d'adénocarcinome A549 cultivée en hypoxie, ont été ré-analysées en prenant en considération les données cliniques sur la survie des patients.

Nous avons déterminé une signature de 716 lncARNs dérégulées entre les tumeurs et les tissus sains. Les tumeurs ont été classées à posteriori en fonction d'un score hypoxique basé sur une signature métagénique<sup>423</sup>. Nous avons observé une corrélation entre l'expression de 321 lncARNs et le statut hypoxique des tumeurs. Par ailleurs, les données cliniques sur le suivi des patients nous ont permis de mettre en évidence une corrélation entre l'expression de 314 lncARNs et une altération de la survie globale de ces patients. L'analyse croisée de ces données a révélé une signature restreinte de 14 lncARNs corrélés au statut hypoxique *in vivo* et à une altération de la survie des patients. Parmi ces candidats, seuls 3 lncARNs sont régulés par l'hypoxie dans la lignée A549 : les lncARNs intergéniques LINC01116, LINC00963 et le lncARN ZRANB1-2 (Figure 1A).

#### LINC01116, un nouvel « Hypoxa-lnc »

Nous nous sommes alors focalisés sur LINC01116 dont l'expression est augmentée dans les tumeurs de la cohorte de patients du CHU de Nice ayant un statut hypoxique élevé (*log rank test pvalue* : 0,015) (Figure 1B). Comme précédemment mentionné, la forte expression de ce transcrit est associée à une diminution très significative de la survie de ces patients (Figure 1C). Ces observations ont été validées, en analysant des données transcriptomiques publiques générées par séquençage ARN et provenant d'une cohorte de patients américains de plus grande ampleur (TCGA, The Cancer Genome Atlas) comportant 522 tumeurs ADC (*log rank test pvalue* :  $8,9 \times 10^{-9}$ ) (Figure 1D-E).

L'analyse de l'expression de LINC01116 par qPCR montre une induction d'un facteur 2 en réponse à l'hypoxie dans les lignées d'adénocarcinomes pulmonaires A549 et H1975 présentant des statuts mutationnels distincts (Figure S1A). De plus, la répression des facteurs de transcription HIF1A et HIF2A par interférence ARN montre que cette régulation par l'hypoxie implique exclusivement le facteur HIF2A/EPAS (Figure S1C). L'analyse de l'expression du transcrite de l'anhydrase carbonique CA9 sert de contrôle positif à la réponse hypoxique régulée par le facteur HIF1A (Figure S1B-D).

Ainsi, le transcrite LINC01116 semble être un intéressant marqueur de mauvais pronostic pour les patients atteints par un adénocarcinome broncho-pulmonaire.

### **Caractérisation structurale du transcrite LINC01116**

Afin d'aborder l'étude fonctionnelle ultérieure de ce lncARN, nous avons préalablement entrepris sa caractérisation structurale. Les informations disponibles dans la base de données LNCipédia suggèrent l'existence de 19 isoformes potentielles du transcrite LINC01116 allant de 383 à 13732 nt.

Nous avons mis en évidence, par des approches de séquençage ARN et de RT-PCR classique couplée à du séquençage Sanger, l'existence dans nos lignées d'ADC d'un transcrite de 1121 nt constitué de 3 exons et se rapprochant de la composition exon-intron de l'isoforme LINC01116:17 (Figure 2).

Afin de préciser la localisation subcellulaire de LINC01116 nous avons réalisé des expériences d'hybridations in situ d'ARN en fluorescence (single-molecule RNA Fluorescent In Situ Hybridization, smRNA-FISH) et de fractionnements subcellulaires couplées à de la PCR quantitative. Les résultats montrent que LINC01116 est majoritairement cytosolique et qu'une partie est également associée à une fraction membranaire (Figure S2).

### **Etude de l'origine cellulaire de LINC01116 dans les tumeurs pulmonaires**

Afin d'analyser plus précisément l'expression de LINC01116 dans de l'épithélium respiratoire sain, nous avons réalisé une expérience de séquençage d'ARN sur cellules uniques à partir de pièces opératoires de turbinectomies (Figure 3AB). Les résultats montrent une expression non détectable de LINC01116 dans les cellules épithéliales nasales mais une forte expression dans un groupe de cellules exprimant également les transcrits du facteur de Von Willebrand (VWF) et de PECAM-1 (CD31, Platelet endothelial cell adhesion molecule), marqueurs spécifiques des cellules endothéliales. Ces résultats sont corroborés par des analyses d'expression par qPCR sur des cultures primaires (HbMEC, HUVEC) ou sur une lignée immortalisée de cellules endothéliales (HdMEC1), qui révèlent des niveaux

d'expression bien plus conséquents que ceux mesurés dans nos lignées d'ADC A549 et H1975 (Figure 3C). De façon intéressante, l'expression de LINC01116 est très faible dans la lignée non tumorale de cellules épithéliales bronchiques BEAS-2B, alors qu'elle est très importante dans la lignée de cancer épidermoïde pulmonaire CALU-1. LINC01116 est également induit par l'hypoxie dans les cellules HUVEC (Figure 3D). De plus, nous avons confirmé que le transcrite exprimé dans les cellules endothéliales est cytosolique et identique à celui exprimé dans les cellules tumorales (Fig S3).

En se basant sur ces résultats, nous avons visualisé par smRNA-FISH l'expression de LINC01116 dans des coupes d'adénocarcinomes. Nous avons observé l'expression du transcrite dans les cellules tumorales exprimant la cadhérine E mais également dans les cellules endothéliales, positives pour VWF, des vaisseaux drainant la tumeur (Fig 3EF). Il est important de noter que le co-marquage de protéines par immunofluorescence et des ARN par FISH nécessite des traitements de déprotéinisation partielle des tissus qui altèrent la qualité des marquages protéiques.

Par conséquent, dans les tumeurs pulmonaires, l'hypoxie augmente l'expression de LINC01116 dans les cellules cancéreuses mais également dans la composante vasculaire du stroma tumoral.

### **Caractérisation fonctionnelle de LINC01116 dans les cellules tumorales et dans les cellules endothéliales**

Nous avons abordé l'étude fonctionnelle de LINC01116 dans les ADC pulmonaires par des approches de perte et de gain de fonction dans la lignée d'adénocarcinome A549.

La transfection d'ARN interférents (siARNs) ciblant LINC01116 diminue d'environ 75% l'expression de ce transcrite dans les cellules A549 (Figure S4AD). Cette inhibition est quasi-totale si l'on analyse plus spécifiquement l'expression cytosolique de LINC01116 (Figure S4BC). Dans ces conditions, l'analyse des conséquences de l'inhibition d'expression de LINC01116 sur la prolifération et la migration cellulaires n'a pas permis de mettre en évidence d'effets notoires sur le contrôle de ces fonctions (Figure S4E-H).

Inversement, nous avons surexprimé LINC01116 par transduction lentivirale dans la lignée A549 (Figure S5AB). L'augmentation d'un facteur 5 du niveau de ce transcrite n'a aucune conséquence sur la prolifération ou sur la migration des cellules analysée dans un système de chambres de Boyden ou par un essai de refermeture de blessure sur un tapis cellulaire (Figure S5C-E). Ces résultats montrent que l'altération de l'expression de ce lncARN n'a aucun impact sur la croissance ou la mobilité des cellules tumorales.

En revanche, la répression transitoire de l'expression de LINC01116 à l'aide de deux siARNs distincts affecte significativement l'adhésion des cellules A549 sur des puits de culture coatés avec du collagène de type I ou de la fibronectine (Figure 4 AB).

Nous avons souhaité étudier les conséquences de la répression d'expression de LINC01116 dans un modèle de culture en 3D *in vitro* plus proche de la croissance d'une tumeur *in vivo*. Les sphéroïdes sont des modèles de culture 3D multicellulaires caractérisés par la mise en place au cours de leur croissance de gradients de nutriments, d'hypoxie et de prolifération mimant l'organisation cellulaire de micro-régions tumorales. Ainsi, le modèle sphéroïde permet de considérer les interactions et l'hétérogénéité cellulaire ainsi que le microenvironnement. Des cellules d'ADC A549 transfectées avec un siARN contrôle ou deux siARNs ciblant LINC01116 ont étéensemencées dans des puits à fond rond à faible adhérence afin d'induire l'aggrégation des cellules. De façon intéressante, nous avons constaté que les sphéroïdes formés à partir des cellules dans lesquelles LINC01116 est réprimé présentent un défaut de compaction se traduisant par un plus grand volume apparent (Figure S6AB). Ces observations suggèrent qu'en plus de présenter une diminution d'adhésion sur des protéines matricielles, les cellules ayant un niveau réduit de LINC01116 se caractérisent également par un défaut de connections intercellulaires.

Afin de déterminer les effets de la répression durable de l'expression de lncARNs candidats, nous avons développé au laboratoire une approche de CRISPR interférence (CRISPRi) (cf : Annexe). Ce système consiste à recruter, à l'aide de séquences d'ARN guides, un complexe répresseur de la transcription sur le promoteur du gène cible. Ce complexe est constitué de la fusion d'une forme inactive de l'enzyme Cas9 (dCas9) au domaine répresseur de la transcription de KRAB (Krüppel Associated Box) (Figure 4C). Après avoir établi une lignée A549 exprimant de façon stable le complexe répresseur dCas9-KRAB, les cellules ont été transduites pour exprimer des ARN guides contrôle (NEG#1) ou ciblant la région proche du site d'initiation de la transcription de LINC01116 (sg-1 et sg-2). Après sélection, les cellules présentent une réduction de 60 à 70% de l'expression du transcrit (Figure 4D). Les conséquences de cette répression ont été analysées et confirment une nette diminution de l'adhésion des cellules tumorales sur du collagène de type I ou de la fibronectine (Figure 4EF).

De façon similaire à ce que nous avons observé dans les cellules d'ADC, l'interférence ARN ciblant LINC01116 n'impacte pas la prolifération et la migration des cellules endothéliales HUVEC (Figure S7), mais elle a un effet drastique sur leur adhésion sur différentes protéines matricielles incluant le collagène I, la fibronectine ou encore le collagène IV constituant majeur de la lame basale des vaisseaux (Figure 5AB). De plus, la

diminution du niveau d'expression de LINC01116 induite par des oligonucléotides antisens de type Gapmers, via un mécanisme impliquant la RNase H, produit les mêmes effets (Figure S8). Cette altération des capacités d'adhésion des cellules endothéliales sur différents supports a été confirmée dans la lignée immortalisée de cellules endothéliales microvasculaires dermiques HMEC1<sup>426(p1)</sup> (Figure S9A-C). Afin de préciser les conséquences de l'inhibition de LINC01116 nous avons réalisé des essais de tubulogenèse *in vitro* qui consistent à analyser l'organisation de cellules endothéliales sous forme de tubes sur du matrigel, une matrice de membrane basale extraite d'un sarcome murin et riche en laminine et collagène IV. Les cellules traitées avec les siARNs contre LINC01116 produisent des réseaux de tubes moins abondants et moins réticulés se traduisant par un nombre restreint de noeuds, de mailles et de segments de ces réseaux (Figure S9DE).

Afin d'évaluer l'impact de LINC01116 sur les interactions intercellulaires de cellules HUVEC nous avons analysé la fonctionnalité d'un endothélium *in vitro*. L'inhibition du transcrit LINC01116 augmente la perméabilité de l'endothélium quantifiée par la diffusion accrue de dextran-FITC (Figure 6A) traduisant un défaut dans la fonction de barrière de l'endothélium.

Nous avons également observé que la migration transendothéliale de cellules tumorales A549 est réduite à travers un endothélium dans lequel l'expression du transcrit LINC01116 a été inhibée (Figure 6BC). Ce défaut de migration pourrait être lié à un défaut de formation du gradient de facteurs de croissance (SVF+EGF) diffusant de la chambre inférieure vers la chambre supérieure afin d'attirer les cellules tumorales, qui serait conséquent à une perméabilité intercellulaire accrue de l'endothélium. En effet, lorsqu'on casse ce gradient par ajout de facteurs de croissance dans la chambre supérieure, les cellules tumorales ne migrent plus vers la chambre inférieure (Figure 6DE).

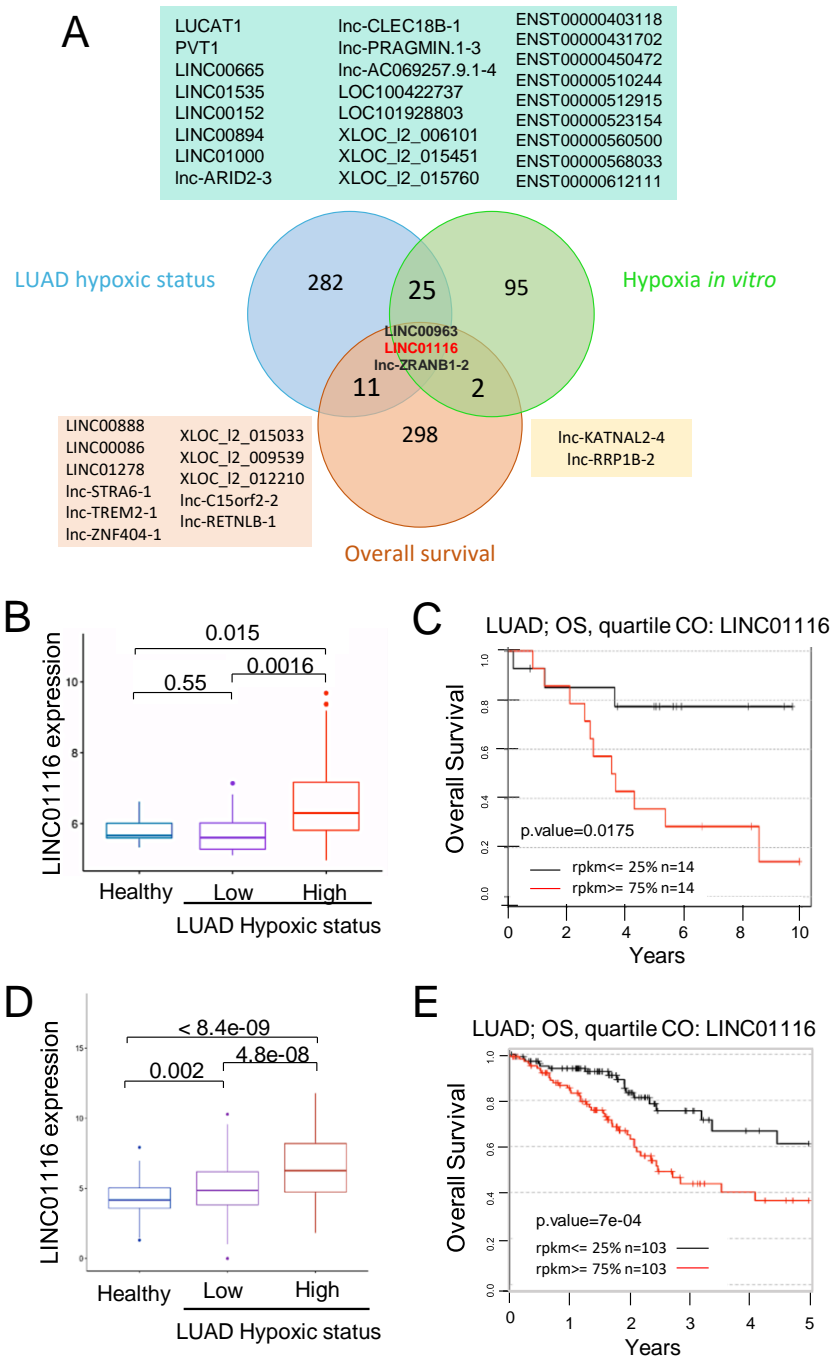
### **Mécanismes moléculaires affectés par l'inhibition du transcrit LINC01116 dans les cellules endothéliales et les cellules tumorales**

Afin de préciser les mécanismes moléculaires altérés par la répression de l'expression de LINC01116 et impliqués dans l'adhésion cellulaire nous avons dans un premier temps analysé des événements moléculaires induits plus ou moins précocément suite à l'engagement des intégrines lors de l'adhésion de cellules HUVEC sur une protéine matricielle. L'activation de l'intégrine  $\alpha 5\beta 1$  par son ligand, la fibronectine, entraîne l'activation de la tyrosine kinase p125FAK par phosphorylation de son résidu tyrosine 397 puis la formation de structure d'adhérences faisant le lien avec le cytosquelette appelées adhésions focales, riches en paxilline, un substrat de FAK<sup>427</sup>. Dans les cellules traitées avec un siARN contrôle, l'analyse de l'état de phosphorylation de la tyrosine 397 de FAK par Western blot

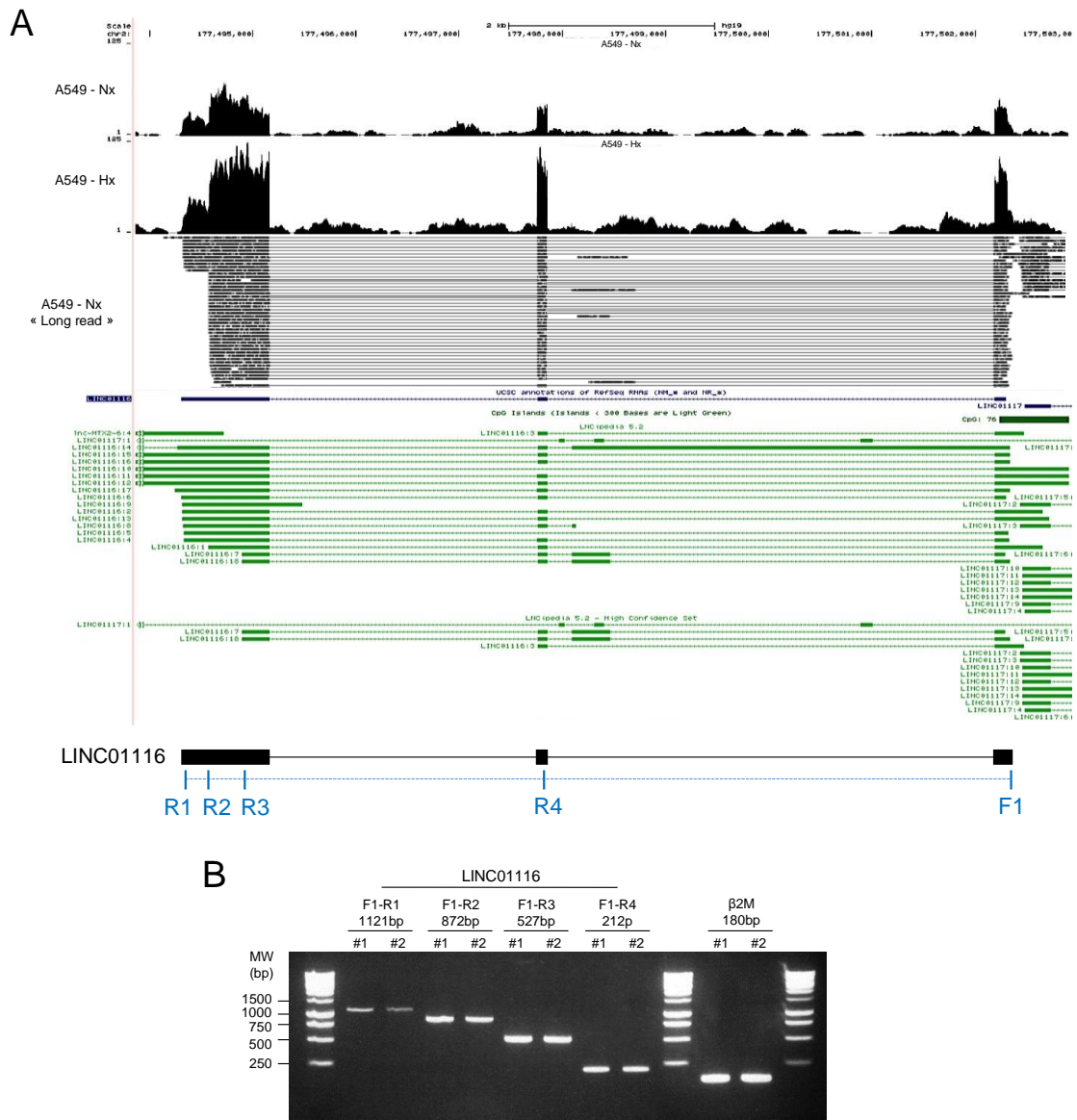
montre une activation qui débute après 45 min d'adhésion, puis qui se renforce à 90 min et se maintient au delà 4h. Dans les cellules réprimées pour l'expression de LINC01116, la phosphorylation de FAK, donc son activation est fortement réduite (Figure S10A). Ceci se traduit par une forte diminution de l'étalement des cellules et par un nombre restreint d'adhésions focales visualisées par le marquage en immunofluorescence de la paxilline (Figure S10B).

Par ailleurs, des travaux ont mis en évidence une faible induction de HEY2, un gène cible de la voie NOTCH1 en réponse à l'adhésion de cellules HUVEC sur de la fibronectine <sup>427</sup>. Nous avons confirmé cette induction, et montré qu'elle est absente dans les cellules exprimant de faibles niveaux de LINC01116 (Figure S10A). De plus, l'induction par l'adhésion, de l'expression des ARNm de HEY2 et de HEY1 et DII4, d'autres gènes cibles de la voie NOTCH1, est également réduite dans les cellules traitées avec des siARNs ciblant LINC01116 (Figure S10C) <sup>83,428-430</sup>. Ce défaut de signalisation de la voie NOTCH1 pourrait traduire un défaut d'adhésion ou un défaut d'interactions entre les cellules <sup>429</sup>.

D'autre part, pour évaluer l'implication de LINC01116 dans le contrôle de l'expression génique nous avons analysé les conséquences de son inhibition par ARN interférence sur le transcriptome de cellules endothéliales HMEC1 puis de cellules tumorales A549 par hybridation de puces à ADN. Dans les cellules HMEC1, le profilage transcriptomique a révélé l'altération significative de 765 et 322 gènes, respectivement dans les cellules traitées avec le si-LINC01116#1 ou le si-LINC01116#2. L'analyse comparative montre que seulement 101 gènes sont communément impactés par ces deux ARNs interférents (Figure 7AB). Leur annotation fonctionnelle à l'aide du logiciel Ingenuity Pathway Analysis (IPA), montre que ces gènes sont impliqués dans les processus d'interactions et de signalisation intercellulaires, de motilité, de survie et de croissance des cellules endothéliales (Figure 7C). Parmi les réseaux géniques altérés dans les cellules où LINC01116 est réprimé, nous avons noté la répression de réseaux régulés par les facteurs de transcription NKX2-3 et PDX1 (Figure 7DE). Dans les cellules A549, nous avons observé une altération significative de 186 gènes dans les cellules transfectées par le si-LINC01116#1 et cultivées en normoxie ou en hypoxie (Figure S11A). Leur annotation fonctionnelle montre qu'ils sont impliqués dans les mêmes processus que ceux identifiés dans les cellules endothéliales (Figure S11B). Cependant, bien que l'on retrouve l'altération des réseaux géniques contrôlés le facteur de transcription NKX2-3 et PDX1 (Figure S11DE), seulement 10 gènes dérégulés dans les cellules tumorales A549 sont communs à la signature génique altérée dans les cellules endothéliales HMEC1 (Figure S11C).

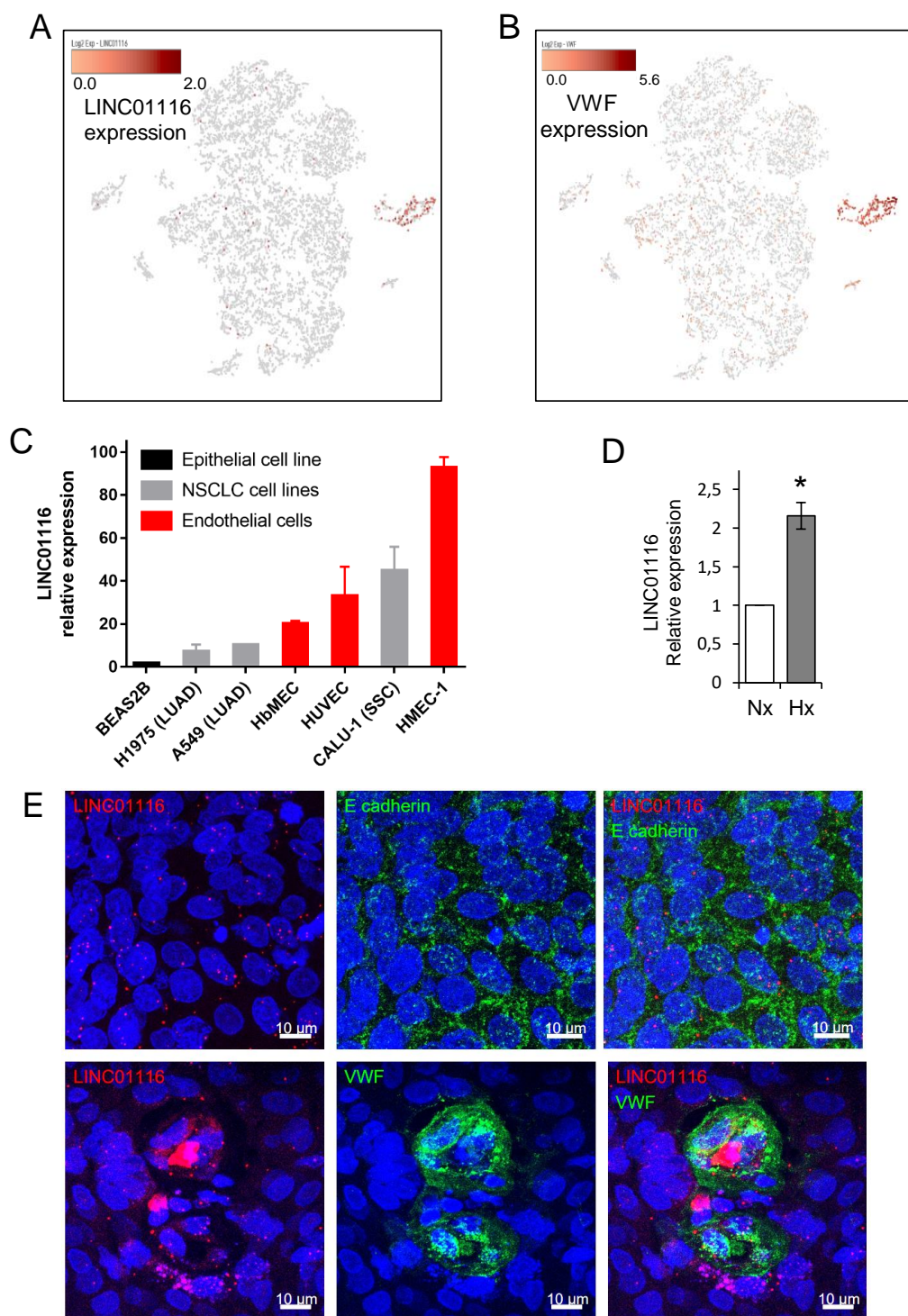


**Figure 1 – LINC0116, un lincARN induit par l’hypoxie dans les ADC pulmonaires et associé à une diminution de la survie des patients.** **A.** Diagramme de Venn présentant les lincARNs candidats dont l’expression est i) modulée en fonction du statut hypoxique des tumeurs d’ADC des patients de la cohorte du CHU de Nice (n=57), et/ou ii) en réponse à la culture de cellules A549 en condition hypoxique *in vitro* (1,2% O<sub>2</sub>), et/ou iii) associée au pronostic des patients. **B.D.** Box-plots montrant l’expression de LINC0116 en fonction du statut hypoxique des tumeurs d’ADC des cohortes de patients du CHU de Nice (n=57, **B**) et du TCGA (n=522, **D**). La pertinence statistique des comparaisons est déterminée par le test de Wilcoxon. **C.E.** Représentation Kaplan-Meier en quartile de la corrélation entre l’expression de LINC0116 et la survie globale des patients des cohortes locale (**C**) et du TCGA (**E**). Les p-values sont calculées en utilisant le log-rank test.



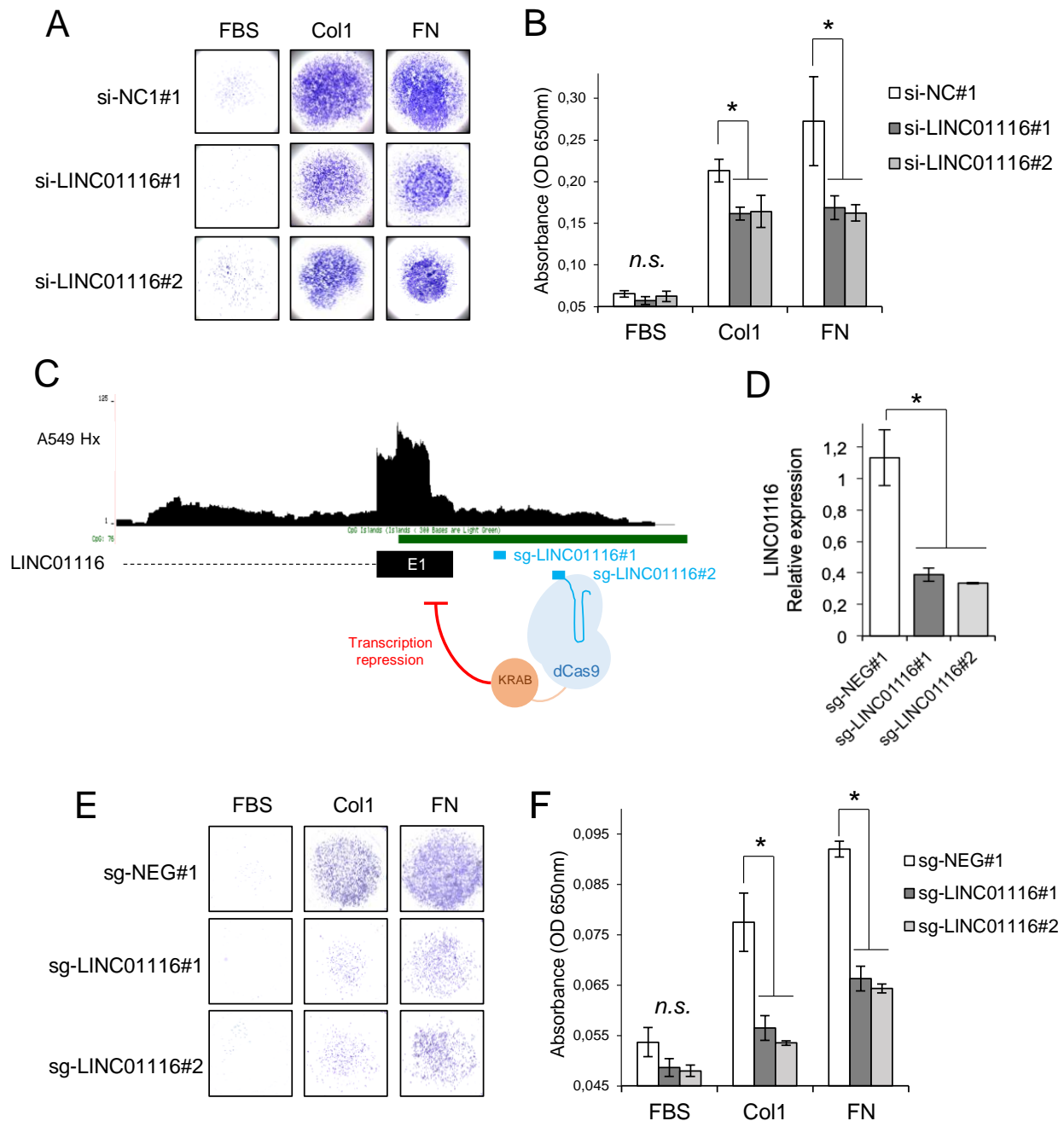
**Figure 2 – Caractérisation structurale de LINC01116 dans la lignée d’ADC A549. : A.** La séquence de LINC01116 a été déterminée par séquençage d’ARN de la lignée A549 cultivée en normoxie (Nx) ou en hypoxie (Hx) pendant 24h. Les reads courts (NextSeq500 Illumina, 150 bp paired-end reads représentés en format big wig) et longs (PromethION, Oxford Nanopore Technology, full length représentés en format bam) générés ont été alignés sur le génome de référence hg19 puis visualisés sur le navigateur UCSC. La région génomique de LINC01116 montre un transcript de 1121pb constitué de 3 exons. **B.** Validation par RT-PCR de la structure intron-exon de LINC01116 à partir de deux préparations d’ADNc de cellules A549 cultivées en hypoxie.





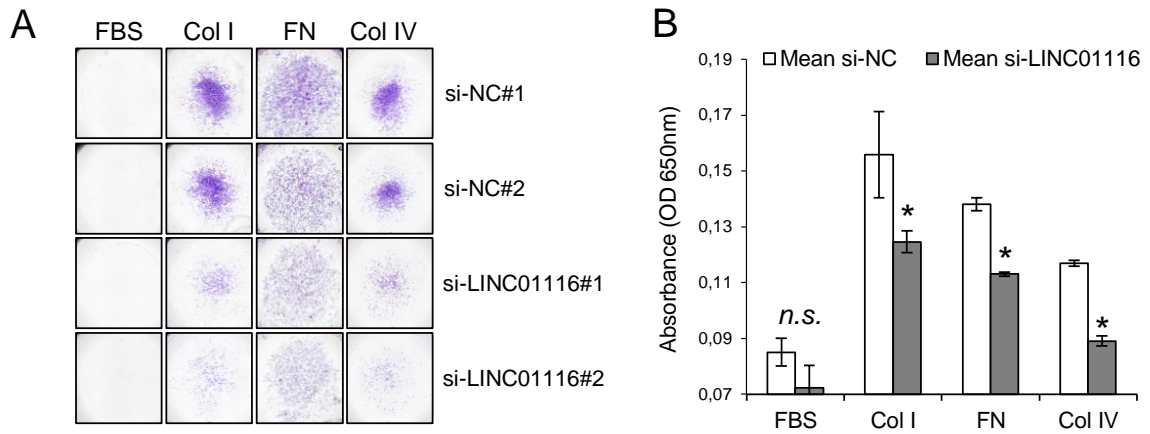
**Figure 3 – LINC01116 est exprimé dans les cellules endothéliales.**

**AB.** Analyse transcriptomique par séquençage ARN sur cellule unique à partir de biopsies de turbinectomies représentée sous forme de t-SNE et montrant la co-expression de LINC01116 et de l'ARNm de VWF dans les cellules endothéliales. **C.** Analyse par RT-qPCR de l'expression relative de LINC01116 dans différentes lignées de CBNPC (ADC ou SCC), d'une lignée épithéliale bronchique (BEAS2B) et de modèles de cultures primaires (HUVEC) ou de lignées (HdMEC1 et HbMEC) de cellules endothéliales. Les résultats représentent la moyenne de 3 expériences indépendantes +/- SD. **D.** Induction de l'expression de LINC01116 par l'hypoxie dans les cellules endothéliales HUVEC (n=3, moyenne +/- SD). **E.** Analyse de l'expression de LINC01116 par smRNA FISH sur des coupes de biopsies d'ADC. Des co-marquages en immunofluorescence ont été réalisés avec des anticorps anti-cadhérine E et anti-VWF afin de marquer respectivement les cellules épithéliales ou les cellules endothéliales. Les noyaux sont marqués au DAPI.



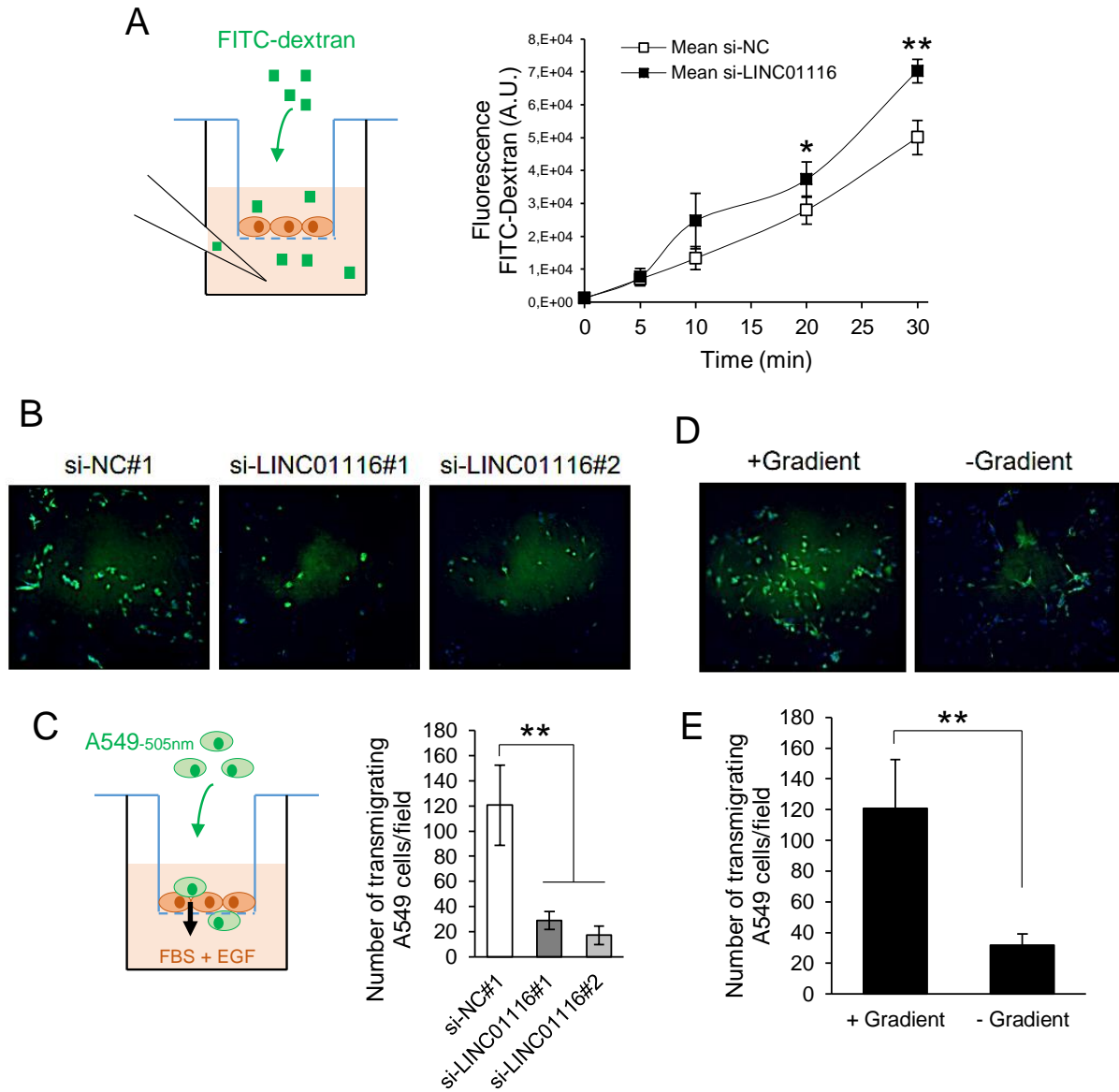
**Figure 4 – Effets de la répression de LINC01116 par ARN interférence ou par CRISPRi sur l'adhésion des cellules d'ADC A549.**

**AB.** Après 48 h de transfection avec les siARNs indiqués, les cellules (30 000/puit) sontensemencées dans des puits d'une plaque 96 puits préalablement coâtés avec du SVF (1%), du collagène 1 (Col1, 20 µg/ml) ou de la fibronectine (FN, 20 µg/ml). Après 1h à 37°C, les puits sont lavés plusieurs fois et les cellules adhérentes sont marquées au crystal violet (0,1% + 25% méthanol), visualisées au Cytation 5 (Biotek) (**A**), lysées au triton 0,2% puis quantifiées par mesure de l'absorbance à 650 nm (**B**). Les résultats représentent la moyenne de 2 expériences indépendantes réalisées en triplicats +/- SD. (\*: pvalue < 0,05). **C.** Schéma représentant les guides ARN utilisés pour réprimer LINC01116 dans l'approche CRISPRi. **D.** Analyse par RT-qPCR de l'efficacité de la répression de LINC01116 par CRISPRi. **E.** Les effets de l'inactivation de LINC01116 par le système CRISPRi ont été analysés et quantifiés comme en **AB**.



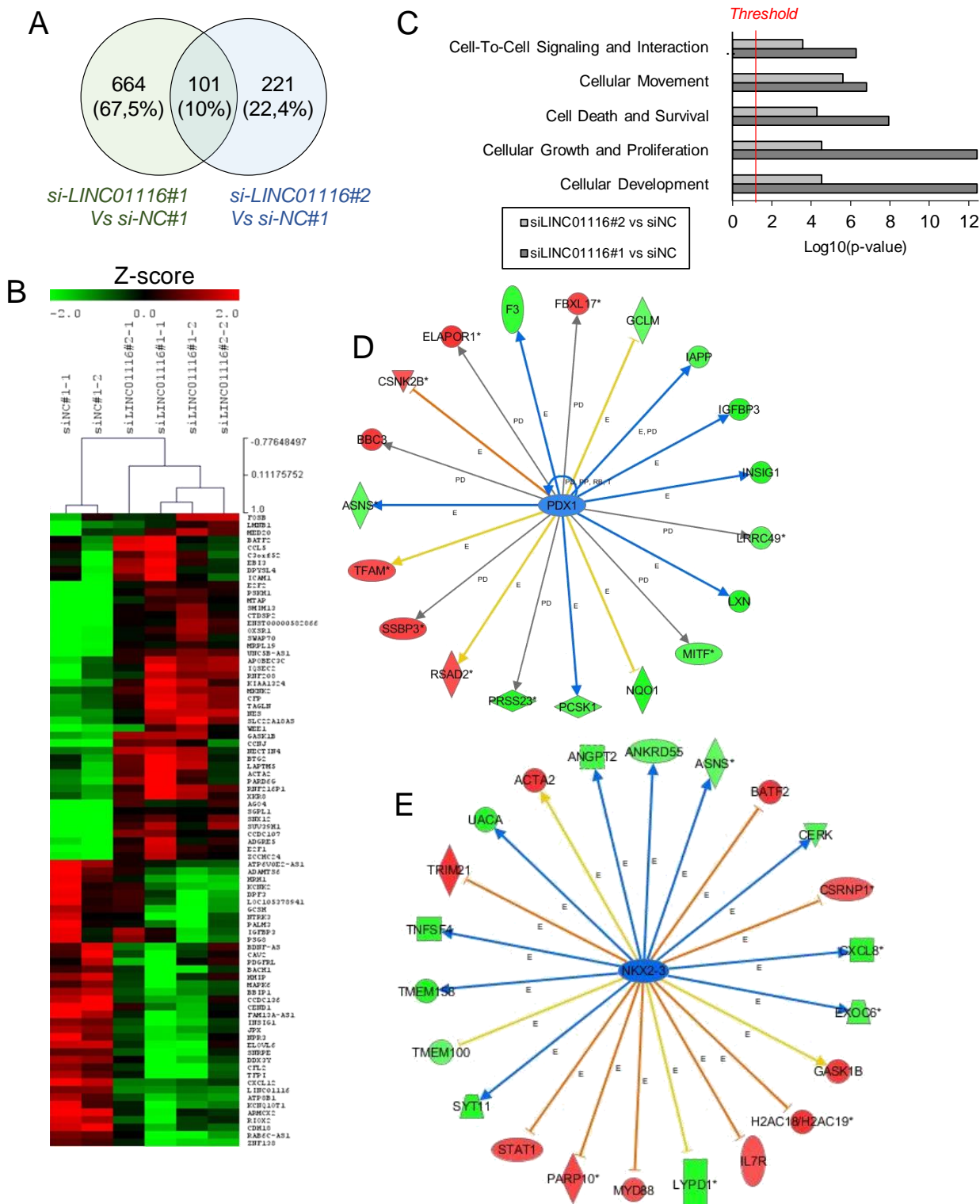
**Figure 5 – Effets de l’inhibition de LINC01116 par ARN interférence sur l’adhésion des cellules endothéliales.**

**AB.** Après 48 h de transfection avec les siARNs indiqués (10 nM), les cellules HUVEC (30 000/puit) sontensemencées dans des puits (plaque 96 puits) préalablement coatés avec du SVF (1%), du collagène 1 (Col1, 20 µg/ml), de la fibronectine (FN, 20 µg/ml) ou du collagène IV (Col IV, 20 µg/ml). Après 2h à 37°C, les puits sont lavés et les cellules adhérentes marquées au crystal violet, visualisées au Cytation (**A**), lysées puis quantifiées par mesure de l’absorbance à 650 nm (**B**). Les histogrammes représentent la moyenne de 3 expériences indépendantes réalisées en triplicats +/- SD. (\*:pvalue < 0,05).



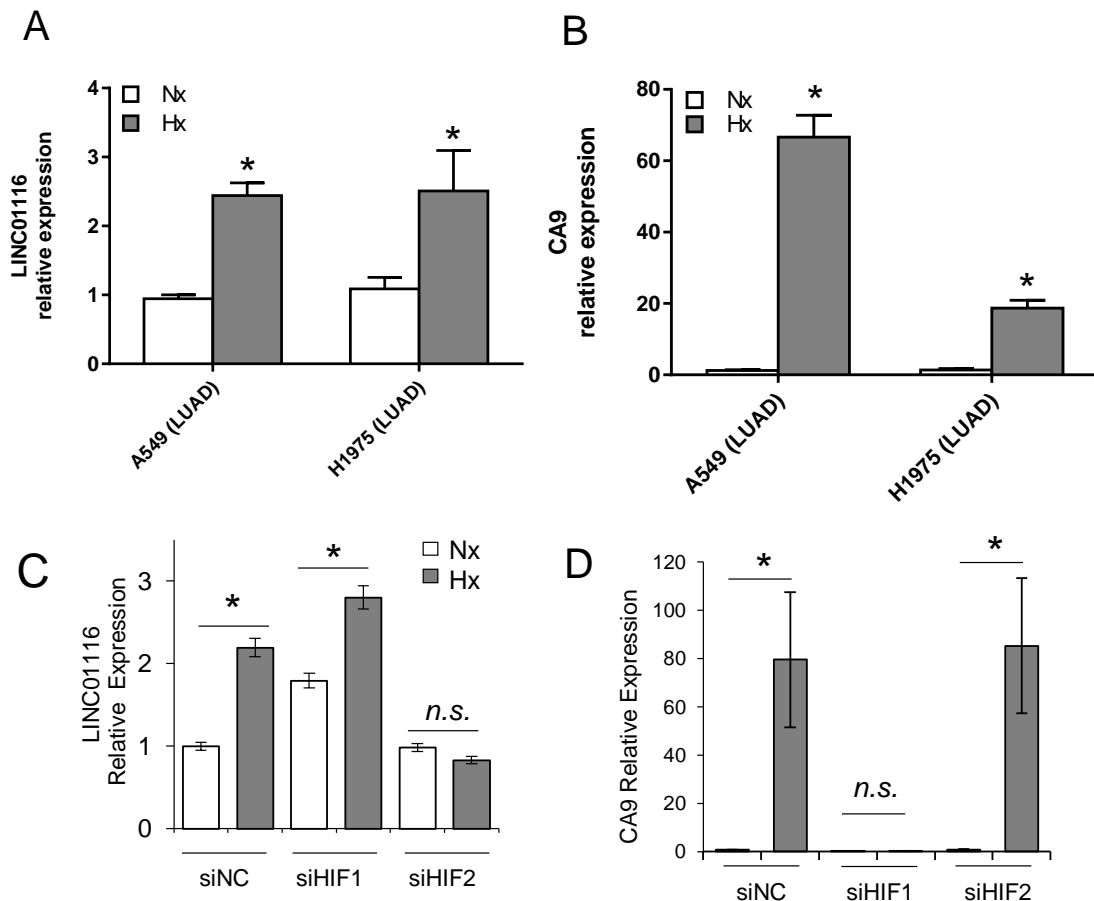
**Figure 6 – Effets de l’inhibition du transcrit LINC0116 sur la perméabilité de l’endothélium *in vitro*.**

**A-C.** Les cellules HUVEC transfectées pendant 48 h avec les siARNs indiqués ont été ensemencées à confluence (30000 cellules/insert) dans des inserts de chambre de Boyden pré-coatés avec du collagène I. **A.** Après 24 h, du Dextran-FITC est ajouté dans la chambre supérieure. Des aliquots de milieu sont prélevés au cours du temps dans la chambre inférieure afin d’évaluer la diffusion du dextran par mesure de la fluorescence. Les résultats montrent la moyenne de triplicats +/- SD et sont représentatifs de 3 expériences indépendantes. **B-C.** Après 24 h, des cellules d’ADC A549 (30000) marquées à l’aide d’un marqueur vital (Bodipy-505nm) sont déposées dans la chambre supérieure dans du milieu dépourvu de sérum puis attirées à travers l’endothélium sur la face inférieure des inserts par ajout de milieu contenant 10% SVF + EGF (10ng/mL) dans la chambre inférieure. Les cellules sont visualisées en microscopie à fluorescence puis quantifiées. **D-E.** Des cellules A549 ont été déposées sur un endothélium de cellules HUVEC non transfectées en présence ou en l’absence de SVF+EGF, et attirées dans la chambre inférieure par du milieu +SVF+EGF. Les résultats représentent la moyenne de deux expériences indépendantes en triplicats +/- SD. (\*:pvalue < 0,05; \*\*:pvalue < 0,001).



**Figure 7 – Effets de l’inhibition de LINC01116 sur le transcriptome des cellules endothéliales HMEC1.**

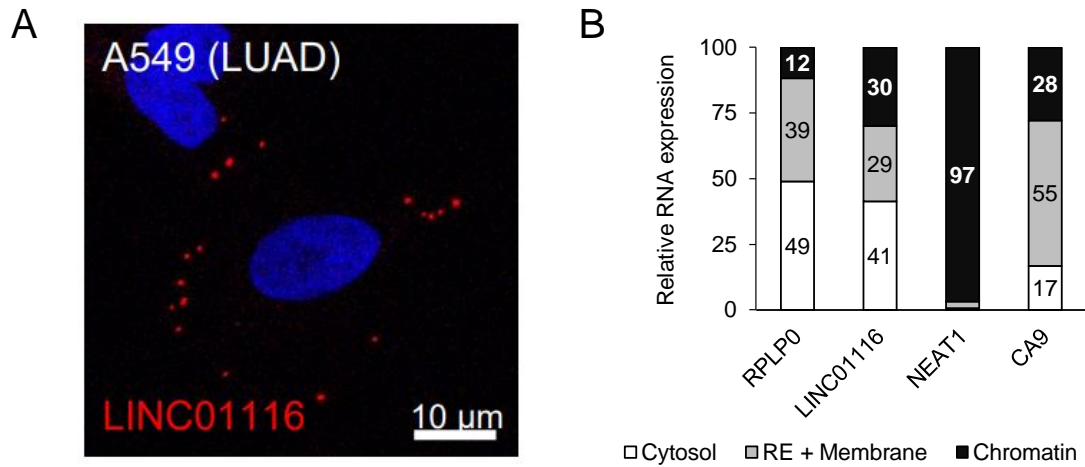
**A.** Diagramme de Venn représentant les gènes différentiellement exprimés entre les cellules HMEC1 transfectées pendant 48h avec des si-LINC01116#1, si-LINC01116#2 comparé à la transfection de si-NC#1 (log Fold Changel > 0,5, adj pvalue<0,05, Intensité >6). **B.** Représentation en échelle de couleur de la classification hiérarchique de 101 gènes différentiellement exprimés entre des cellules HMEC1 transfectées pendant 48h avec des si-LINC01116#1, si-LINC01116#2 ou si-NC#1. **C.** Annotation fonctionnelle réalisée à l’aide du logiciel Ingenuity Pathway (IPA) montrant les principaux processus biologiques dérégulés par la transfection des si-LINC01116#1 et si-LINC01116#2 comparés au si-NC#1 (log Fold Changel > 0,6, adj pvalue<0,05, Intensité >6). **D.E.** Exemples de réseaux géniques altérés par l’inhibition de LINC01116 issus de l’analyse IPA. Les flèches oranges et bleues représentent respectivement les prévisions d’activation ou d’inhibition. Les gènes réprimés sont en vert, les gènes induits en rouge.



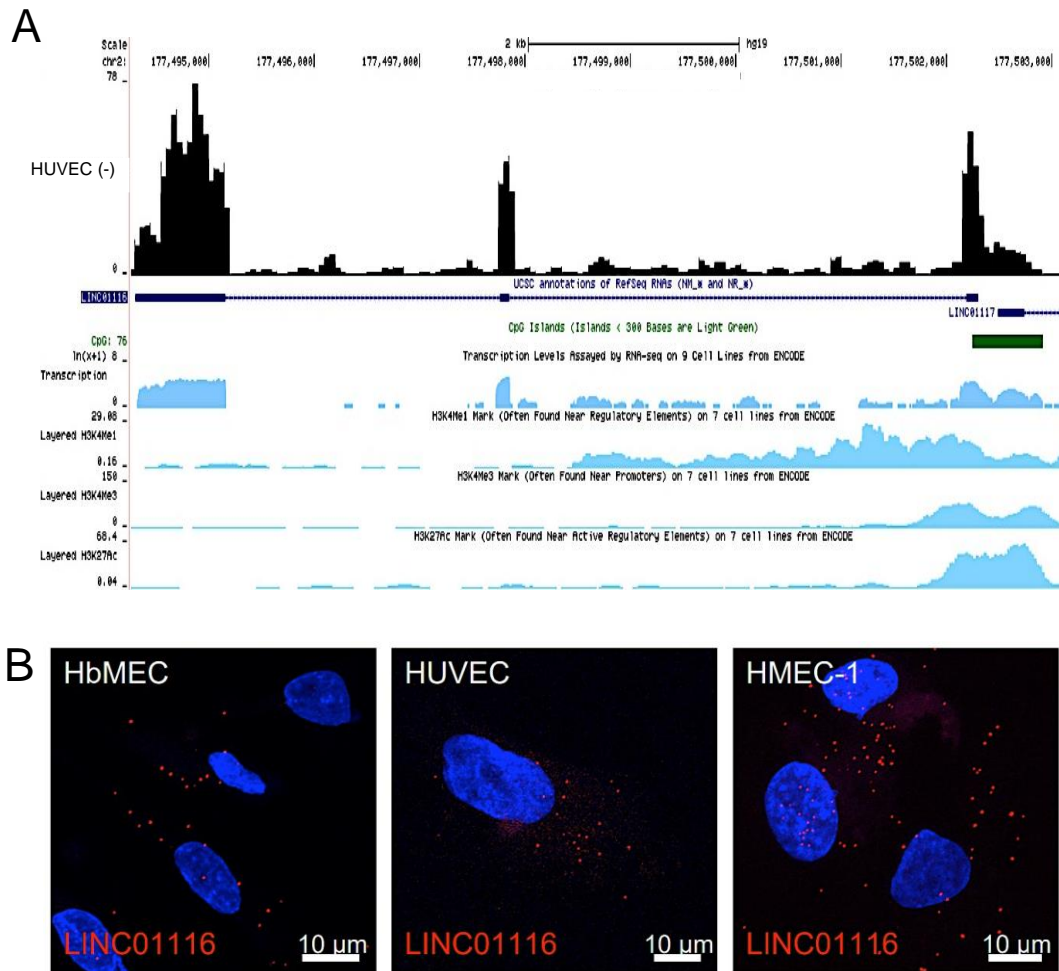
**Figure S1 - Régulation de l'expression de LINC01116 par l'hypoxie dans les cellules d'ADC.**

**AB.** Les cellules des lignées d'ADC A549 et H1975 ont été cultivées en normoxie (Nx) ou en hypoxie (Hx) pendant 24 h et l'expression de LINC01116 et CA9 analysée par RT-qPCR. L'expression de CA9 sert de contrôle positif à la réponse hypoxique. **CD.** Les cellules A549 ont été transfectées avec des siARNs (25 nM) contrôle (siNC) ou ciblant HIF1 ou HIF2, puis cultivées 24 h en normoxie ou en hypoxie. L'expression de LINC01116 et CA9 a été analysée par RT-qPCR et normalisée par rapport à RPLP0.



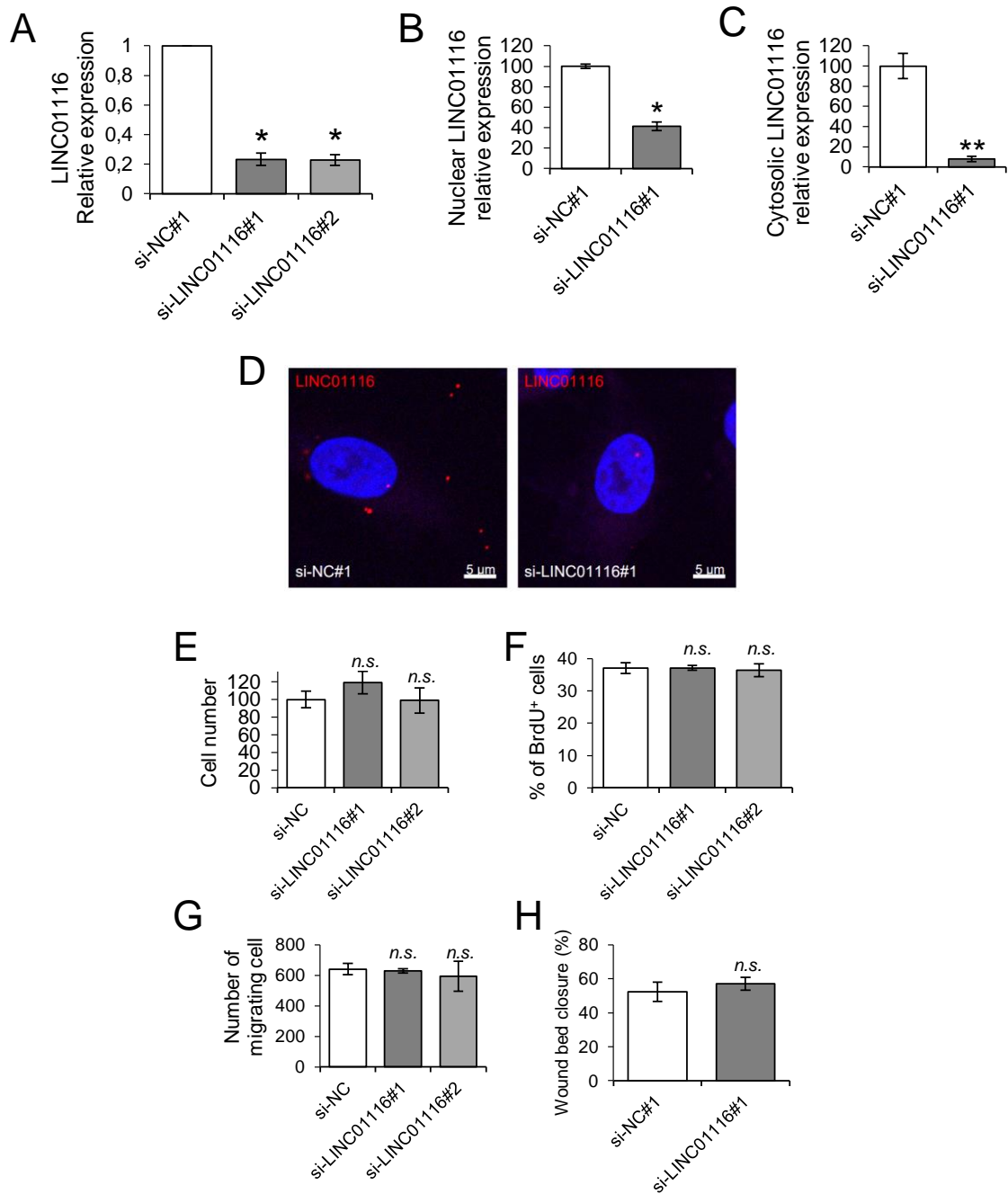


**Figure S2 – Localisation subcellulaire de LINC01116 dans les cellules A549.** Analyse de l'expression de LINC01116 par : **A.** smRNA FISH. Les noyaux sont marqués au DAPI. **B.** Fractionnement subcellulaire couplé à de la RT-qPCR. Les ARNs ont été extraits à partir des fractions cytosolique, membranaire incluant le reticulum endoplasmique et nucléaire, et analysés pour l'expression des lncARNs LINC01116 et NEAT1 (nucléaire), et des ARNm de RPLP0 et CA9. Les résultats sont représentatifs de trois expériences indépendantes.



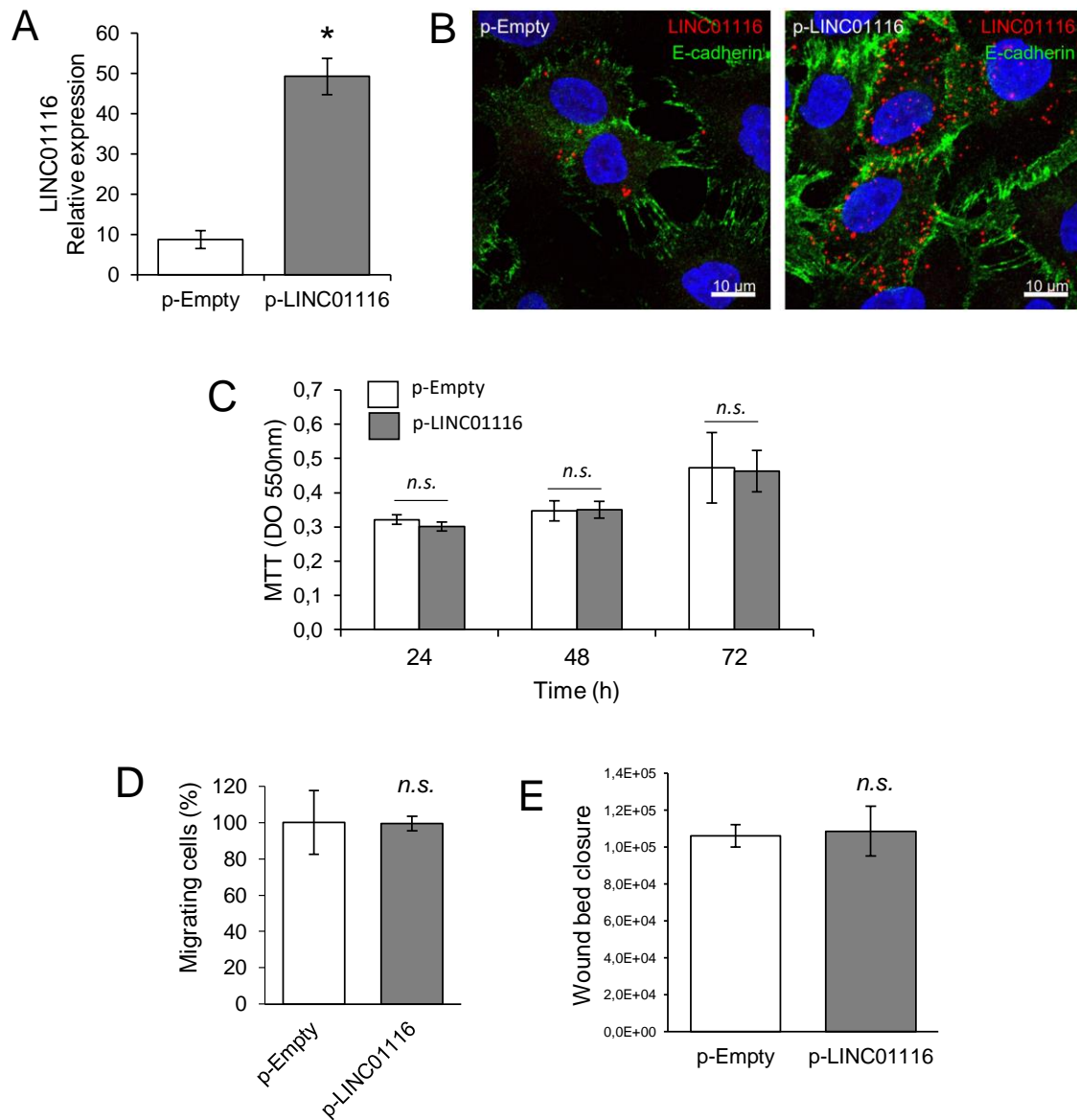
**Figure S3 – Caractérisation structurale de LINC01116 dans les cellules endothéliales. : A.** Analyse de LINC01116 par séquençage d'ARN de cellules HUVEC. Les reads générés sont visualisés sur le navigateur UCSC. Les marques d'histone H3K4Me1, H3K4Me3 et H3K27Ac de cellules HUVEC issues de données publiques du projet ENCODE sont représentées. **B.** Marquage par smRNA FISH à l'aide de sondes ciblant LINC01116 sur des modèles de cellules endothéliales (HbMEC, HUVEC, HdMEC1). Les noyaux ont été marqué au DAPI et les cellules ont été visualisé par microscopie confocale.





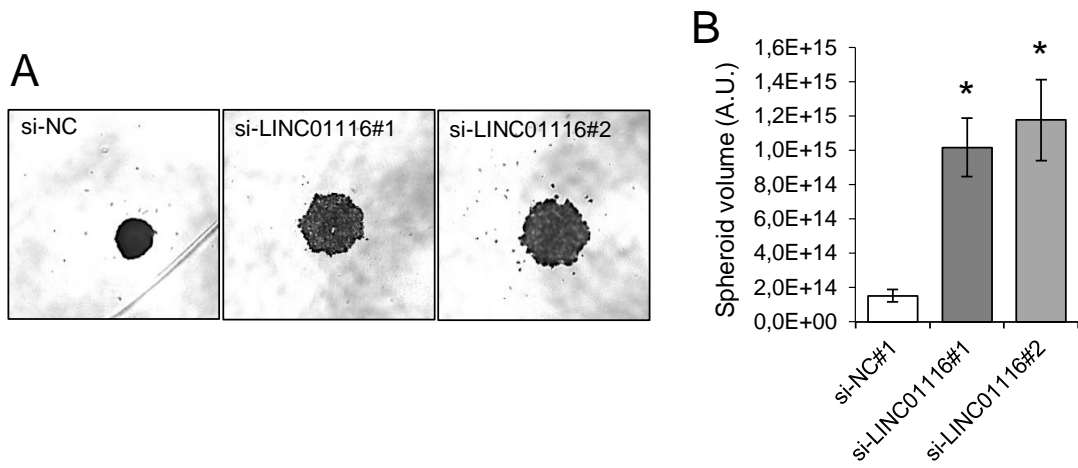
**Figure S4 – Effets de l’inhibition de LINC01116 par ARN interférence sur la croissance et la mobilité des cellules d’ADC A549.**

**A-C.** Analyse par RT-qPCR de l’expression de LINC01116 dans des cellules A549 traitées avec des siARNs contrôle (siNC1) ou ciblant le transcrite 48h après transfection (n=3, moyenne +/- SD). L’expression a été mesurée dans des extraits cellulaires totaux (**A**), ou dans des fractions nucléaire (**B**) ou cytosolique (**C**). **D.** Analyse des effets de l’interférence ARN par smRNA FISH. **EF.** Effets des siARNs ciblant LINC01116 sur la prolifération mesurée par comptage cellulaire 72h post-transfection (**E**), ou par marquage au BrdU des cellules en phase S (**F**), (moyenne de 3 expériences +/- SD). **GH.** Effets des siARNs ciblant LINC01116 sur la migration cellulaire mesurée dans des chambres de Boyden (**G**) ou par « scratch-assay » (**H**) comme précisé dans la section matériels et méthodes (moyenne de 3 expériences +/- SD). (\*: pvalue<0,05).



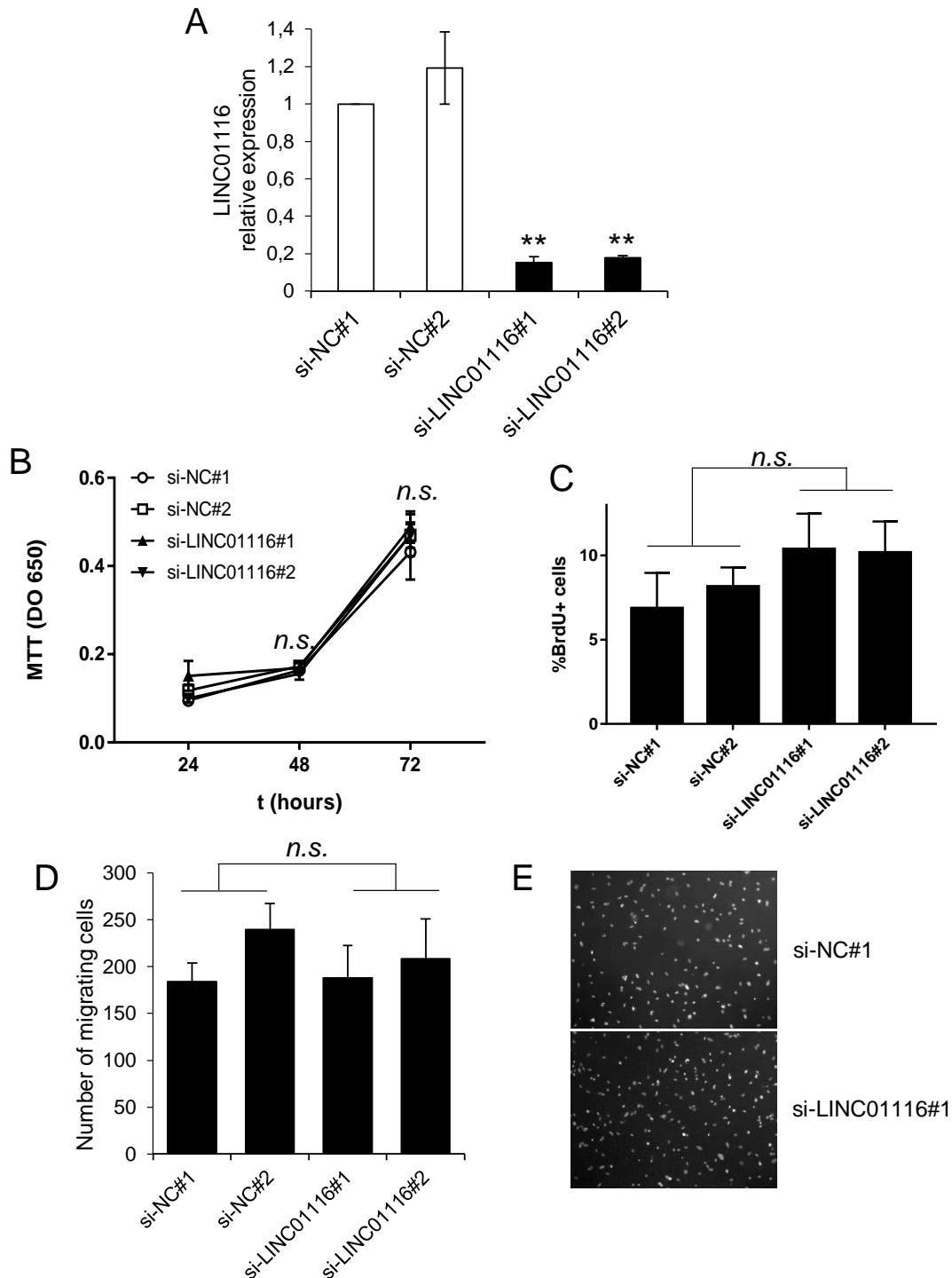
**Figure S5 – Effets de la surexpression de LINC01116 sur la croissance et la mobilité des cellules d'ADC A549.**

Les cellules A549 ont été transduites par infection lentivirale avec un vecteur vide ou exprimant le transcrit LINC01116. La surexpression du transcrit a été analysée par RT-qPCR (A) et par smRNA FISH (B). Analyse des effets de la surexpression de LINC01116 sur la prolifération mesurée par un essai MTT à différents temps après transfection (C). Les résultats sont représentatifs de trois expériences indépendantes. DE. Effets de la surexpression de LINC01116 sur la migration cellulaire mesurée dans des chambres de Boyden (D) ou par « scratch-assay » (E) comme précisé dans la section matériels et méthodes. Les résultats sont représentatifs de 2 expériences indépendantes. (\*:  $p$ value<0,05).



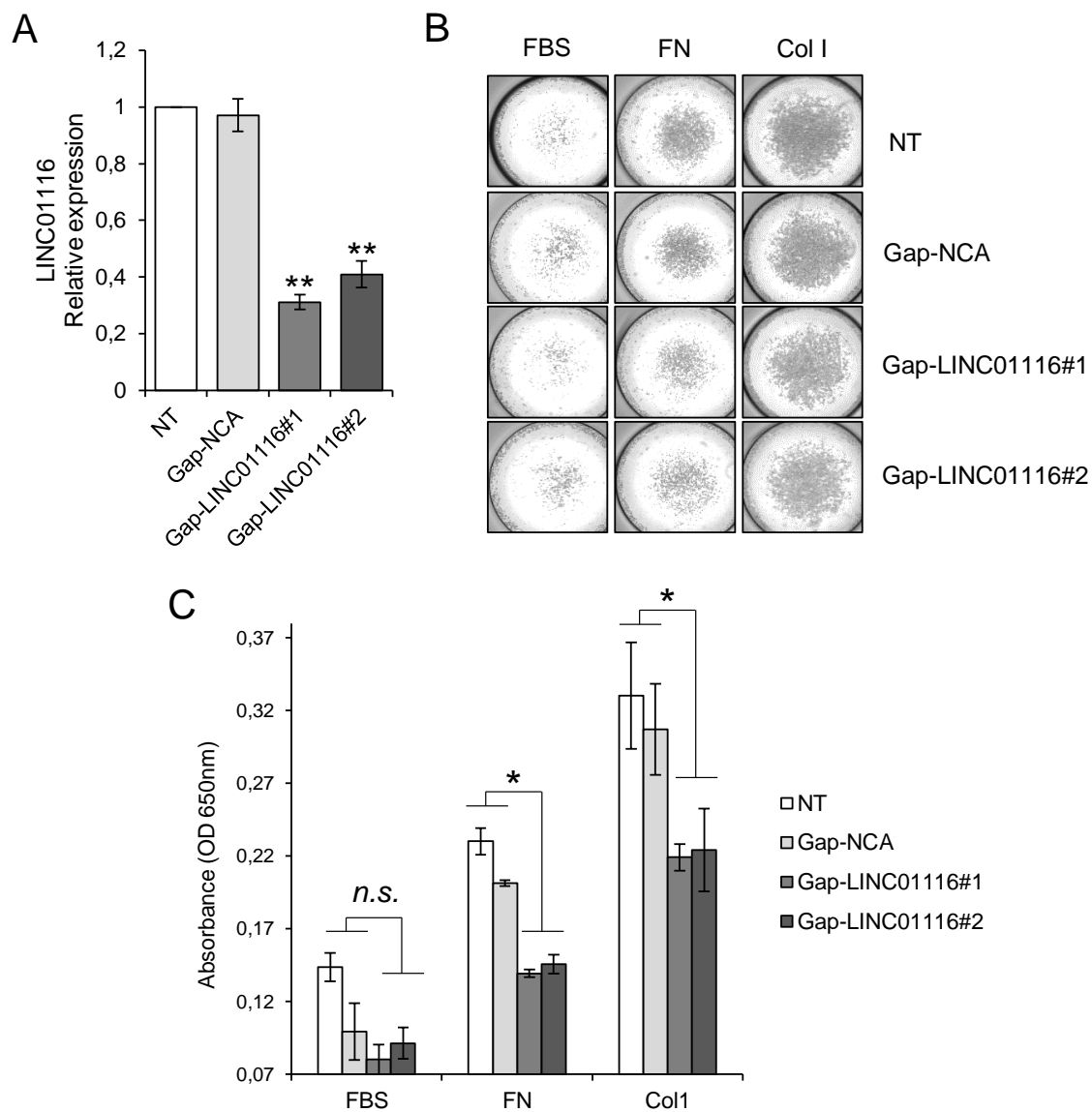
**Figure S6 – Effets de l’ARN interférence ciblant LINC01116 sur la formation de sphéroïdes de cellules d’ADC A549.**

Les cellules A549 sont transfectées avec des siARN (10nM) contrôle (NC1) ou spécifiques de LINC01116. Après 48h les cellules (10 000) sont ensemencées dans des puits à fond rond à faible adhérence afin d’induire la formation de sphéroïdes. Après 24 h les sphéroïdes sont photographiés au cytation et leur volume calculé. Les résultats correspondent à la moyenne du volume de 12 sphéroïdes +/- SD et sont représentatifs de 3 expériences indépendantes. (\*: pvalue<0,05).



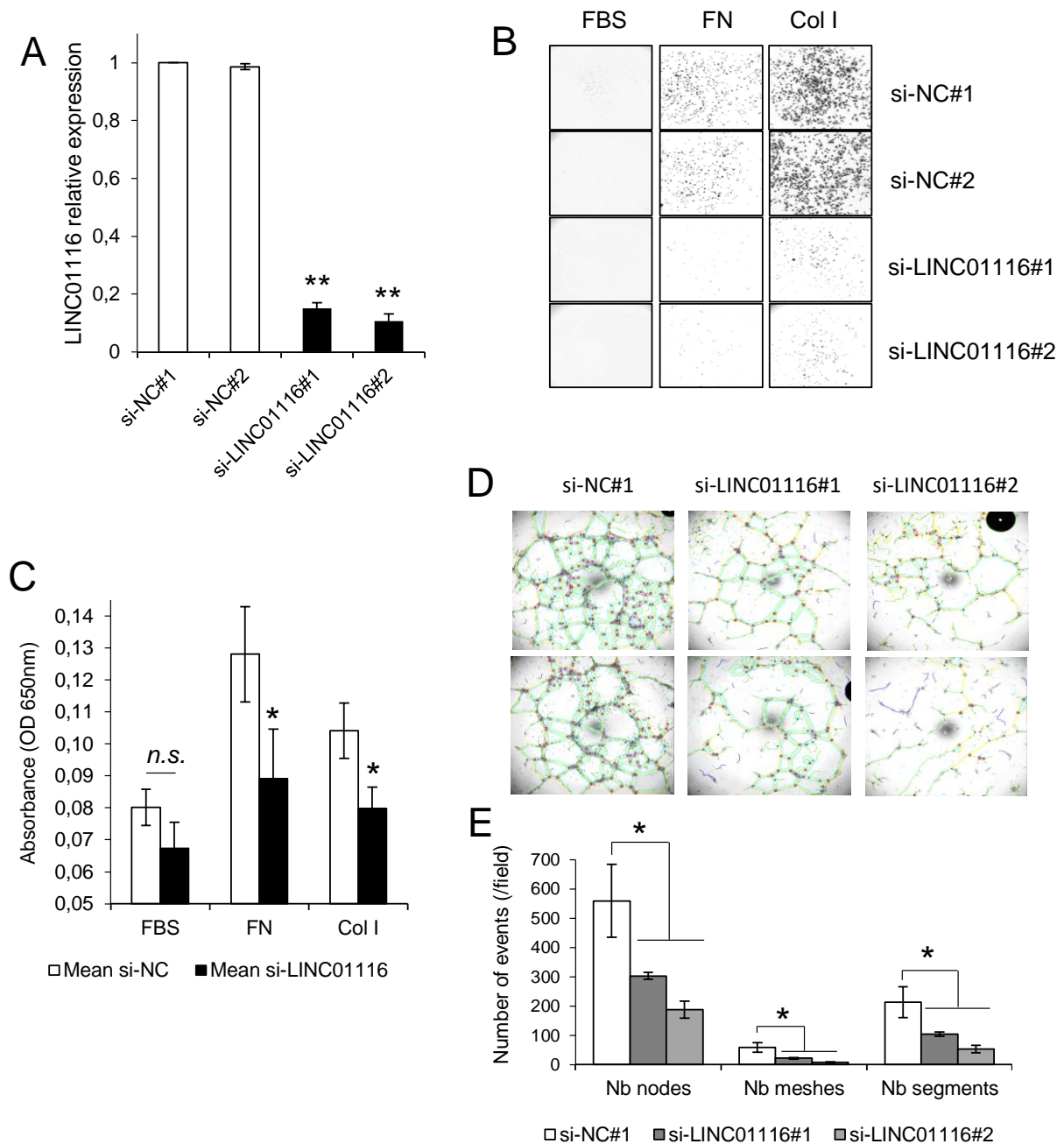
**Figure S7 - Effets de l'inhibition de LINC01116 par ARN interférence sur la croissance et la mobilité des cellules endothéliales HUVEC.**

**A.** Analyse par RT-qPCR de l'expression de LINC01116, 48 h après la transfection de cellules HUVEC par des siARNs contrôle (siNC1, siNC2) ou ciblant le transcrit (n=3, moyenne +/- SD). **BC.** Analyse des effets de l'interférence ARN LINC01116 sur la prolifération mesurée par un essai MTT à différents temps après transfection (**B**) (moyenne de 2 expériences +/- SD), ou par marquage au BrdU des cellules en phase S (**C**). **DE.** Effets des siARNs ciblant LINC01116 sur la migration cellulaire mesurée dans des chambres de Boyden. (moyenne de 3 expériences +/- SD). (\*: pvalue<0,05).



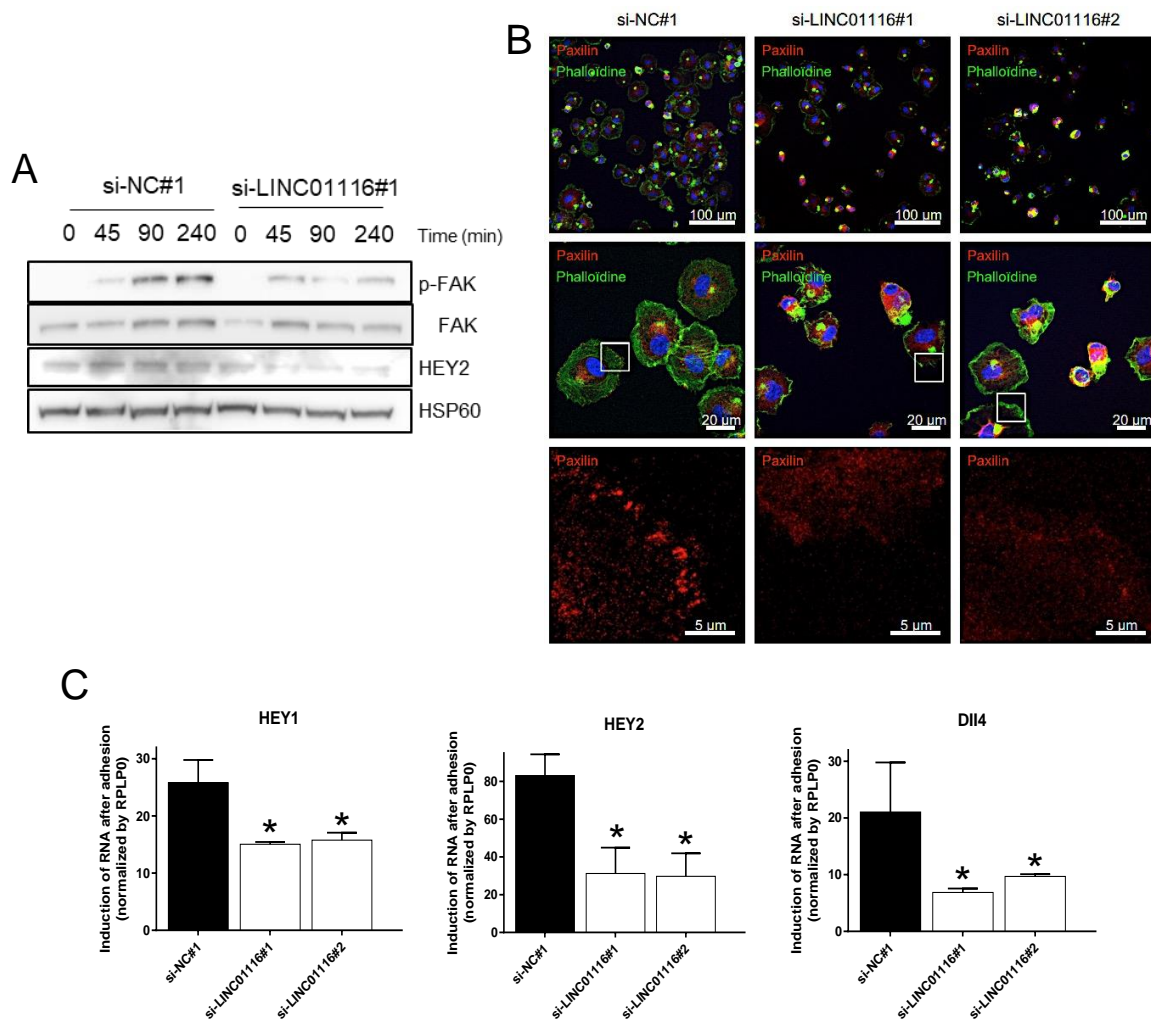
**Figure S8 – Effets d’oligonucléotides antisens de type gapmers ciblant LINC01116 sur l’adhésion des cellules endothéliales HUVEC.**

**A.** L’expression de LINC01116 a été mesurée par RT-qPCR dans des cellules HUVEC non transfectées (NT), ou 48h après la transfection avec un gapmer contrôle (Gap-NCA) ou deux gapmers spécifiques du transcrit LINC01116 (Gap-LINC01116#1 et Gap-LINC01116#2) à la concentration de 20 nM. **BC.** Les cellules (30000/puit) ont été ensemencées dans des puits coatés avec du sérum de veau fœtal (FBS), de la fibronectine (FN) ou du collagène I (Col I). Après 1h à 37°C, les puits ont été lavés et les cellules adhérentes marquées au crystal violet, visualisées au Cytation (**B**), lysées puis quantifiées par mesure de l’absorbance à 650 nm (**C**). Les histogrammes représentent la moyenne de 2 expériences indépendantes réalisées en triplicats +/- SD. (\*:pvalue < 0,05).

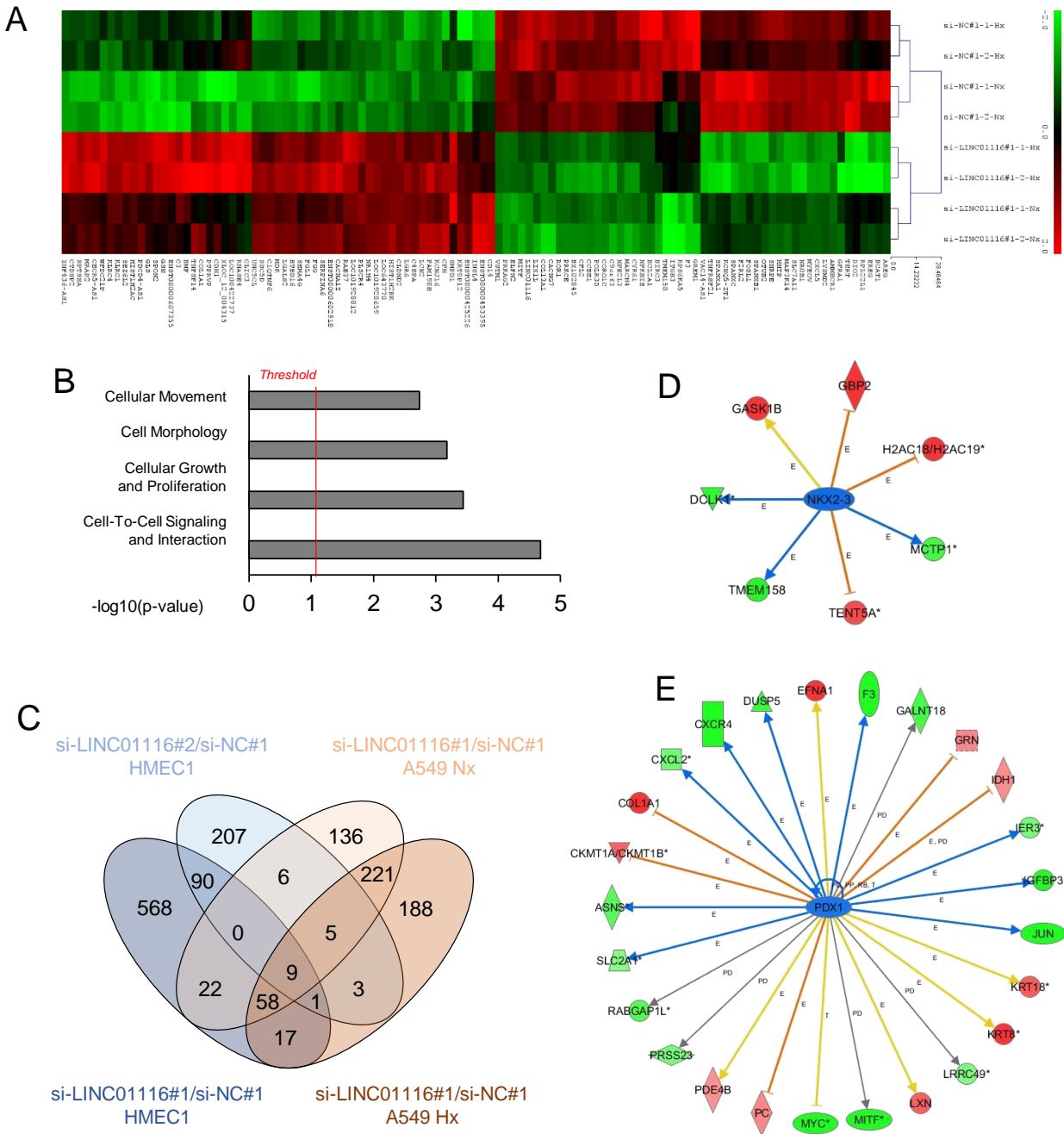


**Figure S9 – Effets de l’inhibition de LINC01116 sur l’adhésion des cellules HMEC1.**

**A.** L’inhibition de LINC01116 a été mesurée par RT-qPCR 48h après la transfection des cellules HMEC1 avec des ARN interférants (10 nM). **B-C.** Adhésion des cellules transfectées, dans des puits coâtés avec du SVF (1%), du collagène I (20 µg/ml) ou de la fibronectine (FN, 20 µg/ml). Après 1h à 37°C, les puits sont lavés et les cellules adhérentes marquées au crystal violet, visualisées au Cytation (**B**), lysées puis quantifiées par mesure de l’absorbance à 650 nm (**C**). (Moyenne de triplicats +/- SD d’une expérience représentative de deux expériences). **D-E. Expérience de formation de tubules :** Les cellules transfectées (10000 cellules) sont ensemencées dans des puits pré-coâtés avec du Matrigel puis incubées 18h et visualisées au Cytation 5 en contraste de phase (**D**). Les images ont été analysées pour quantifier le nombre de branchements, de mailles et de segments (**E**). Moyenne de deux expériences +/- SD. (\*: pvalue < 0,05).



**Figure S10 – Effets de l’inhibition de LINC01116 par ARN interférence sur la formation d’adhésions focales et sur l’activation de la voie NOTCH.** Les cellules HUVEC ont été transfectées avec des siARNs (10 nM) contrôle ou ciblant LINC01116 puis incubées 48h. **A.** Les cellules sont ensemencées sur des boîtes de pétri coatées avec de la fibronectine puis lysées à différents temps d’adhésion et la phosphorylation de FAK sur la tyrosine 397 et l’induction de HEY2 ont été analysées par Western Blot. La membrane hybridée avec phospho-FAK été strippée puis réhybridée avec un anticorps anti-FAK. L’hybridation de HSP60 sert de contrôle de charge. **B.** Les cellules sont ensemencées pendant 45 min sur des lamelles de verre coatées avec de la fibronectine puis marquées en immunofluorescence avec un anticorps anti-paxiline et de la phalloïdine-Alexa Fluor 488. Les noyaux ont été marqué au DAPI. Les résultats sont représentatifs de trois expériences indépendantes. **C.** Les cellules HUVEC ont été ensemencées pendant 4 h sur des boîtes de pétri coatées avec de la fibronectine (20  $\mu$ g/ml) puis lysées en Trizol pour extraire les ARN. L’expression des gènes HEY1, HEY2 et DII4, induits par la voie NOTCH a été analysée par RT-qPCR et normalisée par rapport à RPLP0. Les résultats représentent l’induction de l’expression par rapport à des cellules maintenues en suspension et correspondent à la moyenne de 3 expériences indépendantes réalisées en triplicats +/- SD. (\*:pvalue < 0,001).



**Figure S11 – Effets de l’inhibition de LINC0116 sur le transcriptome des cellules d’ADC A549.**

**A.** Représentation en échelle de couleur de la classification hiérarchique de 186 gènes différentiellement exprimés entre des cellules A549 cultivées en normoxie ou en hypoxie et transfectées pendant 48h avec des si-LINC0116#1 ou si-NC#1 (log Fold Change >1, adj pvalue<0,05, Intensité >6). **B.** Annotation fonctionnelle réalisée à l’aide du logiciel Ingenuity Pathway (IPA) montrant les principaux processus biologiques dérégulés par la transfection du si-LINC0116#1 comparé au si-NC#1 (log Fold Change > 0,8, adj pvalue<0,05, Intensité >6). **C.** Diagramme de Venn représentant les gènes différentiellement exprimés dans les cellules transfectées avec les si-LINC0116#1, si-LINC0116#2 ou si-NC#1 dans les HMEC1 (log Fold Change > 0,5, adj pvalue<0,05, Intensité >6) et les A549 cultivées en normoxie (Nx) ou en hypoxie (Hx) (log Fold Change > 0,8, adj pvalue<0,05, Intensité >6). **D.E.** Exemples de réseaux géniques altérés par l’inhibition de LINC0116 issus de l’analyse IPA dans les cellules A549. Les flèches oranges et bleues représentent respectivement les prévisions d’activation ou d’inhibition. Les gènes réprimés sont en vert, les gènes induits en rouge.



## ***Discussion et perspectives***

### **LINC01116, un hypoxa-lnc marqueur de l'agressivité des adénocarcinomes broncho-pulmonaires.**

Comme nous l'avons vu précédemment, de nombreux exemples relient les lncARNs à l'agressivité des cancers. Dans cette étude, nous nous sommes focalisés sur LINC01116 qui est induit par l'hypoxie *in vitro* et corrélé au statut hypoxique de tumeurs d'ADC ainsi qu'à un mauvais pronostic des patients issus de deux cohortes (CHU de Nice et TCGA). Nous avons démontré que LINC01116 est induit par le facteur de transcription HIF2/EPAS1 dans une lignée d'adénocarcinomes pulmonaires (ADC). Ce facteur de transcription responsable en partie de la réponse hypoxique des cellules joue un rôle important dans la progression tumorale <sup>431</sup>, suggérant que les gènes qu'il régule ont également une fonction dans le développement tumoral.

La base de données LNCipédia (<https://lncipedia.org/>) décrit l'existence de 19 isoformes potentielles de transcrits LINC01116. Nous avons établi, par des approches de séquençage ARN générant des reads courts et longs puis par des validations par RT-PCR, que l'isoforme majoritaire dans nos modèles d'étude correspond à une isoforme cytosolique de 1121nt ayant la même structure exon-intron que l'isoforme LINC01116:17. De récents travaux ont déterminé par Northern Blot que dans les gliomes LINC01116 est un transcrit de 980nt <sup>432</sup>. Toutefois, ces auteurs n'ont pas précisé la structure exon-intron de ce transcrit. Ces différences pourraient s'expliquer par l'expression de variants spécifiques d'un tissu donné et pouvant exercer des fonctions différentes.

D'autre part, l'analyse du locus de LINC01116 nous a permis de déterminer qu'il s'agit d'un lncARN divergent (cf. III.3.D) dont le promoteur contient un large îlot CpG et promeut également l'expression de LINC01117. Des analyses *in silico* de culpabilité par association réalisées sur les données de séquençage ARN des ADC de la cohorte du TCGA ont montré que LINC01117 est le gène dont l'expression présente la meilleure corrélation avec celle de LINC01116 avec un coefficient de corrélation de Pearson de 0,85 (pvalue  $8,37e^{-167}$ ). Cependant, LINC01117 ayant une expression très faible, nous n'avons pas poursuivi l'étude de ce transcrit et nous nous sommes focalisés sur la fonction de LINC01116. L'exemple de lncARNs divergents composés d'un lncARN fortement exprimé et fonctionnel s'opposant à un lncARN faiblement exprimé correspondant à de la transcription pervasive a déjà été reporté dans la littérature <sup>172</sup>.

## Caractérisation fonctionnelle de LINC01116 dans les cellules tumorales et dans les cellules endothéliales

Concomitamment à l'initiation de ce projet, des études concernant la fonction de LINC01116 ont été publiées dans diverses revues scientifiques. En 2017, une équipe a renommé LINC01116, TALNEC2 pour « Tumor Associated LncRNA Expressed in Chromosome 2 ». Dans cette étude, les auteurs ont montré que l'expression de ce transcrit est forte chez les patients souffrant de glioblastomes et est associée à un mauvais pronostic. Ils ont également montré l'implication de TALNEC2 dans la régulation de la prolifération, la migration et l'invasion des cellules de glioblastomes *in vitro* et dans la croissance tumorale des glioblastomes *in vivo* <sup>433</sup>. Des études plus récentes ont mis en évidence son expression dans divers cancers agressifs (cancer du sein, ostéosarcome, cancer gastrique, gliome et CBNPC) <sup>432,434-438</sup>. Ces études ont montré que des siARNs ciblant LINC01116 affectent d'une part la prolifération des ostéosarcomes, des gliomes et des cancers du seins <sup>432,434,435,437</sup> et d'autre part la migration des ostéosarcomes, des cancers gastriques et des cancers de la tête et du cou en mettant en jeu divers mécanismes et des associations avec des partenaires différents <sup>434-436,438</sup>. Cette hétérogénéité des effets observés est à rapprocher de la variété d'isoformes de ces transcrits précédemment mentionnée, d'autant plus que de nombreuses études ne caractérisent pas précisément les variants exprimés dans leurs modèles. Ces travaux montrent que la caractérisation fonctionnelle et mécanistique de ce transcrit n'est pas encore clairement établie. Ainsi, dans nos modèles, l'inhibition transitoire de LINC01116 dans les cellules d'ADC a révélé qu'il n'était pas impliqué dans le contrôle de la prolifération ou de la migration des cellules cancéreuses mais plutôt dans les mécanismes d'adhésion cellulaire sur différentes protéines matricielles et dans les interactions cellule-cellule nécessaires à la formation de sphéroïdes.

L'adhésion cellulaire est particulièrement importante dans les mécanismes pro-angiogéniques des cellules endothéliales qui ont de forts niveaux de LINC01116 et dont l'inhibition diminue également l'adhérence. En effet, LINC01116 est exprimé dans les cellules cancéreuses mais plus encore dans les cellules endothéliales, ce qui pourrait contribuer à établir la corrélation existant entre son expression dans les tumeurs et la survie des patients. De plus, l'hypoxie stimule l'angiogenèse tumorale, associée aux cancers agressifs, et stimule également l'expression de LINC01116 dans les cellules endothéliales.

Nos expériences d'inhibitions transitoires démontrent la fonction de LINC01116 dans les mécanismes d'adhésion cellulaire que ce soit dans les cellules cancéreuses ou dans les cellules endothéliales. D'ailleurs, l'utilisation des gapmers est un outil d'intérêt puisqu'ils sont

actuellement utilisés en tant qu'outils de thérapie génique et ont déjà été utilisés dans l'invalidation de lncARNs <sup>439-443</sup>.

Le défaut d'adhésion des cellules dans lesquelles LINC01116 est inhibé transitoirement a été observé sur différentes protéines matricielles (fibronectine, collagènes de type I et IV) entraînant l'activation de différentes intégrines. Ces données suggèrent que les mécanismes régulés par ce lncARN sont probablement plus impliqués dans l'intégration ou la transduction des signaux régulant l'adhésion cellulaire que dans la liaison des ligands à la surface des cellules.

D'un point de vue méthodologique, la répression transitoire de l'expression d'un lncARN peut être un facteur limitant pour mettre en évidence les conséquences à long terme de cette inhibition. Par ailleurs, les siARNs et les antisens gapmers, fréquemment utilisés dans les études fonctionnelles des lncARNs, se caractérisent par d'importants effets non spécifiques dits effets "off-target". Afin de pallier à ces limitations techniques, nous avons développé une approche d'invalidation stable par CRISPR Interférence (CRISPRi) en une seule étape de transduction virale. Dans les cellules tumorales A549, la répression stable de LINC01116 confirme les effets de l'inhibition transitoire sur l'adhésion cellulaire. Cet outil est particulièrement intéressant pour réaliser des études fonctionnelles sur des cellules primaires cultivées sur un nombre de générations plus limité. Toutefois, dans ces conditions, l'expression de LINC01116 n'a été inhibée que de 40% dans les cellules HUVEC, ce qui n'a pas été suffisant pour reproduire les résultats obtenus par les méthodes d'inhibition transitoire inhibant le transcrite à 80%. Dans ces cellules la faible inhibition du lncARN pourrait être liée à des interactions avec des partenaires protéiques qui stabiliseraient le transcrite, ou à la fixation de nombreux facteurs de transcription sur son promoteur qui limiteraient l'accès des guides ARNs ou à un effet délétère à moyen/long terme sur les cellules HUVEC fortement réprimées pour LINC01116.

Nous avons également utilisé une approche de gain de fonction par surexpression stable du transcrite par transduction lentivirale. Dans les cellules d'ADC, cette surexpression n'impacte pas la prolifération, la migration ou l'adhésion des cellules. Nous avons confirmé la localisation cytosolique du transcrite transduit. Toutefois, nous ne pouvons pas savoir s'il est engagé dans les mêmes structures et interactions. En effet, la fonctionnalité du transcrite est potentiellement limitée par l'abondance de ses partenaires moléculaires et notamment par sa structure secondaire qui peut être verrouillée par des associations avec des partenaires protéiques. En revanche, l'expression de LINC01116 étant déjà conséquente dans les HUVEC, nous n'avons pas réussi à augmenter son niveau dans ces cellules.

L'étude fonctionnelle de LINC01116 *in vivo* nécessiterait de travailler sur des modèles murins, mais la conservation inter-espèces des lncARNs est limitée <sup>163</sup>. Cependant, des recherches dans la base de données ENCODE nous ont permis de trouver un locus murin comprenant un lncARN appelé Gm13652. Ce transcrit de 3000 nt possède la même structure intron-exon que son potentiel orthologue humain, LINC01116. Malheureusement, diverses tentatives pour détecter l'expression de cet orthologue murin à partir de données publiques de séquençage d'ARN ou générées au laboratoire à partir de cellules endothéliales murines, de cellules de carcinomes pulmonaires murins de Lewis (LLC <sup>444</sup>) et d'un modèle d'ADC murin induit par la forme mutée de *Kras* se sont avérées infructueuses.

Nous avons également essayé de cloner, sans succès, l'ADNc de *Gm13652* à l'aide d'amorces spécifiques à partir de cellules de LLC. Les annotations des lncARNs chez la souris restent encore largement incomplètes et les données publiques de séquençage ARN manquent généralement de profondeur. Toutefois, leur actualisation régulière par le consortium GENCODE pourrait, à l'avenir, permettre d'identifier des modèles d'études pour cloner ce potentiel orthologue murin de LINC01116 <sup>144</sup>.

Concernant la fonction de ce lncARN dans les cellules d'ADC, nous envisageons de réaliser des xénogreffes chez la souris Nude, de cellules tumorales, humaines inhibées stablement pour LINC01116. Pour étudier la composante vasculaire et nous rapprocher du contexte tumoral, nous envisageons de former des sphéroïdes à partir de cellules cancéreuses et de cellules endothéliales inhibées pour le transcrit LINC01116. Bien qu'imparfaite, cette technique permet de recréer des réseaux tubulaires au sein des sphéroïdes mimant partiellement la néo-angiogenèse tumorale <sup>445</sup>.

### **Mécanismes moléculaires impactés par l'inhibition du transcrit LINC01116**

Afin de comprendre les mécanismes d'actions de LINC01116, nous avons recherché les gènes et les voies de signalisations impactés par son inhibition. Nous avons analysé le transcriptome de cellules endothéliales HMEC1 et de cellules cancéreuses A549 inhibées ou non pour le transcrit LINC01116. Dans les cellules HMEC1, nous n'avons identifié qu'une centaine de gènes communément altérés par les deux siARNs ciblant LINC01116 par comparaison à un siARN contrôle, ce qui représente seulement 10% de la totalité des gènes modulés par ces ARNs interférents. Ces différences peuvent être liées à des effets non spécifiques ("off-target") des si-ARNs ciblant LINC01116 <sup>446,447</sup>. Cependant, les effets sur l'adhésion cellulaire ont été confirmés par l'utilisation d'oligonucléotides antisens de type gapmer et par CRISPRi.

L'analyse des gènes communément modulés par les deux siARNs a mis en évidence quelques réseaux géniques impactés par l'inhibition de LINC01116. C'est le cas d'un ensemble de gènes contrôlé par le facteur de transcription NKX2-3 qui est réprimé par l'inhibition de LINC01116 <sup>448</sup>. NKX2-3 est impliqué dans les mécanismes d'adhésion des cellules sur leur support et dans l'angiogenèse. Il induit l'expression de certains gènes tels que l'IL8 (*CXCL8*) et l'angiopoïétine 2 (*ANGPT2*) qui sont réprimés par l'inhibition de LINC01116 dans les cellules endothéliales. Un second réseau génique est également réprimé par l'inhibition de LINC01116 et concerne les gènes modulés par le facteur de transcription PDX1. PDX1 est exprimé dans les cellules de l'endoderme et participe au développement du pancréas <sup>449</sup>. Plus précisément, il est impliqué dans la formation de l'épithélium et des communications intercellulaires en modulant l'expression de la E-cadherine <sup>450</sup>. De plus, il est également induit dans les cancers pancréatiques et gastriques <sup>451,452</sup>. Ainsi bien que ce facteur de transcription soit peu documenté dans nos modèles d'études, son expression dans d'autres types de cancers et sa fonction dans la formation de jonctions intercellulaires pourrait expliquer sa dérégulation lors de l'inhibition de LINC01116. Ces réseaux de gènes sont également modulés dans les cellules A549 inhibées pour LINC01116 (Figure S11DE).

Cependant, il est nécessaire de constater que ces profilages de l'expression génique ont donné peu de renseignements sur des cibles moléculaires potentiellement impliquées dans un défaut d'adhésion. Une des explications possible réside dans l'état dans lequel se trouvent les cellules que nous avons étudiées. En effet, ces expériences ont été réalisées sur des cellules endothéliales HMEC1 ou tumorales A549, 48h après transfection par des siARNs et 72h après leur ensemencement sur des boîtes de culture, c'est à dire dans des conditions où les cellules ont déjà établi des contacts avec leur support et avec les cellules voisines bien avant que l'inhibition de LINC01116 n'ait lieu. A l'avenir, nous souhaitons réaliser un profilage transcriptomique sur des cellules, réprimées ou non pour LINC01116, puis engagées dans un processus dynamique d'adhésion ou de formation d'interactions cellule-cellule. Nous pensons qu'analyser l'expression des gènes dans ces conditions sera beaucoup plus informatif pour comprendre les mécanismes moléculaires altérés par l'absence de LINC01116 et affectant ces processus cellulaires. Une preuve de concept de cette hypothèse de travail a été apportée par l'analyse ciblée, par RT-qPCR, de l'expression de gènes de la voie NOTCH1 dans les cellules HUVEC 4h après adhésion sur de la fibronectine. Cette analyse a mis en évidence la répression des gènes HEY1, HEY2 et DII4 dans les cellules inhibées pour LINC01116 (Figure S10AC). Ces altérations n'avaient pas été indentifiées par nos analyses transcriptomiques sur cellules statiques.

Dans cette optique, nous avons également souhaité analyser le profil d'expression génique de sphéroïdes présentant des défauts de compaction (Figure S6). Malheureusement, malgré de nombreuses tentatives, nous n'avons pas réussi à extraire des ARNs de qualité suffisante pour analyser leur transcriptome.

Par ailleurs, l'adhésion cellulaire sur différentes protéines matricielles engage différentes intégrines et conduit à l'induction de différents gènes cibles de la voie NOTCH. Ainsi, l'adhésion sur laminine 111, active les intégrines  $\alpha6\beta1$  conduisant à une forte induction de l'expression de DII4<sup>428</sup>. De plus, la voie Notch est à la fois régulée par l'adhésion cellulaire et impliquée dans la signalisation induite par les interactions cellule-cellule<sup>428-430</sup>. Il serait donc intéressant d'observer l'impact de l'inhibition de LINC01116 sur l'activation de la voie Notch lors d'adhésion sur différentes protéines matricielles et dans des essais de tubulogénèse.

### **Caractérisation du mode d'action de LINC01116 par l'identification de ses interactants moléculaires**

Si l'on exclut les transcrits antisens, les pseudogènes qui servent de leurre pour la fixation de microRNAs et les lincRNAs ayant des activités "enhancer", la plupart des lincRNAs contrôlent l'expression génique par divers mécanismes transcriptionnels, post-transcriptionnels ou post-traductionnels, qui dépendent notamment de leur localisation subcellulaire (nucléaire, cytosolique, granules...) <sup>453</sup> et qui impliquent le recrutement de ribonucléoprotéines (RNP ou RBP RNA-binding proteins) régulatrices et/ou l'association avec des séquences génomiques ou d'autres transcrits <sup>204</sup>. Les RBPs s'associant avec un lincARN peuvent soit agir sur le métabolisme du transcrit (épissage, stabilité...) soit participer à son mode d'action.

Dans ce contexte, l'identification des partenaires protéiques et des gènes cibles (ADN ou ARN) des lincARNs, par la combinaison d'approches biochimiques et d'analyses protéomiques et/ou de séquençage à haut-débit, constitue une problématique centrale dans le domaine des lincARNs.

Ainsi, la localisation subcellulaire cytosolique et en partie membranaire de LINC01116 (Figures S2AB et S5B) suggère que ce transcrit pourrait réguler l'expression génique par des mécanismes post-transcriptionnels tels que le contrôle de la stabilité des ARNm ou par des mécanismes post-traductionnels en servant de plateforme d'ancrage pour le contrôle de l'activation ou de la stabilité de protéines. C'est le cas du linc-DC qui maintient STAT3 sous une forme active phosphorylée en bloquant sa déphosphorylation par SHP1<sup>256</sup>, et de LINK-A qui fixe les kinases BRK et LRRK2 pour activer HIF1 $\alpha$  dans les cancers mammaires triple-négatifs<sup>341</sup>.

Afin d'identifier les protéines interagissant avec LINC01116, nous avons réalisés des expériences de pull-down d'ARN à l'aide d'un lncARN transcrit *in vitro* et biotinylé à son extrémité 3' (Figure 39AB). Ces expériences nous ont permis de mettre en évidence différentes protéines partenaires dont les protéines ILF3 (Inducible Interleukin Factor 3) et PABPC1. Ces interactions ont été validées dans les cellules d'ADC A549 et dans les cellules endothéliales HUVECs et HMEC1.

La protéine ILF3 possède deux isoformes majoritaires, une isoforme exclusivement nucléaire de 110 kDa, et une isoforme de 90 kDa (NF-90) (Fig. 39C) qui permet le transport nucléo-cytosolique d'ARN en se fixant sur leurs structures tige-boucle et la formation de granules de stress cytosoliques<sup>191,258</sup>. L'interaction de NF-90 avec le transcrit LINC01116 a été validée, dans un second temps, par pull-down d'ARN puis révélation par Western Blot (Fig. 39D). Toutefois, ces expériences de pull-down sont réalisées en conditions acellulaires et ne sont pas forcément révélatrices des interactions endogènes. Nous avons donc corroboré ces résultats par des expériences d'immunoprécipitation d'ARN (RIP) sur des complexes natifs (Figure 39EF) et visualisé ces interactions par co-marquage des ARNs et des protéines par smRNA FISH couplé à de l'immunofluorescence (Figure 39G). NF-90 est une protéine candidate intéressante qui pourrait contrôler le métabolisme de LINC01116 en régulant son transport et/ou sa stabilité. Des études ont également démontré l'implication de NF-90 dans les processus pro-angiogéniques et les capacités d'adhésion des cellules endothéliales notamment en régulant la stabilité de l'ARNm du VEGFA<sup>454-456</sup>. Une étude récente suggère également l'implication de LINC01116 dans le contrôle de la stabilité du transcrit du VEGFA par un mécanisme d'éponge au microARN miR-31-5p<sup>432</sup>. Nous avons recherché la possible existence d'un complexe ribonucléoprotéique impliquant NF-90, LINC01116 et l'ARNm du VEGFA dans nos modèles. Nous avons bien validé la co-immunoprécipitation de LINC01116 et du transcrit VEGFA par un anticorps dirigé contre ILF3 (Figure 40A-C). Cependant, l'expérience inverse de RAP (RNA Antisens Purification, Figure 25), consistant à précipiter après réticulation covalente le transcrit LINC01116 à l'aide d'oligonucléotides antisens biotinylés n'a pas révélé de co-précipitation de l'ARNm du VEGFA (Figure 40D).

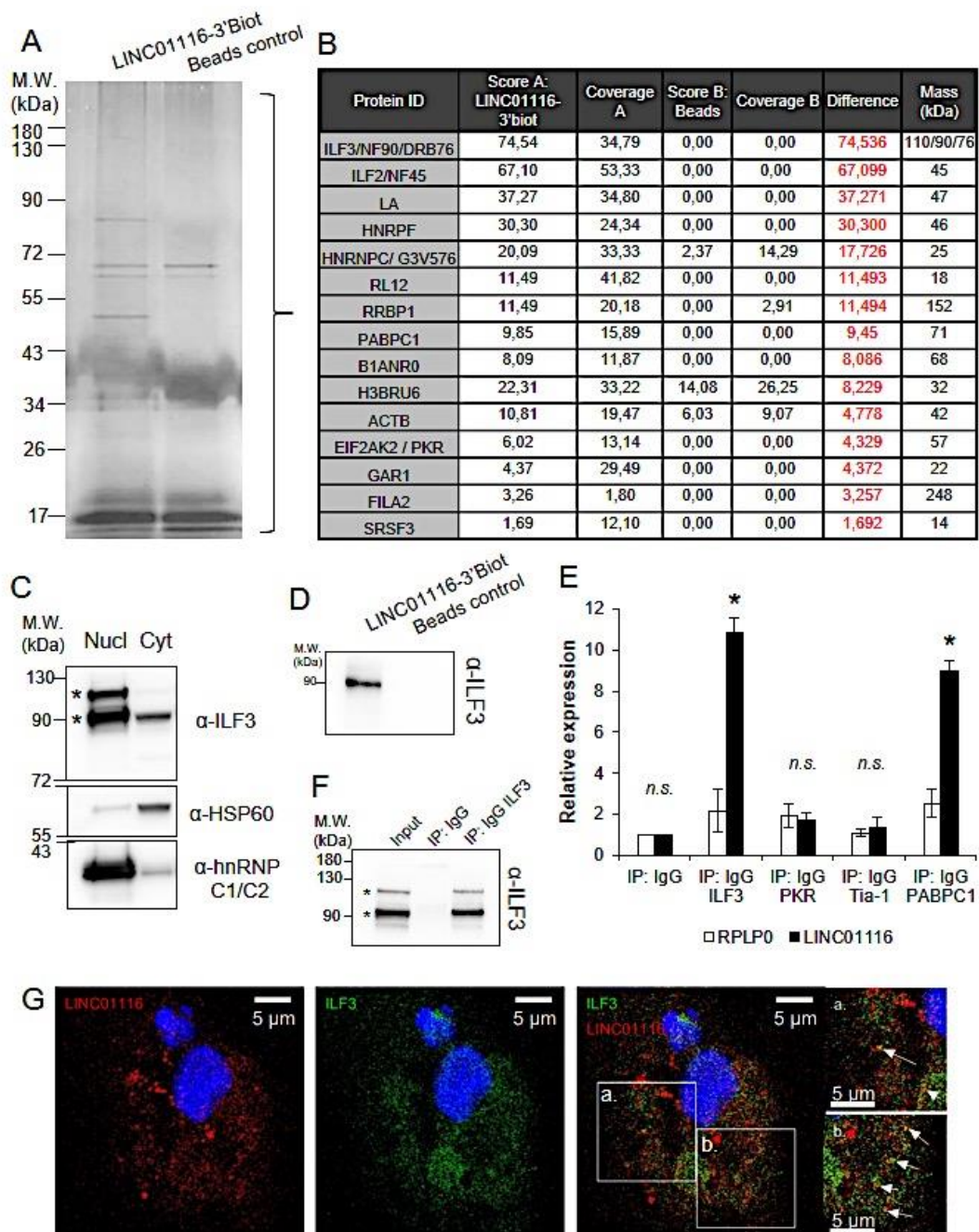
Ainsi, nos résultats n'ont pas permis de préciser le rôle de l'interaction entre la protéine ILF3/NF-90 et LINC01116. D'autres investigations seront nécessaires pour déterminer si cette association régule la stabilité de LINC01116 ou sa localisation subcellulaire, son action sur l'adhésion des cellules, ou si LINC01116 s'associe à d'autres transcrits.

La protéine PABPC1 est impliquée dans le contrôle de la stabilité des ARNs et dans la formation de granules d'ARN cytoplasmiques (p-bodies et granules de stress)<sup>263,266</sup>. Dans

certaines conditions, PABPC1 peut faire la navette entre le cytosol et le noyau <sup>457</sup>. De façon intéressante, plusieurs publications récentes lui attribuent diverses fonctions impliquant une relocalisation au niveau de la membrane plasmique <sup>458,459</sup>. Ainsi, PABPC1 interagit avec la métalloprotéase membranaire ADAM15 et des ARNm en cours de traduction lors de la ré-adhésion de chondrocytes ou de synoviocytes <sup>458</sup>. Par ailleurs, dans les cellules épithéliales, PABPC1 participe à un complexe macromoléculaire important pour l'intégrité des jonctions adhérentes intercellulaires et impliquant notamment la cadhérine E, le complexe RISC et des ARNm <sup>459</sup>. Dans ce contexte, on peut émettre l'hypothèse que PABPC1 pourrait être impliquée dans la localisation de LINC01116 à la membrane plasmique et participer à la médiation de sa fonction dans l'adhésion cellulaire et/ou dans les connections intercellulaires.

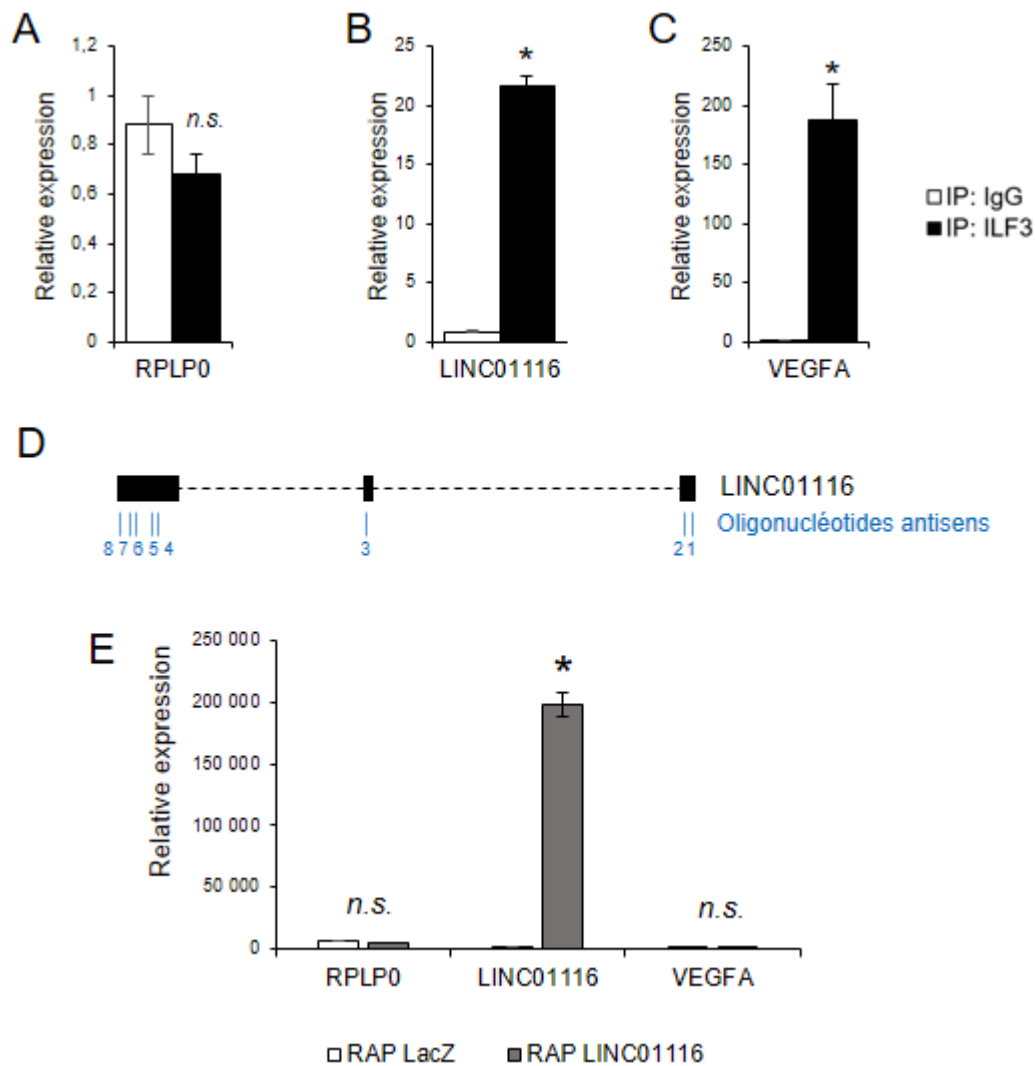
Des études fonctionnelles de ces protéines partenaires par des approches de perte ou de gain de fonctions permettront de préciser leur éventuelle implication dans la régulation du métabolisme de LINC01116 et dans les processus d'adhésion ou de connection intercellulaire dans lesquels il pourrait être impliqué.





**Figure 39:** Recherche d'interactants protéiques de LINC01116 dans les cellules HUVEC.

**A-B.** L'identification des partenaires protéiques de LINC01116 a été réalisée par des expériences de pulldown d'ARN. Du lysat cellulaire de cellules HUVEC a été incubé en présence ou en absence de LINC1116 transcrit in vitro et biotinylé en 3', puis les complexes ribonucléoprotéiques formés ont été précipités à l'aide de billes magnétiques couplées à de la streptavidine. Les protéines précipitées par le transcrit ont été séparées sur gel d'électrophorèse puis révélées par coloration à l'argent (**A**). L'identification des peptides tryptiques par analyses en spectrométrie de masse a permis d'établir une liste de protéines présentées sous forme de tableau (**B**). **C.** Analyse de la localisation subcellulaire d'ILF3 par fractionnement subcellulaire + Western blot. Des aliquots de fractions nucléaires (Nucl) et cytosoliques (Cyt) ont été séparés sur gel puis transférés sur membrane et révélés à l'aide de différents anticorps. La révélation des protéines HSP60 et hnRNP1/C2, ayant des localisations respectivement cytosolique et nucléaire, a permis de valider la qualité du fractionnement. **D.** Validation par pulldown ARN couplé à du Western Blot de la précipitation de l'isoforme cytosolique d'ILF3 par LINC01116. **E-F.** Expériences d'immunoprécipitation d'ARN. Les complexes ARN-protéines endogènes présents dans des extraits de cellules HUVEC ont été immunoprécipités à l'aide d'anticorps ciblant les protéines ILF3, PKR, Tia-1 et PABPC1. Les ARNs co-précipités ont été extraits et la présence des transcrits LINC01116 et RPLP0 a été révélée par RT-qPCR (**E**). La précipitation spécifique de la protéine ILF3 par son anticorps a été validée par Western Blot (**F**). Ces résultats représentent la moyenne de trois expériences indépendantes +/- SD, (\*pvalue <0,05). **G.** Co-localisation d'ILF3 et LINC01116 visualisée par smRNA-FISH couplé à de l'immunofluorescence.



**Figure 40:** Recherche d'interactions potentielles entre la protéine ILF3, LINC01116 et l'ARNm de VEGFA

**A-C.** Expériences d'Immunoprecipitation d'ARN (RIP) à l'aide d'anticorps anti-ILF3. La présence de LINC01116 et des ARNm de RPLP0 et VEGFA a été révélée par RT-qPCR. **DE. Recherche d'ARN interagissant avec LINC01116 par Purification d'ARN par antisens (RAP).** Des cellules HUVEC ont été fixées en formaldéhyde 4% puis lysées dans un tampon 1% NP-40. Les complexes macromoléculaires endogènes contenant LINC01116 ont été précipités à l'aide d'un mélange de 8 oligonucléotides antisens biotinylés dirigés contre LINC01116 (**D**). Afin d'évaluer le bruit de fond expérimental, nous avons utilisé un lot de sondes complémentaires du transcrit LacZ absent des cellules eucaryotes. Les ARN ont été extraits à partir des complexes macromoléculaires précipités et la présence de LINC01116 et des ARNm de RPLP0 et VEGFA a été analysée par RT-qPCR (**E**). Les résultats représentés sont la moyenne de deux expériences indépendantes, +/- SD, (\*p-value <0,05).

## **V. CONCLUSION**

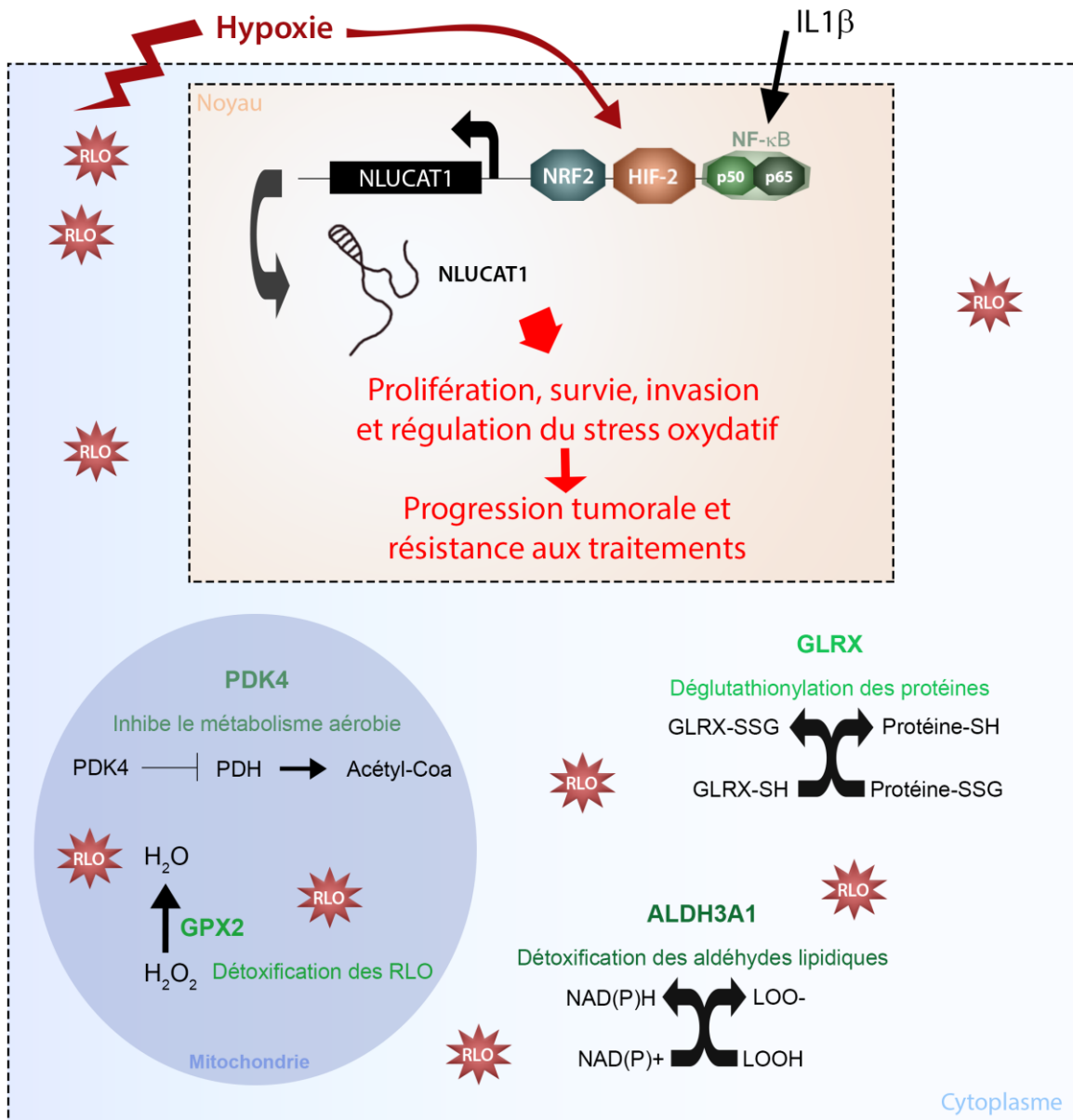
## **Conclusion**

Mes travaux de thèse ont porté sur l'étude de deux lncARNs, NLUCAT1 et LINC01116, induits par l'hypoxie et corrélés à un mauvais pronostic des patients ayant un adénocarcinome broncho-pulmonaire.

L'étude du transcrit NLUCAT1 a été réalisée conjointement avec le Dr Laura Moreno Leon. J'ai poursuivi ce travail en précisant sa structure exon-intron, sa régulation par l'hypoxie impliquant les facteurs de transcription HIF2, NF- $\kappa$ B et NRF2, et sa localisation nucléaire notamment en développant l'approche de smRNA-FISH. Concernant la caractérisation fonctionnelle du transcrit, j'ai confirmé, dans les cellules invalidées par CRISPR/Cas9 pour le transcrit NLUCAT1, les défauts de prolifération au sein de sphéroïdes, la sensibilité aux agents causant un stress oxydatif (cisplatine, ménadione, peroxyde d'hydrogène) et la diminution des propriétés migratoires et invasives. Mes travaux ont également consisté à démontrer la non implication du variant court cytosolique SCAL1 dans la médiation des effets de NLUCAT1. Enfin, j'ai pu valider les cibles moléculaires de NLUCAT1 impliquées dans le contrôle du stress oxydatif : les gènes *ALDH3A1*, *GPX2*, *GLRX* et *PDK4*. Ainsi, cette étude a mis en évidence l'implication de NLUCAT1 dans différents processus pro-tumoraux et notamment dans la résistance au cisplatine (Figure 41).

Par ailleurs, j'ai abordé l'étude du mode d'action de NLUCAT1 par l'identification de ses interactants nucléiques et protéiques en développant diverses approches biochimiques de précipitations couplées à des analyses à haut débit par séquençage ou par spectrométrie de masse. La mise en place de ces méthodologies de ChIRP, de RIP, de RAP et de pull-down d'ARN, qui a été chronophage, a permis de lever des verrous technologiques indispensables pour aborder cette problématique centrale dans le domaine des lncARNs.

Actuellement, l'étude du mécanisme d'action de NLUCAT1 se poursuit par la validation des interactions identifiées, la caractérisation des sites d'interactions entre le transcrit et ses partenaires et par l'étude des protéines partenaires par des approches de pertes/gains de fonction.



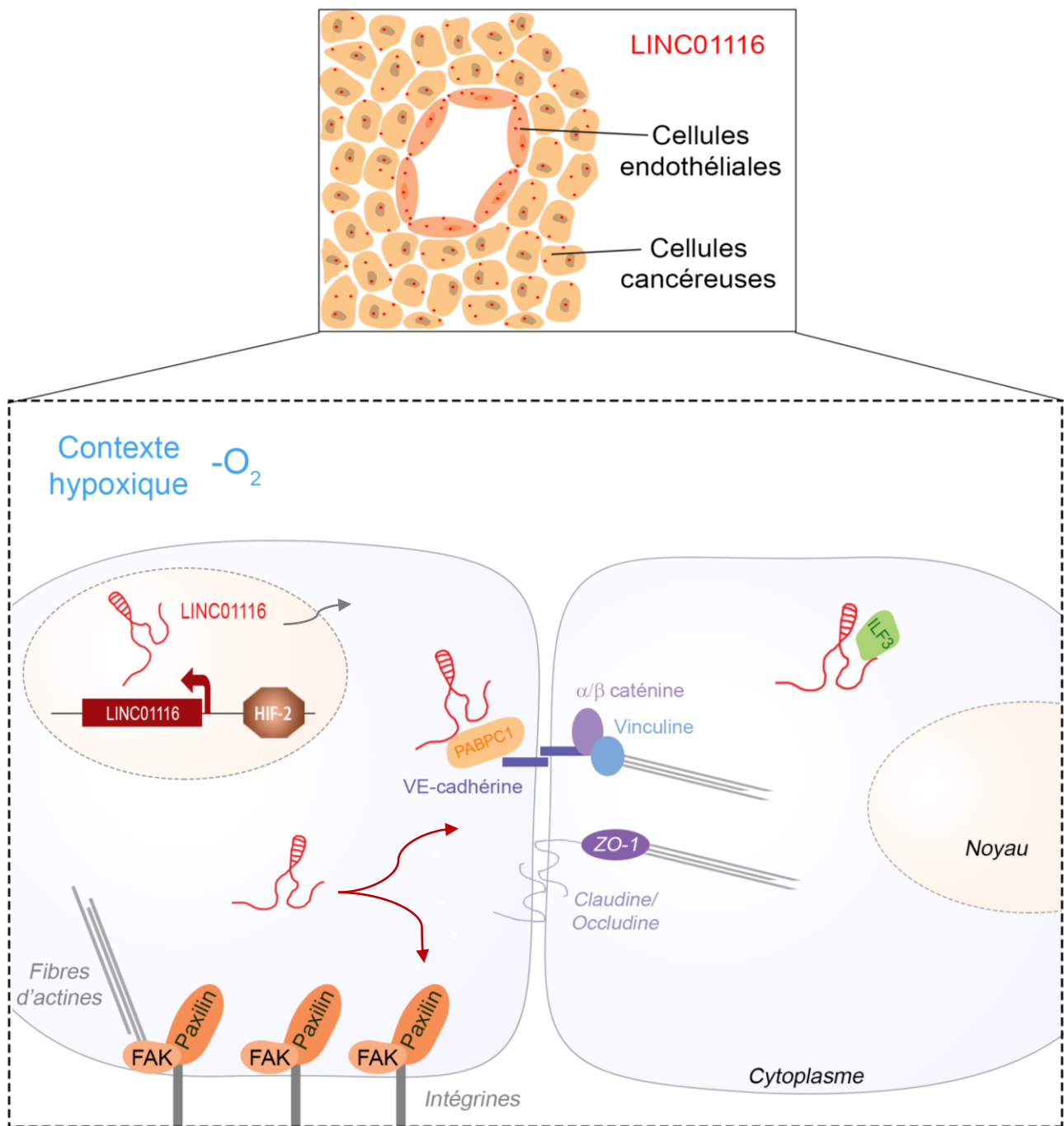
**Figure 41:** NLUCAT1, un lncARN nucléaire protumoral qui régule la réponse anti-oxydante au sein des CBNPC.

L'hypoxie et les radicaux libres oxygénés ( $O_2^-$ ,  $H_2O_2$ ,  $HO^+$ ) participent à l'activation des facteurs de transcription HIFs, NF- $\kappa$ B et NRF2 capables d'induire la transcription de NLUCAT1. NLUCAT1 pourrait exercer un rétrocontrôle positif sur les voies régulées par ces facteurs (notamment la réponse anti-oxydante) afin de favoriser la survie, la prolifération et l'invasion cellulaire. Le stress oxydatif, notamment par la régulation des radicaux libres, est pris en charge par différentes enzymes spécialisées tel que les enzymes GPX2, GLRX, ALDH3A1 et PDK4 dont l'expression est réprimée en absence de NLUCAT1.

Mon second projet a porté sur le transcrit LINC01116 induit par l'hypoxie dans les cellules d'ADC *in vitro* mais également corrélé au statut hypoxique des tumeurs et à une diminution de la survie des patients. J'ai montré que ce transcrit cytosolique est exprimé dans les cellules tumorales mais également dans les cellules endothéliales du stroma tumoral. La répression de l'expression de ce lncARN à l'aide de divers outils (siARNs, gapmers, CRISPRi), m'a permis de mettre en évidence son implication dans les mécanismes d'adhésion et de communications intercellulaires des cellules cancéreuses et des cellules endothéliales (Figure 42).

Concernant les mécanismes moléculaires impliqués dans la médiation des effets de LINC01116, j'ai montré que son inhibition altère l'expression de gènes cibles de la voie Notch. D'autres profilages transcriptomiques seront nécessaires pour mieux caractériser les mécanismes moléculaires régulés par LINC01116. Par ailleurs, en utilisant les méthodologies que j'ai mis en place, j'ai identifié deux protéines s'associant avec LINC01116 : i) la protéine ILF3 impliquée dans le transport des ARNs et dans la régulation post-transcriptionnelle de leur stabilité dans les p-bodies, et ii) la protéine PABPC1 jouant également un rôle dans la stabilisation des ARNs et leur transport jusqu'à la membrane plasmique. Là encore, des expériences complémentaires devront être réalisées afin d'élucider les mécanismes régulés par ces complexes ribonucléoprotéiques.

A terme, ces expériences fonctionnelles et l'implication pro-tumorale de ces molécules pourront permettre d'envisager ces transcrits comme des marqueurs pronostics et/ou de nouvelles cibles thérapeutiques. Ils pourront également être ciblés par des oligonucléotides (Gapmer) ou des vecteurs d'expression CRISPR/Cas9 en tant que traitements adjuvants des cancers broncho-pulmonaires.



**Figure 42:** LINC01116, un hypoxa-lnc cytosolique impliqué dans l'adhésion et les jonctions intercellulaires.

LINC01116 est induit par le facteur de transcription HIF2 en hypoxie dans les cellules cancéreuses et les cellules endothéliales. Après exportation dans le cytosol, LINC01116 impacte l'adhésion des cellules en modulant la signalisation FAK/Paxillin et les jonctions intercellulaires. LINC01116 s'associe notamment aux protéines ILF3 et PABPC1 dans le cytosol des cellules, certainement pour médier sa fonction par un mécanisme encore inconnu.

## **VI. ANNEXE**



# **One step CRISPRi efficiently represses gene expression in mammalian cells**

**Caroline Lacoux <sup>1</sup>, Marine Gautier-Isola <sup>1</sup>, Julien Fassy <sup>1</sup>, Roger Rezzonico <sup>2</sup>, Bernard Mari <sup>1</sup>, and Georges Vassaux <sup>2\*</sup>**

<sup>1</sup> Université Côte d'Azur, CNRS, IPMC, FHU-OncoAge, Valbonne, France.

<sup>2</sup> Université Côte d'Azur, INSERM, CNRS, IPMC, Valbonne, France.

\*Corresponding author : [vassaux@ipmc.cnrs.fr](mailto:vassaux@ipmc.cnrs.fr)

Keywords : One-step CRISPRi, two-step CRISPRi, transcriptional repression, loss of function.

Running head: Comparison between one- and two- step CRISPRi

## **ABSTRACT**

CRISPRi exploits a deactivated Cas9 (dCas9) fused to a Krüppel-associated box (KRAB) repression domain. CRISPRi is a technique allowing loss of function studies. It is particularly adapted to the study of long-non-coding RNAs. In cells, a RNA guide complementary to relevant sequences in the targeted locus directs the KRAB-dCas9 protein to repress transcription. This is usually a two-step process involving two consecutive lentiviral infection, the first to generate stably expressing KRAB-dCas9 cells and the second to express the RNA guide. A one-step approach has also been proposed in which two distinct expression cassettes encoding the RNA guide and KRAB-dCas9 are included. The whole cistron is packaged into a lentivirus and CRISPRi-mediated gene repression can be achieved in a single step. We evaluated the efficacy of these two strategies using currently available tools and we demonstrate in mammalian cells that they are equally efficient. We propose an experimental workflow to implement this one-step strategy that can be used as a platform to apply CRISPRi to primary cells or to cells with a limited life-span.

Loss of function (LOF) approaches are key to elucidate the role of specific genes in a biological process. In eukaryotic cells, LOF is mainly achieved through three distinct methods: RNA interference (RNAi), antisense oligonucleotides (ASO) and a modification of the CRISPR/Cas9 system referred to as CRISPR interference (CRISPRi) (1). These three methods can efficiently reduce the expression of targeted genes. However, biases have been reported, the most important being off-target effects, in which the level of expression of non-targeted genes is affected independently of the specific effect on the targeted gene. The existence of off-target effects in LOF has been widely documented (1-5) and a recent report has demonstrated convincingly that the CRISPRi methodology is associated with a lower level of off-target effects and is particularly adapted to study Long non-coding RNA (LncRNA) (1).

CRISPRi is based on the utilization of a deactivated Cas9 (dCas9) fused to a Krüppel-associated box (KRAB) repression domain. In cells, a single guide RNA (sgRNA) complementary to sequences present at the targeted locus directs the KRAB-dCas9 protein to transcriptomic regulatory elements/promoter regions, preventing or reducing the initiation of the transcription of the targeted gene (1, 6). In most studies, CRISPRi is a two-step process. Lentiviral vectors are used to generate stably expressing KRAB-dCas9 cells, which are selected on the basis of their level of KRAB-dCas9 expression (either in bulk or as individual clones). The LOF is achieved by infecting these cells with a lentivirus encoding the sgRNA selective to the targeted gene. This two-step strategy is very efficient and the most important bias in the methodology lies in the utilization of clones generated from single-cells expressing KRAB-dCas9 (1). This suggests that polyclonal populations of KRAB-dCas9-expressing cells should be preferred for CRISPRi experiments (1).

A one-step approach (OSA-CRISPRi) has also been proposed, in which an expression cassette encoding the sgRNA is included in the plasmid encoding the KRAB-dCas9 expression cassette. The whole cistron is packaged into a lentivirus, and CRISPRi-mediated

gene repression is achieved in a single step. In addition to being less time-consuming, OSA-CRISPRi offers the possibility of applying CRISPRi to primary cells or to cells with a limited life span and/or reduced culture passage potential. Purified lentiviruses may also be used *in vivo* for LOF studies in live animals. OSA-CRISPRi has been used to repress the expression of target genes in a high-throughput screen (7) but, to our knowledge, a direct comparison between OSA-CRISPRi and the two-step approach (TSA-CRISPRi) has never been reported. A crucial aspect in the relative efficacies of OSA-CRISPRi and TSA-CRISPRi is the fact that, for the latter, the relative level of sgRNA to KRAB-dCas9 can be adjusted to maximize target gene inhibition of expression, while in OSA-CRISPRi this ratio is fixed by the relative strengths of the different promoters driving expression of the guide RNA and KRAB-dCas9. In the present paper, we compare directly the efficacies of OSA-CRISPRi and TSA-CRISPRi in mammalian cells expressing the nuclear LncRNA NLUCAT1 (8).

We first cloned guides RNA directed towards NLUCAT1 using the Gibson assembly method (NEBuilder HiFi DNA Assembly Master Mix, NEB # E2621). The plasmids CROPseq-Guide-Puro (Addgene plasmid # 86708) and pLenti-(BB)-EF1a-KRAB-dCas9-P2A-EGFP (Addgene plasmid # 60954), were digested using the restriction enzyme BsmBI (NEB # R0580) for 2h at 50°C. The relevant fragments (around 8 kB) were gel-purified using the Qiagen Gel purification kit and stored at – 20°C in 20-fmol aliquots. Aliquoted, BsmBI-digested plasmid was mixed with 0.55 µL guide oligonucleotide (200nM, see Table below) in 10µl total volume, combined with 10µl 2X NEBuilder HiFi Assembling Master mix and the mixture was incubated at 50°C for 20 minutes. 8µL of NEBuilder Assembling mixture were incubated with 100 µL of Stabl2 competent *E coli*. The mixture was heat-shocked at 42°C for 45 seconds and transferred to ice for 2 minutes. SOC medium (900 µl) was added to the Stabl2-NEBuilder mixture and the mix was incubated at 37°C for 1 hour. Transformed bacterial cells (350µl) were plated onto LB agarose plates containing ampicillin (100µg/mL) and incubated overnight at 37°C. Individual colonies were picked and grown overnight in 5 mL of Terrific Broth medium containing 150µg/mL ampicillin and low-endotoxin, small scale preparation of

plasmid DNA were performed using the ToxOut EndoFree Plasmid Mini Kit from BioVision (#K1326-250).

Lentiviruses were produced using a standard Lipofectamine 2000™ transfection protocol, using one million HEK293 cells seeded in a T25 flask (25 cm<sup>2</sup>) in DMEM medium supplemented with 10% bovine serum. A mixture of four plasmids (3 µg pMDLg/pRRE, 1.4 µg pRSV-Rev, 2 µg pVSV-G and 2.5 µg of the plasmid containing the expression cassette to package) was transfected. Forty-eight hours later, the medium was collected, centrifuged for 5 minutes at 3000 rpm, and 2.5 mL supernatant containing the viral particles was collected and used to infect cells or aliquoted and stored at -80°C.

In the TSA-CRISPRi protocol, we first infected A549 cells with a lentivirus produced from the plasmid pHR SFFV KRAB-dCas9-P2A-mCherry (addgene #60954) allowing the expression of a fusion protein KRAB-dCas9-mCherry. Cells were then sorted by preparative FACS and the top 5% brightest cells were taken. We selected sgRNAs against the lncRNA NLUCAT1 as well as a non-relevant sgRNA and we cloned into the CROPseq-Guide-Puro plasmid. For OSA-CRISPRi, functional sgRNAs were cloned into the pLenti-(BB)-EF1a-KRAB-dCas9-P2A-EGFP plasmid. Infection of cells with the CROPseq-Guide-Puro plasmid- or with the pLenti-(BB)-EF1a-KRAB-dCas9-P2A-EGFP plasmid-derived lentivirus marked the starting point of the experiment.

The comparison between OSA-CRISPRi and TSA-CRISPRi was performed by inactivating NLUCAT1 in lung adenocarcinoma A549 cells. After implementing the CRISPRi protocols using two independent validated sgRNAs (**Table 1**), cells were collected, RNA extracted and quantitative RT-qPCR was performed on total RNA (as previously described (8)) collected 4 and 11 days following lentivirus infection (**Figure 1A and 1B**). The list of primers used is presented in **Table 2**. Both OSA-CRISPRi and TSA-CRISPRi induced a marked and stable reduction of NLUCAT1 expression at the two time-points. We then examined KRAB-dCas9 levels in the different conditions. Levels of KRAB-dCas9 were equivalent in all TSA-CRISPRi

conditions as they were generated from the same cellular bulk population and were thus used as reference. Levels of KRAB-dCas9 in the different OSA-CRISPRi conditions were systematically lower than those obtained in TSA-CRISPRi (**Figure 1C**).

Finally, we determined the level of expression of the different sgRNAs at two time points (**Figure 2A and B**). The guides were expressed at a level roughly three orders and one order of magnitude lower than the reporter gene (RPLP0) for OSA-CRISPRi and TSA-CRISPRi, respectively. The expression of the sgRNAs was stable for at least one week. This large difference in guide expression is due to the design of the expression cassette of the sgRNA in the TSA-CRISPRi. Three sgRNAs (two functional and one serving as a tag to identify the guide in the CROP-seq approach (9)) are generated (**Figure 2C**). Considering that the functional sgRNA in both approaches is driven by the same U6 promoter (**Figure 2C and D**), it is tempting to assume that functional guides are expressed in the same range level in the two approaches and at a level well-below that of the RNA-polymerase II-driven “tag” transcript. Altogether, our data demonstrate that the inhibition of target gene expression achieved using OSA-CRISPRi is at least as efficient as that achieved with TSA-CRISPRi, even in conditions in which KRAB-dCas9 (**Figure 1C**) and sgRNAs expression levels (**Figure 2C and D**) are not fully optimized.

To validate the one-step CRISPRi approach in cells with a limited life-span and/or reduced culture passage potential, we subjected human microvascular endothelial cells (HMEC1) (10) to OSA-CRISPRi against NLUCAT1. **Figure 3** shows that the level of NLUCAT1 expression was reduced by around 80% with both sgRNAs. In conclusion, we propose that OSA-CRISPRi represents an invaluable tool for loss-of-function experiments, particularly those involving lncRNA, in fields as diverse as tissue engineering and programming of primary cells.

## **Acknowledgement**

We would like to acknowledge the technical support of the MICA imaging facility of the University Côte d'Azur, and to thank Dr D Gillot (Rennes, France) for help in lentivirus production and Franck Aguila (IPMC, Valbonne) for artwork. The plasmids pHR-SFFV-KRAB-dCas9-P2A-mCherry, CROPseq-Guide-Puro and pLenti-(BB)-EF1a-KRAB-dCas9-P2A-EGFP were provided by J Weissman, C Bock and J Ferrer, respectively. This work was supported by Canceropole PACA and Plan Cancer 2018 « ARN non-codants en cancérologie: du fondamental au translationnel » (number 18CN045).

## REFERENCES

1. Stojic, L., Lun, A. T. L., Mangei, J., Mascalchi, P., Quarantotti, V., Barr, A. R., Bakal, C., Marioni, J. C., Gergely, F., and Odom, D. T. (2018) Specificity of RNAi, LNA and CRISPRi as loss-of-function methods in transcriptional analysis. *Nucleic Acids Res* 46, 5950-5966
2. Jackson, A. L., Bartz, S. R., Schelter, J., Kobayashi, S. V., Burchard, J., Mao, M., Li, B., Cavet, G., and Linsley, P. S. (2003) Expression profiling reveals off-target gene regulation by RNAi. *Nature biotechnology* 21, 635-637
3. Kamola, P. J., Kitson, J. D., Turner, G., Maratou, K., Eriksson, S., Panjwani, A., Warnock, L. C., Douillard Guilloux, G. A., Moores, K., Koppe, E. L., Wixted, W. E., Wilson, P. A., Gooderham, N. J., Gant, T. W., Clark, K. L., Hughes, S. A., Edbrooke, M. R., and Parry, J. D. (2015) In silico and in vitro evaluation of exonic and intronic off-target effects form a critical element of therapeutic ASO gapmer optimization. *Nucleic Acids Res* 43, 8638-8650
4. Kuscu, C., Arslan, S., Singh, R., Thorpe, J., and Adli, M. (2014) Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nature biotechnology* 32, 677-683
5. Thakore, P. I., D'Ippolito, A. M., Song, L., Safi, A., Shivakumar, N. K., Kabadi, A. M., Reddy, T. E., Crawford, G. E., and Gersbach, C. A. (2015) Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat Methods* 12, 1143-1149
6. Kampmann, M. (2018) CRISPRi and CRISPRa Screens in Mammalian Cells for Precision Biology and Medicine. *ACS Chem Biol* 13, 406-416
7. Miguel-Escalada, I., Bonas-Guarch, S., Cebola, I., Ponsa-Cobas, J., Mendieta-Esteban, J., Atla, G., Javierre, B. M., Rolando, D. M. Y., Farabella, I., Morgan, C. C., Garcia-Hurtado, J., Beucher, A., Moran, I., Pasquali, L., Ramos-Rodriguez, M., Appel, E. V. R., Linneberg, A., Gjesing, A. P., Witte, D. R., Pedersen, O., Grarup, N., Ravassard, P., Torrents, D., Mercader, J. M., Piemonti, L., Berney, T., de Koning, E. J. P., Kerr-Conte, J., Pattou, F., Fedko, I. O., Groop, L., Prokopenko, I., Hansen, T., Marti-Renom, M. A., Fraser, P., and Ferrer, J. (2019) Human pancreatic islet three-dimensional chromatin architecture provides insights into the genetics of type 2 diabetes. *Nat Genet* 51, 1137-1148
8. Moreno Leon, L., Gautier, M., Allan, R., Ilié, M., Nottet, N., Pons, N., Paquet, A., Truchi, M., Fassy, J., Magnone, V., Kinnebrew, G., Radovich, M., Cheok, M. H., Barbry, P., Vassaux, G., Marquette, C.-H., Ponzio, G., Ivan, M., Pottier, N., Hofman, P., Mari, B., and Rezzonico, R. (2019) The nuclear hypoxia-regulated NLUCAT1 long non-coding RNA contributes to an aggressive phenotype in lung adenocarcinoma through regulation of oxidative stress. *Oncogene* 38, 7146-7165
9. Datlinger, P., Rendeiro, A. F., Schmidl, C., Krausgruber, T., Traxler, P., Klughammer, J., Schuster, L. C., Kuchler, A., Alpar, D., and Bock, C. (2017) Pooled CRISPR screening with single-cell transcriptome readout. *Nat Methods* 14, 297-301
10. Ades, E. W., Candal, F. J., Swerlick, R. A., George, V. G., Summers, S., Bosse, D. C., and Lawley, T. J. (1992) HMEC-1: establishment of an immortalized human microvascular endothelial cell line. *J Invest Dermatol* 99, 683-690



## FIGURE LEGENDS:

### **Figure 1: Impact of OSA- and TSA-CRISPRi on NLUCAT1 expression in A549 cells.**

NLUCAT1 expression was measured by RT-qPCR at 4 (A) and 11 (B) days after infection. Data represented the percentage of expression of NLUCAT1 in cells treated with NLUCAT1 sgRNA3 or sgRNA5 compared to cells treated with a non-relevant guide using either OSA- or TSA-CRISPRi. (C) Data represented the percentage of dCas9 expression in OSA-CRISPRi condition compared to TSA-CRISPRi condition. Data are means  $\pm$  SD of duplicates and are representative of three independent experiments.

### **Figure 2: Effects of OSA-CRISPRi and TSA-CRISPRi on NLUCAT1 sgRNAs expression in A549 cells.**

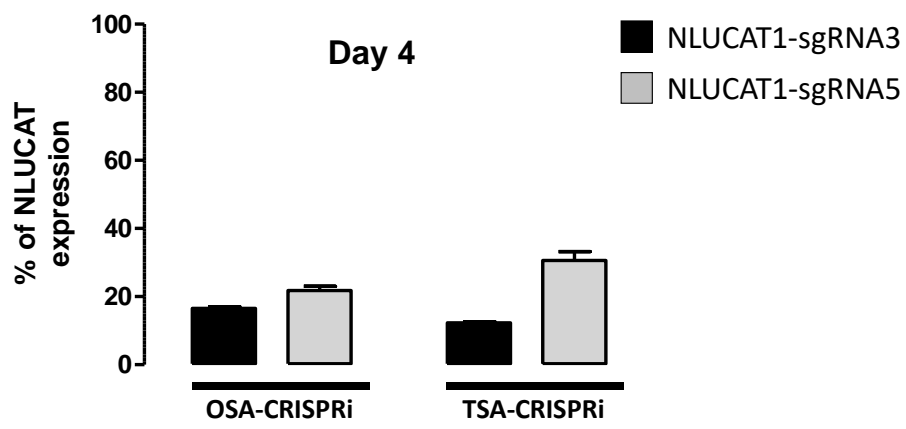
(A, B) Level of expression of NLUCAT1 sgRNA3 and sgRNA5 was analyzed by RT-qPCR at 4 (A) and 11 days (B) after infection. Data are expressed as fold of expression compared to that of a reference gene (RPLP0). (C,D) Schematic representation of expression cassettes used in the TSA- (C) and OSA-CRISPRi (D) experiments. (C) Infection of KRAB-dCas9-expressing mammalian cells with a lentivirus generated from the sgRNA-containing CROPseq-Guide-Puro plasmid leads to the generation of two sgRNAs driven by the U6 promoter (that will repress target gene transcription) and an EF1 $\alpha$  promoter-driven transcript containing the puromycin resistance coding sequence, the WPRE sequence and the sgRNA sequence that can be used as a tag for single cell RNA-seq approaches. (D) In the OSA-CRISPRi, cells are infected with lentiviruses generated from the sgRNA-containing pLenti-(BB)-EF1 $\alpha$ -KRAB-dCas9-P2A-EGFP plasmid. The EF1 $\alpha$  promoter drives the expression of the KRAB-dCas9, while the U6 promoter drives the expression of the sgRNA.

### **Figure 3: OSA-CRISPRi inhibited NLUCAT1 expression in HMEC1 cells..**

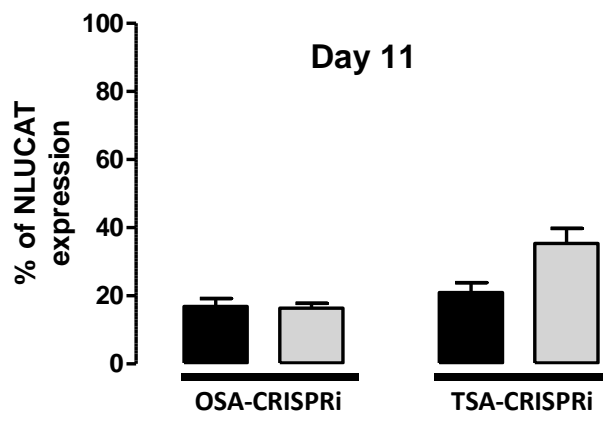
NLUCAT1 expression was analyzed in HMEC1 cells 11 days after infection with OSA-CRISPRi vector containing NLUCAT1 sgRNA3 or sgRNA5. Data are means  $\pm$  SD of duplicates and are representative of two independent experiments.

Figure 1

A



B



C

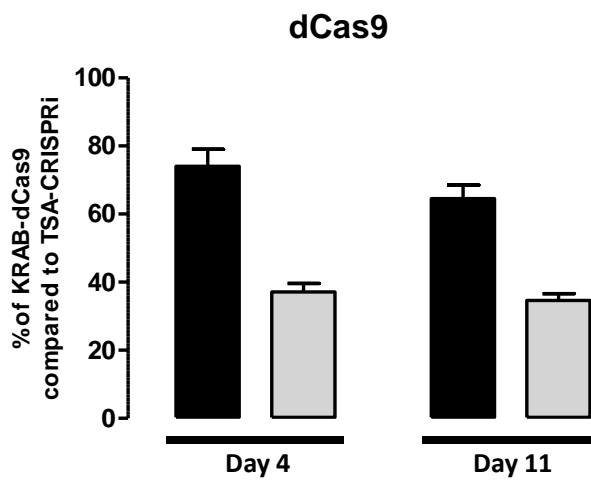


Figure 2

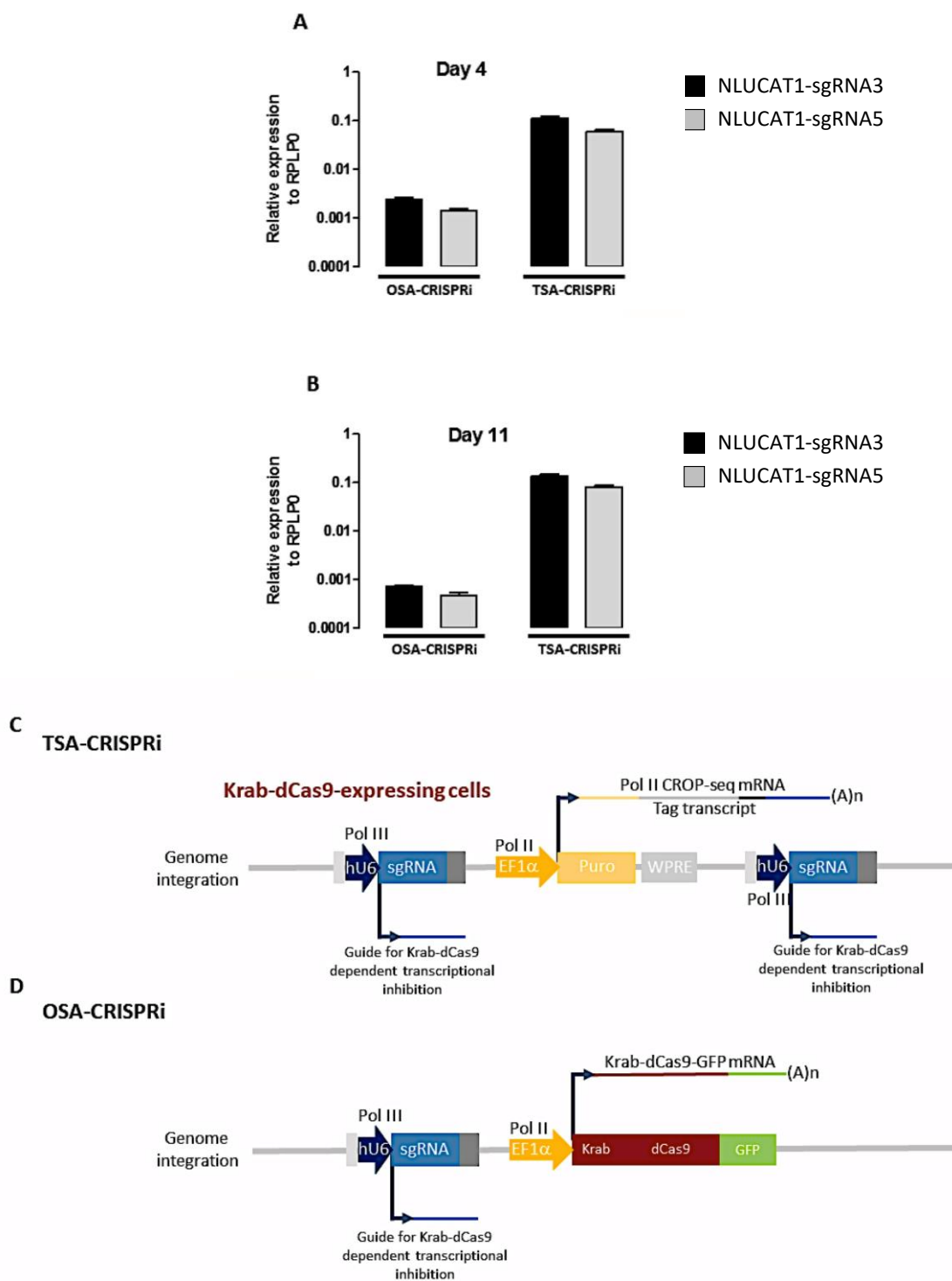
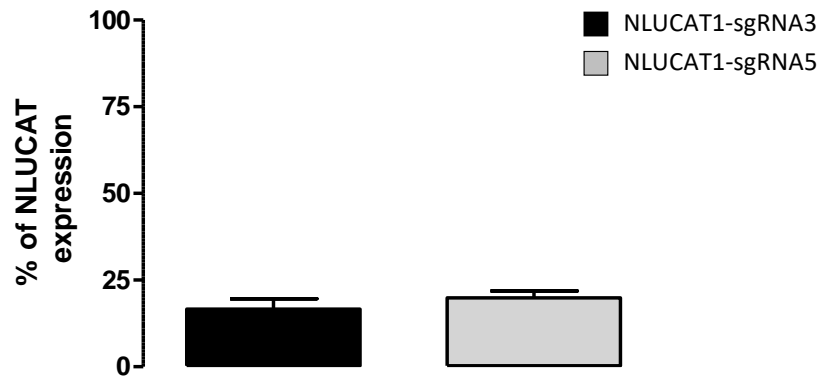


Figure 3



**Table 1: List of DNA oligonucleotides used for sgRNA cloning**

Sequences in bold correspond to the actual sgRNA sequence. Underlined sequences are identical to the plasmids and allow annealing with the receiving plasmids (CROPSeq-Guide-Puro or pLenti-(BB)-EF1a-KRAB-dCas9-P2A-EGFP).

<b>sgRNA</b>	<b>Sequences</b>
NLUCAT1 sgRNA3	<u>TGGAAAGGACGAAACACCG</u> <b>AAGCTCGGATTGCCTTAGAC</b> <u>GTTTTAGAGC</u> <u>TAGAAATAGCAAGTTAAAATAAGGC</u>
NLUCAT1 sgRNA5	<u>TGGAAAGGACGAAACACCG</u> <b>GACTTGCCACCCTGGTTGAT</b> <u>GTTTTAGAGC</u> <u>TAGAAATAGCAAGTTAAAATAAGGC</u>
Non relevant sgRNA	<u>TGGAAAGGACGAAACACCG</u> <b>CGCCAAACGTGCCCTGACGG</b> <u>GTTTTAGAG</u> <u>CTAGAAATAGCAAGTTAAAATAAGGC</u>

**Table 2:** List of RT-qPCR primers used in this study.

<b>Gene</b>	<b>Primers</b>	<b>Sequences</b>
RPLP0	Forward:	GCATCAGTACCCCATTCTATCAT
	Reverse:	AGGTGTAATCCGTCTCCACAGA
dCas9	Forward:	CAGATTTTCCAGCCGTCTGCTC
	Reverse:	AGTACCCACCATCTACCACC
NLUCAT1	Forward:	TCAAGCTCGGATTGCCTT
	Reverse:	AGGAAAGGAGCCAGAAGTCAG
NLUCAT1 sgRNA3	Forward:	AAGCTCGGATTGCCTTAGAC
	Reverse:	GCACCGACTCGGTGCCACTT
NLUCAT1 sgRNA5	Forward:	GACTTGCCACCCTGGTTGAT
	Reverse:	GCACCGACTCGGTGCCACTT
Non relevant sgRNA	Forward:	CGCCAAACGTGCCCTGACG
	Reverse:	GCACCGACTCGGTGCCACTT

## **VII. REFERENCES BIBLIOGRAPHIQUES**

1. Kotton DN, Morrisey EE. Lung regeneration: mechanisms, applications and emerging stem cell populations. *Nat Med*. 2014;20(8):822-832. doi:10.1038/nm.3642
2. Ruiz García S, Deprez M, Lebrigand K, et al. Novel dynamics of human mucociliary differentiation revealed by single-cell RNA sequencing of nasal epithelial cultures. *Dev Camb Engl*. 2019;146(20). doi:10.1242/dev.177428
3. Dela Cruz CS, Tanoue LT, Matthay RA. Lung Cancer: Epidemiology, Etiology, and Prevention. *Clin Chest Med*. 2011;32(4). doi:10.1016/j.ccm.2011.09.001
4. Cesar A. Moran MD. Slide Show: Non–Small-Cell Lung Cancer Pathology. Cancer Network. Accessed April 1, 2020. <https://www.cancernetwork.com/slideshows/slide-show-non-small-cell-lung-cancer-pathology>
5. Wellmann KF. [SMOKING AND HEALTH. ON THE REPORT OF THE ADVISORY COMMITTEE TO THE SURGEON GENERAL OF THE PUBLIC HEALTH SERVICE]. *Dtsch Med Wochenschr* 1946. 1964;89:1085-1086.
6. Gross AJ. Presentation: The risk of lung cancer in nonsmokers in the United States and its reported association with environmental tobacco smoke. *J Clin Epidemiol*. 1995;48(5):587-598.
7. Yoshida K, Gowers KHC, Lee-Six H, et al. Tobacco smoking and somatic mutations in human bronchial epithelium. *Nature*. 2020;578(7794):266-272. doi:10.1038/s41586-020-1961-1
8. Centers for Disease Control and Prevention (US), National Center for Chronic Disease Prevention and Health Promotion (US), Office on Smoking and Health (US). *How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General*. Centers for Disease Control and Prevention (US); 2010. Accessed August 19, 2020. <http://www.ncbi.nlm.nih.gov/books/NBK53017/>
9. Avino P, Scungio M, Stabile L, Cortellessa G, Buonanno G, Manigrasso M. Second-hand aerosol from tobacco and electronic cigarettes: Evaluation of the smoker emission rates and doses and lung cancer risk of passive smokers and vapers. *Sci Total Environ*. 2018;642:137-147. doi:10.1016/j.scitotenv.2018.06.059
10. Scott A, Lugg ST, Aldridge K, et al. Pro-inflammatory effects of e-cigarette vapour condensate on human alveolar macrophages. *Thorax*. 2018;73(12):1161-1169. doi:10.1136/thoraxjnl-2018-211663
11. Vogeltanz-Holm N, Schwartz GG. Radon and lung cancer: What does the public really know? *J Environ Radioact*. 2018;192:26-31. doi:10.1016/j.jenvrad.2018.05.017
12. Rivera GA, Wakelee H. Lung Cancer in Never Smokers. *Adv Exp Med Biol*. 2016;893:43-57. doi:10.1007/978-3-319-24223-1\_3
13. Kulhánová I, Morelli X, Le Tertre A, et al. The fraction of lung cancer incidence attributable to fine particulate air pollution in France: Impact of spatial resolution of air pollution models. *Environ Int*. Published online 30 October 2018. doi:10.1016/j.envint.2018.09.055
14. Marant Micallef C, Shield KD, Vignat J, et al. Cancers in France in 2015 attributable to occupational exposures. *Int J Hyg Environ Health*. 2019;222(1):22-29. doi:10.1016/j.ijheh.2018.07.015
15. Provost D, Cantagrel A, Lebaillly P, et al. Brain tumours and exposure to pesticides: a case-control study in southwestern France. *Occup Environ Med*. 2007;64(8):509-514. doi:10.1136/oem.2006.028100
16. Li L, Tian T, Zhang X. The impact of radiation on the development of lung cancer. *J Theor Biol*. 2017;428:147-152. doi:10.1016/j.jtbi.2017.06.020



17. Sellers TA, Bailey-Wilson JE, Elston RC, et al. Evidence for mendelian inheritance in the pathogenesis of lung cancer. *J Natl Cancer Inst.* 1990;82(15):1272-1279.
18. Côté ML, Liu M, Bonassi S, et al. Increased risk of lung cancer in individuals with a family history of the disease: A pooled analysis from the International Lung Cancer Consortium. *Eur J Cancer Oxf Engl 1990.* 2012;48(13):1957-1968. doi:10.1016/j.ejca.2012.01.038
19. Bailey-Wilson JE, Amos CI, Pinney SM, et al. A Major Lung Cancer Susceptibility Locus Maps to Chromosome 6q23–25. *Am J Hum Genet.* 2004;75(3):460-474.
20. Amos CI, Pinney SM, Li Y, et al. A susceptibility locus on chromosome 6q greatly increases risk lung cancer risk among light and never smokers. *Cancer Res.* 2010;70(6):2359-2367. doi:10.1158/0008-5472.CAN-09-3096
21. Fang S, Pinney SM, Bailey-Wilson JE, et al. Ordered subset analysis identifies loci influencing lung cancer risk on chromosomes 6q and 12q. *Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol.* 2010;19(12):3157-3166. doi:10.1158/1055-9965.EPI-10-0792
22. Marshall AL, Christiani DC. Genetic susceptibility to lung cancer—light at the end of the tunnel? *Carcinogenesis.* 2013;34(3):487-502. doi:10.1093/carcin/bgt016
23. Schwartz AG, Cote ML. Epidemiology of Lung Cancer. In: Ahmad A, Gadgeel S, eds. *Lung Cancer and Personalized Medicine: Current Knowledge and Therapies.* Advances in Experimental Medicine and Biology. Springer International Publishing; 2016:21-41. doi:10.1007/978-3-319-24223-1\_2
24. Denholm R, Schüz J, Straif K, et al. Is Previous Respiratory Disease a Risk Factor for Lung Cancer? *Am J Respir Crit Care Med.* 2014;190(5):549-559. doi:10.1164/rccm.201402-0338OC
25. Marcus JL, Leyden WA, Chao CR, et al. Immunodeficiency, AIDS-related pneumonia, and risk of lung cancer among HIV-infected individuals: *AIDS.* 2017;31(7):989-993. doi:10.1097/QAD.0000000000001434
26. Rezazadeh A, Laber DA, Ghim S-J, Jenson AB, Kloecker G. The Role of Human Papilloma Virus in Lung Cancer: A Review of the Evidence. *Am J Med Sci.* 2009;338(1):64-67. doi:10.1097/MAJ.0b013e3181a393ba
27. Hanahan D, Weinberg RA. Hallmarks of Cancer: The Next Generation. *Cell.* 2011;144(5):646-674. doi:10.1016/j.cell.2011.02.013
28. Hanahan D, Weinberg RA. The Hallmarks of Cancer. *Cell.* 2000;100(1):57-70. doi:10.1016/S0092-8674(00)81683-9
29. Imielinski M, Berger AH, Hammerman PS, et al. Mapping the Hallmarks of Lung Adenocarcinoma with Massively Parallel Sequencing. *Cell.* 2012;150(6):1107-1120. doi:10.1016/j.cell.2012.08.029
30. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature.* 2013;500(7463):415-421. doi:10.1038/nature12477
31. Marjanovic ND, Hofree M, Chan JE, et al. Emergence of a High-Plasticity Cell State during Lung Cancer Evolution. *Cancer Cell.* 2020;38(2):229-246.e13. doi:10.1016/j.ccell.2020.06.012
32. LaFave LM, Kartha VK, Ma S, et al. Epigenomic State Transitions Characterize Tumor Progression in Mouse Lung Adenocarcinoma. *Cancer Cell.* 2020;38(2):212-228.e13. doi:10.1016/j.ccell.2020.06.006

33. Maynard A, McCoach CE, Rotow JK, et al. Therapy-Induced Evolution of Human Lung Cancer Revealed by Single-Cell RNA Sequencing. *Cell*. 2020;182(5):1232-1251.e22. doi:10.1016/j.cell.2020.07.017
34. Soh J, Okumura N, Lockwood WW, et al. Oncogene Mutations, Copy Number Gains and Mutant Allele Specific Imbalance (MASI) Frequently Occur Together in Tumor Cells. *PLOS ONE*. 2009;4(10):e7464. doi:10.1371/journal.pone.0007464
35. Zito Marino F, Bianco R, Accardo M, et al. Molecular heterogeneity in lung cancer: from mechanisms of origin to clinical implications. *Int J Med Sci*. 2019;16(7):981-989. doi:10.7150/ijms.34739
36. Haber DA, Settleman J. Drivers and passengers. *Nature*. 2007;446(7132):145-146. doi:10.1038/446145a
37. Collisson EA, Campbell JD, Brooks AN, et al. Comprehensive molecular profiling of lung adenocarcinoma. *Nature*. 2014;511(7511):543-550. doi:10.1038/nature13385
38. Cafarotti S, Lococo F, Froesh P, Zappa F, Andrè D. Target Therapy in Lung Cancer. *Adv Exp Med Biol*. 2016;893:127-136. doi:10.1007/978-3-319-24223-1\_6
39. Midha A, Dearden S, McCormack R. EGFR mutation incidence in non-small-cell lung cancer of adenocarcinoma histology: a systematic review and global map by ethnicity (mutMapII). *Am J Cancer Res*. 2015;5(9):2892-2911.
40. Herbst RS, Heymach JV, Lippman SM. Lung Cancer. *N Engl J Med*. 2008;359(13):1367-1380. doi:10.1056/NEJMra0802714
41. Siegelin MD, Borczuk AC. Epidermal growth factor receptor mutations in lung adenocarcinoma. *Lab Investig J Tech Methods Pathol*. 2014;94(2):129-137. doi:10.1038/labinvest.2013.147
42. Boriack-Sjodin PA, Margarit SM, Bar-Sagi D, Kuriyan J. The structural basis of the activation of Ras by Sos. *Nature*. 1998;394(6691):337-343. doi:10.1038/28548
43. Scheffzek K, Ahmadian MR, Kabsch W, et al. The Ras-RasGAP complex: structural basis for GTPase activation and its loss in oncogenic Ras mutants. *Science*. 1997;277(5324):333-338. doi:10.1126/science.277.5324.333
44. Karachaliou N, Mayo C, Costa C, et al. KRAS Mutations in Lung Cancer. *Clin Lung Cancer*. 2013;14(3):205-214. doi:10.1016/j.clcc.2012.09.007
45. Yu HA, Sima CS, Shen R, et al. Prognostic Impact of KRAS Mutation Subtypes in 677 Patients with Metastatic Lung Adenocarcinomas. *J Thorac Oncol*. 2015;10(3):431-437. doi:10.1097/JTO.0000000000000432
46. Kerr EM, Gaude E, Turrell FK, Frezza C, Martins CP. Mutant Kras copy number defines metabolic reprogramming and therapeutic susceptibilities. *Nature*. 2016;531(7592):110-113. doi:10.1038/nature16967
47. Skoulidis F, Byers LA, Diao L, et al. Co-occurring genomic alterations define major subsets of KRAS-mutant lung adenocarcinoma with distinct biology, immune profiles, and therapeutic vulnerabilities. *Cancer Discov*. 2015;5(8):860-877. doi:10.1158/2159-8290.CD-14-1236
48. Linzer DIH, Levine AJ. Characterization of a 54K Dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. *Cell*. 1979;17(1):43-52. doi:10.1016/0092-8674(79)90293-9

49. Haupt Y, Maya R, Kazaz A, Oren M. Mdm2 promotes the rapid degradation of p53. *Nature*. 1997;387(6630):296-299. doi:10.1038/387296a0
50. Levine AJ, Momand J, Finlay CA. The p53 tumour suppressor gene. *Nature*. 1991;351(6326):453-456. doi:10.1038/351453a0
51. Duffy MJ, Synnott NC, Crown J. Mutant p53 as a target for cancer treatment. *Eur J Cancer*. 2017;83:258-265. doi:10.1016/j.ejca.2017.06.023
52. Gibbons DL, Byers LA, Kurie JM. Smoking, p53 Mutation, and Lung Cancer. *Mol Cancer Res MCR*. 2014;12(1):3-13. doi:10.1158/1541-7786.MCR-13-0539
53. Kadara H, Sivakumar S, Jakubek Y, et al. Driver Mutations in Normal Airway Epithelium Elucidate Spatiotemporal Resolution of Lung Cancer. *Am J Respir Crit Care Med*. 2019;200(6):742-750. doi:10.1164/rccm.201806-1178OC
54. Pao W, Girard N. New driver mutations in non-small-cell lung cancer. *Lancet Oncol*. 2011;12(2):175-180. doi:10.1016/S1470-2045(10)70087-5
55. Baik CS, Myall NJ, Wakelee HA. Targeting BRAF-Mutant Non-Small Cell Lung Cancer: From Molecular Profiling to Rationally Designed Therapy. *The Oncologist*. 2017;22(7):786-796. doi:10.1634/theoncologist.2016-0458
56. Soda M, Choi YL, Enomoto M, et al. Identification of the transforming EML4–ALK fusion gene in non-small-cell lung cancer. *Nature*. 2007;448(7153):561-566. doi:10.1038/nature05945
57. Comprehensive molecular profiling of lung adenocarcinoma. *Nature*. 2014;511(7511):543-550. doi:10.1038/nature13385
58. Ji H, Ramsey MR, Hayes DN, et al. LKB1 modulates lung cancer differentiation and metastasis. *Nature*. 2007;448(7155):807-810. doi:10.1038/nature06030
59. Sanchez-Cespedes M, Parrella P, Esteller M, et al. Inactivation of LKB1/STK11 Is a Common Event in Adenocarcinomas of the Lung. *Cancer Res*. 2002;62(13):3659-3662.
60. Singh A, Misra V, Thimmulappa RK, et al. Dysfunctional KEAP1-NRF2 interaction in non-small-cell lung cancer. *PLoS Med*. 2006;3(10):e420. doi:10.1371/journal.pmed.0030420
61. Qian J, Olbrecht S, Boeckx B, et al. A pan-cancer blueprint of the heterogeneous tumor microenvironment revealed by single-cell profiling. *Cell Res*. Published online 19 June 2020:1-18. doi:10.1038/s41422-020-0355-0
62. Sinjab A, Han G, Hara K, et al. *Resolving the Spatial and Cellular Architecture of Lung Adenocarcinoma by Multi-Region Single-Cell Sequencing*. *Cancer Biology*; 2020. doi:10.1101/2020.09.04.283739
63. Bremnes RM, Al-Shibli K, Donnem T, et al. The Role of Tumor-Infiltrating Immune Cells and Chronic Inflammation at the Tumor Site on Cancer Development, Progression, and Prognosis: Emphasis on Non-small Cell Lung Cancer. *J Thorac Oncol*. 2011;6(4):824-833. doi:10.1097/JTO.0b013e3182037b76
64. Chouaib S. Au carrefour du cancer. *Bull Cancer (Paris)*. 2013;100(6):569-574. doi:10.1684/bdc.2013.1754
65. Jain P, Jain C, Velcheti V. Role of immune-checkpoint inhibitors in lung cancer. *Ther Adv Respir Dis*. 2018;12:1753465817750075. doi:10.1177/1753465817750075

66. Coussens LM, Werb Z. Inflammation and cancer. *Nature*. 2002;420(6917):860-867. doi:10.1038/nature01322
67. Gomes M, Teixeira AL, Coelho A, Araújo A, Medeiros R. The role of inflammation in lung cancer. *Adv Exp Med Biol*. 2014;816:1-23. doi:10.1007/978-3-0348-0837-8\_1
68. Carmeliet P. Angiogenesis in health and disease. *Nat Med*. 2003;9(6):653-660. doi:10.1038/nm0603-653
69. Alitalo K, Carmeliet P. Molecular mechanisms of lymphangiogenesis in health and disease. *Cancer Cell*. 2002;1(3):219-227. doi:10.1016/S1535-6108(02)00051-X
70. *Foundations of Manual Lymph Drainage*. Elsevier; 2005. doi:10.1016/B978-0-323-03064-9.X5001-4
71. Paupert J, Sounni NE, Noël A. Lymphangiogenesis in post-natal tissue remodeling: lymphatic endothelial cell connection with its environment. *Mol Aspects Med*. 2011;32(2):146-158. doi:10.1016/j.mam.2011.04.002
72. Saharinen P, Tammela T, Karkkainen MJ, Alitalo K. Lymphatic vasculature: development, molecular regulation and role in tumor metastasis and inflammation. *Trends Immunol*. 2004;25(7):387-395. doi:10.1016/j.it.2004.05.003
73. Goldmann E. The Growth of Malignant Disease in Man and the Lower Animals, with special reference to the Vascular System. *Proc R Soc Med*. 1908;1(Surg Sect):1-13.
74. Tammela T, Alitalo K. Lymphangiogenesis: Molecular Mechanisms and Future Promise. *Cell*. 2010;140(4):460-476. doi:10.1016/j.cell.2010.01.045
75. Brown M, Assen FP, Leithner A, et al. Lymph node blood vessels provide exit routes for metastatic tumor cell dissemination in mice. *Science*. 2018;359(6382):1408-1411. doi:10.1126/science.aal3662
76. Pereira ER, Kedrin D, Seano G, et al. Lymph node metastases can invade local blood vessels, exit the node, and colonize distant organs in mice. *Science*. 2018;359(6382):1403-1407. doi:10.1126/science.aal3622
77. Nathanson SD, Kwon D, Kapke A, Alford SH, Chitale D. The role of lymph node metastasis in the systemic dissemination of breast cancer. *Ann Surg Oncol*. 2009;16(12):3396-3405. doi:10.1245/s10434-009-0659-2
78. Carmeliet P, Jain RK. Angiogenesis in cancer and other diseases. *Nature*. 2000;407(6801):249-257. doi:10.1038/35025220
79. Caporarello N, Lupo G, Olivieri M, et al. Classical VEGF, Notch and Ang signalling in cancer angiogenesis, alternative approaches and future directions. *Mol Med Rep*. 2017;16(4):4393-4402. doi:10.3892/mmr.2017.7179
80. Niki T, Iba S, Tokunou M, Yamada T, Matsuno Y, Hirohashi S. Expression of vascular endothelial growth factors A, B, C, and D and their relationships to lymph node status in lung adenocarcinoma. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2000;6(6):2431-2439.
81. Iso Tatsuya, Hamamori Yasuo, Kedes Larry. Notch Signaling in Vascular Development. *Arterioscler Thromb Vasc Biol*. 2003;23(4):543-553. doi:10.1161/01.ATV.0000060892.81529.8F
82. Hellström M, Phng L-K, Hofmann JJ, et al. DLL4 signalling through Notch1 regulates formation of tip cells during angiogenesis. *Nature*. 2007;445(7129):776-780. doi:10.1038/nature05571

83. Patel NS, Li J-L, Generali D, Poulsom R, Cranston DW, Harris AL. Up-regulation of Delta-like 4 Ligand in Human Tumor Vasculature and the Role of Basal Expression in Endothelial Cell Function. *Cancer Res.* 2005;65(19):8690-8697. doi:10.1158/0008-5472.CAN-05-1208
84. Rodriguez-Vita J, Tetzlaff F, Fischer A. Notch controls endothelial cells. *Oncoscience.* 2017;4(5-6):45-46. doi:10.18632/oncoscience.355
85. Laakkonen P, Waltari M, Holopainen T, et al. Vascular Endothelial Growth Factor Receptor 3 Is Involved in Tumor Angiogenesis and Growth. *Cancer Res.* 2007;67(2):593-599. doi:10.1158/0008-5472.CAN-06-3567
86. Saif MW, Knost JA, Chiorean EG, et al. Phase 1 study of the anti-vascular endothelial growth factor receptor 3 monoclonal antibody LY3022856/IMC-3C5 in patients with advanced and refractory solid tumors and advanced colorectal cancer. *Cancer Chemother Pharmacol.* 2016;78(4):815-824. doi:10.1007/s00280-016-3134-3
87. Tampellini M, Sonetto C, Scagliotti GV. Novel anti-angiogenic therapeutic strategies in colorectal cancer. *Expert Opin Investig Drugs.* 2016;25(5):507-520. doi:10.1517/13543784.2016.1161754
88. Spratlin JL, Cohen RB, Eadens M, et al. Phase I Pharmacologic and Biologic Study of Ramucirumab (IMC-1121B), a Fully Human Immunoglobulin G1 Monoclonal Antibody Targeting the Vascular Endothelial Growth Factor Receptor-2. *J Clin Oncol.* 2010;28(5):780-787. doi:10.1200/JCO.2009.23.7537
89. Hurwitz H, Saini S. Bevacizumab in the treatment of metastatic colorectal cancer: safety profile and management of adverse events. *Semin Oncol.* 2006;33(5 Suppl 10):S26-34. doi:10.1053/j.seminoncol.2006.08.001
90. Falchook GS, Goldman JW, Desai J, et al. A first-in-human phase I study of VGX-100, a selective anti-VEGF-C antibody, alone and in combination with bevacizumab in patients with advanced solid tumors. *J Clin Oncol.* 2014;32(15\_suppl):2524-2524. doi:10.1200/jco.2014.32.15\_suppl.2524
91. Oaknin A, Floquet A, Le Tourneau C, et al. Single agent vanucizumab (RO5520985) for platinum (Pt)-resistant recurrent ovarian cancer (OC): Results from a single arm extension phase of the phase I FIH study. *J Clin Oncol.* 2015;33(15\_suppl):5549-5549. doi:10.1200/jco.2015.33.15\_suppl.5549
92. Kalluri R, Zeisberg M. Fibroblasts in cancer. *Nat Rev Cancer.* 2006;6(5):392-401. doi:10.1038/nrc1877
93. Königshoff M. Lung Cancer in Pulmonary Fibrosis: Tales of Epithelial Cell Plasticity. *Respiration.* 2011;81(5):353-358. doi:10.1159/000326299
94. Nair N, Calle AS, Zahra MH, et al. A cancer stem cell model as the point of origin of cancer-associated fibroblasts in tumor microenvironment. *Sci Rep.* 2017;7. doi:10.1038/s41598-017-07144-5
95. Kinoshita T, Goto T. Molecular Mechanisms of Pulmonary Fibrogenesis and Its Progression to Lung Cancer: A Review. *Int J Mol Sci.* 2019;20(6). doi:10.3390/ijms20061461
96. Jang I, Beningo KA. Integrins, CAFs and Mechanical Forces in the Progression of Cancer. *Cancers.* 2019;11(5). doi:10.3390/cancers11050721
97. Kidd S, Spaeth E, Watson K, et al. Origins of the Tumor Microenvironment: Quantitative Assessment of Adipose-Derived and Bone Marrow-Derived Stroma. *PLOS ONE.* 2012;7(2):e30563. doi:10.1371/journal.pone.0030563

98. Zeisberg EM, Potenta S, Xie L, Zeisberg M, Kalluri R. Discovery of Endothelial to Mesenchymal Transition as a Source for Carcinoma-Associated Fibroblasts. *Cancer Res.* 2007;67(21):10123-10128. doi:10.1158/0008-5472.CAN-07-3127
99. Bartoschek M, Oskolkov N, Bocci M, et al. Spatially and functionally distinct subclasses of breast cancer-associated fibroblasts revealed by single cell RNA sequencing. *Nat Commun.* 2018;9(1):5150. doi:10.1038/s41467-018-07582-3
100. Qian L-W, Mizumoto K, Maehara N, et al. Co-cultivation of pancreatic cancer cells with orthotopic tumor-derived fibroblasts: fibroblasts stimulate tumor cell invasion via HGF secretion whereas cancer cells exert a minor regulative effect on fibroblasts HGF production. *Cancer Lett.* 2003;190(1):105-112. doi:10.1016/s0304-3835(02)00517-7
101. Wilson TR, Fridlyand J, Yan Y, et al. Widespread potential for growth-factor-driven resistance to anticancer kinase inhibitors. *Nature.* 2012;487(7408):505-509. doi:10.1038/nature11249
102. Grinnell F. Fibroblast biology in three-dimensional collagen matrices. *Trends Cell Biol.* 2003;13(5):264-269. doi:10.1016/s0962-8924(03)00057-6
103. Gaggioli C, Hooper S, Hidalgo-Carcedo C, et al. Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells. *Nat Cell Biol.* 2007;9(12):1392-1400. doi:10.1038/ncb1658
104. Monteran L, Erez N. The Dark Side of Fibroblasts: Cancer-Associated Fibroblasts as Mediators of Immunosuppression in the Tumor Microenvironment. *Front Immunol.* 2019;10. doi:10.3389/fimmu.2019.01835
105. Santos AM, Jung J, Aziz N, Kissil JL, Puré E. Targeting fibroblast activation protein inhibits tumor stromagenesis and growth in mice. *J Clin Invest.* 2009;119(12):3613-3625. doi:10.1172/JCI38988
106. Ostermann E, Garin-Chesa P, Heider KH, et al. Effective Immunoconjugate Therapy in Cancer Models Targeting a Serine Protease of Tumor Fibroblasts. *Clin Cancer Res.* 2008;14(14):4584-4592. doi:10.1158/1078-0432.CCR-07-5211
107. Özdemir BC, Pentcheva-Hoang T, Carstens JL, et al. Depletion of Carcinoma-Associated Fibroblasts and Fibrosis Induces Immunosuppression and Accelerates Pancreas Cancer with Reduced Survival. *Cancer Cell.* 2014;25(6):719-734. doi:10.1016/j.ccr.2014.04.005
108. Al Tameemi W, Dale TP, Al-Jumaily RMK, Forsyth NR. Hypoxia-Modified Cancer Cell Metabolism. *Front Cell Dev Biol.* 2019;7. doi:10.3389/fcell.2019.00004
109. Harris AL. Hypoxia — a key regulatory factor in tumour growth. *Nat Rev Cancer.* 2002;2(1):38-47. doi:10.1038/nrc704
110. Ivan M, Kondo K, Yang H, et al. HIF $\alpha$  Targeted for VHL-Mediated Destruction by Proline Hydroxylation: Implications for O<sub>2</sub> Sensing. *Science.* 2001;292(5516):464-468. doi:10.1126/science.1059817
111. Loboda A, Jozkowicz A, Dulak J. HIF-1 versus HIF-2 — Is one more important than the other? *Vascul Pharmacol.* 2012;56(5):245-251. doi:10.1016/j.vph.2012.02.006
112. Koh MY, Lemos R, Liu X, Powis G. The Hypoxia-Associated Factor Switches Cells from HIF-1 $\alpha$ - to HIF-2 $\alpha$ -Dependent Signaling Promoting Stem Cell Characteristics, Aggressive Tumor Growth and Invasion. *Cancer Res.* 2011;71(11):4015-4027. doi:10.1158/0008-5472.CAN-10-4142

113. Cummins EP, Berra E, Comerford KM, et al. Prolyl hydroxylase-1 negatively regulates I $\kappa$ B kinase- $\beta$ , giving insight into hypoxia-induced NF $\kappa$ B activity. *Proc Natl Acad Sci*. 2006;103(48):18154-18159. doi:10.1073/pnas.0602235103
114. Koong AC, Chen EY, Giaccia AJ. Hypoxia causes the activation of nuclear factor kappa B through the phosphorylation of I kappa B alpha on tyrosine residues. *Cancer Res*. 1994;54(6):1425-1430.
115. Toth RK, Warfel NA. Strange Bedfellows: Nuclear Factor, Erythroid 2-Like 2 (Nrf2) and Hypoxia-Inducible Factor 1 (HIF-1) in Tumor Hypoxia. *Antioxidants*. 2017;6(2). doi:10.3390/antiox6020027
116. Vander Heiden MG, Cantley LC, Thompson CB. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science*. 2009;324(5930):1029-1033. doi:10.1126/science.1160809
117. Razungles J, Cavaillès V, Jalaguier S, Teyssier C. L'effet Warburg: De la théorie du cancer aux applications thérapeutiques en cancérologie. *médecine/sciences*. 2013;29(11):1026-1033. doi:10.1051/medsci/20132911020
118. Liberti MV, Locasale JW. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends Biochem Sci*. 2016;41(3):211-218. doi:10.1016/j.tibs.2015.12.001
119. Pugh CW, Ratcliffe PJ. Regulation of angiogenesis by hypoxia: role of the HIF system. *Nat Med*. 2003;9(6):677-684. doi:10.1038/nm0603-677
120. Liu Yuxiang, Cox Shanna R., Morita Toshisuke, Kourembanas Stella. Hypoxia Regulates Vascular Endothelial Growth Factor Gene Expression in Endothelial Cells. *Circ Res*. 1995;77(3):638-643. doi:10.1161/01.RES.77.3.638
121. Koch S, Claesson-Welsh L. Signal Transduction by Vascular Endothelial Growth Factor Receptors. *Cold Spring Harb Perspect Med*. 2012;2(7). doi:10.1101/cshperspect.a006502
122. Nahomi RB, Nagaraj RH. The role of HIF-1 $\alpha$  in the TGF- $\beta$ 2-mediated epithelial-to-mesenchymal transition of human lens epithelial cells. *J Cell Biochem*. 2018;119(8):6814-6827. doi:10.1002/jcb.26877
123. Cox TR, Rumney RMH, Schoof EM, et al. The hypoxic cancer secretome induces pre-metastatic bone lesions through lysyl oxidase. *Nature*. 2015;522(7554):106-110. doi:10.1038/nature14492
124. Wong CC-L, Gilkes DM, Zhang H, et al. Hypoxia-inducible factor 1 is a master regulator of breast cancer metastatic niche formation. *Proc Natl Acad Sci U S A*. 2011;108(39):16369-16374. doi:10.1073/pnas.1113483108
125. Jing X, Yang F, Shao C, et al. Role of hypoxia in cancer therapy by regulating the tumor microenvironment. *Mol Cancer*. 2019;18(1):157. doi:10.1186/s12943-019-1089-9
126. Lee S, Ryu H, Son A -rang, et al. TGF- $\beta$  and Hypoxia/Reoxygenation Promote Radioresistance of A549 Lung Cancer Cells through Activation of Nrf2 and EGFR. *Oxid Med Cell Longev*. 2016;2016:6823471. doi:10.1155/2016/6823471
127. Ohta T, Iijima K, Miyamoto M, et al. Loss of Keap1 Function Activates Nrf2 and Provides Advantages for Lung Cancer Cell Growth. *Cancer Res*. 2008;68(5):1303-1309. doi:10.1158/0008-5472.CAN-07-5003
128. Michels S, Wolf J. Stratified Treatment in Lung Cancer. *Oncol Res Treat*. 2016;39(12):760-766. doi:10.1159/000453406

129. Galluzzi L, Vitale I, Michels J, et al. Systems biology of cisplatin resistance: past, present and future. *Cell Death Dis.* 2014;5(5):e1257-e1257. doi:10.1038/cddis.2013.428
130. Baker S, Dahele M, Lagerwaard FJ, Senan S. A critical review of recent developments in radiotherapy for non-small cell lung cancer. *Radiat Oncol.* 2016;11(1):115. doi:10.1186/s13014-016-0693-8
131. Steven A, Fisher SA, Robinson BW. Immunotherapy for lung cancer: Immunotherapy for lung cancer. *Respirology.* 2016;21(5):821-833. doi:10.1111/resp.12789
132. Lamers CHJ, Willemsen R, van Elzakker P, et al. Immune responses to transgene and retroviral vector in patients treated with ex vivo–engineered T cells. *Blood.* 2011;117(1):72-82. doi:10.1182/blood-2010-07-294520
133. Ajina A, Maher J. Strategies to address chimeric antigen receptor tonic signalling. *Mol Cancer Ther.* 2018;17(9):1795-1815. doi:10.1158/1535-7163.MCT-17-1097
134. Benmebarek M-R, Karches CH, Cadilha BL, Lesch S, Endres S, Kobold S. Killing Mechanisms of Chimeric Antigen Receptor (CAR) T Cells. *Int J Mol Sci.* 2019;20(6). doi:10.3390/ijms20061283
135. Berinstein NL, Berinstein JA. 42 - Therapeutic cancer vaccines. In: Plotkin SA, Orenstein WA, Offit PA, eds. *Vaccines (Sixth Edition)*. W.B. Saunders; 2013:1018-1031. doi:10.1016/B978-1-4557-0090-5.00045-8
136. Holt GE, Podack ER, Raez LE. Immunotherapy as a strategy for the treatment of non-small-cell lung cancer. *Ther Lond Engl 2004.* 2011;8(1):43-54. doi:10.2217/thy.10.84
137. Watson JD, Crick FH. Genetical implications of the structure of deoxyribonucleic acid. *Nature.* 1953;171(4361):964-967. doi:10.1038/171964b0
138. Portin P. The birth and development of the DNA theory of inheritance: sixty years since the discovery of the structure of DNA. *J Genet.* 2014;93(1):293-302. doi:10.1007/s12041-014-0337-4
139. Ohno S. An argument for the genetic simplicity of man and other mammals. *J Hum Evol.* 1972;1(6):651-662. doi:10.1016/0047-2484(72)90011-5
140. Napoli C, Lemieux C, Jorgensen R. Introduction of a Chimeric Chalcone Synthase Gene into *Petunia* Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell.* 1990;2(4):279-289.
141. Lee RC, Feinbaum RL, Ambros V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell.* 1993;75(5):843-854. doi:10.1016/0092-8674(93)90529-Y
142. ENCODE Project Consortium. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science.* 2004;306(5696):636-640. doi:10.1126/science.1105136
143. Kawai J, Shinagawa A, Shibata K, et al. Functional annotation of a full-length mouse cDNA collection. *Nature.* 2001;409(6821):685-690. doi:10.1038/35055500
144. Moore JE, Purcaro MJ, Pratt HE, et al. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature.* 2020;583(7818):699-710. doi:10.1038/s41586-020-2493-4
145. Amin N, McGrath A, Chen Y-PP. Evaluation of deep learning in non-coding RNA classification. *Nat Mach Intell.* 2019;1(5):246-256. doi:10.1038/s42256-019-0051-2



146. Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. The structural basis of ribosome activity in peptide bond synthesis. *Science*. 2000;289(5481):920-930. doi:10.1126/science.289.5481.920
147. Ender C, Krek A, Friedländer MR, et al. A human snoRNA with microRNA-like functions. *Mol Cell*. 2008;32(4):519-528. doi:10.1016/j.molcel.2008.10.017
148. Clark BFC. The crystal structure of tRNA. *J Biosci*. 2006;31(4):453-457. doi:10.1007/BF02705184
149. Pingault L. Analyses structurales et fonctionnelles de l'espace génique du chromosome 3B du blé tendre (*Triticum aestivum* L.). In: ; 2014.
150. Denli AM, Tops BBJ, Plasterk RHA, Ketting RF, Hannon GJ. Processing of primary microRNAs by the Microprocessor complex. *Nature*. 2004;432(7014):231-235. doi:10.1038/nature03049
151. Bartel DP. MicroRNA Target Recognition and Regulatory Functions. *Cell*. 2009;136(2):215-233. doi:10.1016/j.cell.2009.01.002
152. Weber MJ. New human and mouse microRNA genes found by homology search. *FEBS J*. 2005;272(1):59-73. doi:10.1111/j.1432-1033.2004.04389.x
153. Wu K-L, Tsai Y-M, Lien C-T, Kuo P-L, Hung J-Y. The Roles of MicroRNA in Lung Cancer. *Int J Mol Sci*. 2019;20(7). doi:10.3390/ijms20071611
154. Grosso S, Doyen J, Parks SK, et al. MiR-210 promotes a hypoxic phenotype and increases radioresistance in human lung cancer cell lines. *Cell Death Dis*. 2013;4:e544. doi:10.1038/cddis.2013.71
155. Puisségur M-P, Mazure NM, Bertero T, et al. miR-210 is overexpressed in late stages of lung cancer and mediates mitochondrial alterations associated with modulation of HIF-1 activity. *Cell Death Differ*. 2011;18(3):465-478. doi:10.1038/cdd.2010.119
156. Lucas KJ, Myles KM, Raikhel AS. Small RNAs: a new frontier in mosquito biology. *Trends Parasitol*. 2013;29(6):295-303. doi:10.1016/j.pt.2013.04.003
157. Sen GL, Blau HM. A brief history of RNAi: the silence of the genes. *FASEB J Off Publ Fed Am Soc Exp Biol*. 2006;20(9):1293-1299. doi:10.1096/fj.06-6014rev
158. Siomi MC, Sato K, Pezic D, Aravin AA. PIWI-interacting small RNAs: the vanguard of genome defence. *Nat Rev Mol Cell Biol*. 2011;12(4):246-258. doi:10.1038/nrm3089
159. Bartolomei MS, Zemel S, Tilghman SM. Parental imprinting of the mouse H19 gene. *Nature*. 1991;351(6322):153-155. doi:10.1038/351153a0
160. 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. doi:10.1128/MCB.10.1.28
161. Jarroux J, Morillon A, Pinskaya M. History, Discovery, and Classification of lncRNAs. In: Rao MRS, ed. *Long Non Coding RNA Biology*. Advances in Experimental Medicine and Biology. Springer; 2017:1-46. doi:10.1007/978-981-10-5203-3\_1
162. Espinoza CA, Goodrich JA, Kugel JF. Characterization of the structure, function, and mechanism of B2 RNA, an ncRNA repressor of RNA polymerase II transcription. *RNA N Y N*. 2007;13(4):583-596. doi:10.1261/rna.310307
163. Hezroni H, Koppstein D, Schwartz MG, Avrutin A, Bartel DP, Ulitsky I. Principles of Long Noncoding RNA Evolution Derived from Direct Comparison of Transcriptomes in 17 Species. *Cell Rep*. 2015;11(7):1110-1122. doi:10.1016/j.celrep.2015.04.023

164. Ulitsky I, Bartel DP. lincRNAs: Genomics, Evolution, and Mechanisms. *Cell*. 2013;154(1):26-46. doi:10.1016/j.cell.2013.06.020
165. Hawkes EJ, Hennelly SP, Novikova IV, Irwin JA, Dean C, Sanbonmatsu KY. COOLAIR Antisense RNAs Form Evolutionarily Conserved Elaborate Secondary Structures. *Cell Rep*. 2016;16(12):3087-3096. doi:10.1016/j.celrep.2016.08.045
166. Ma L, Bajic VB, Zhang Z. On the classification of long non-coding RNAs. *RNA Biol*. 2013;10(6):925-933. doi:10.4161/rna.24604
167. Tang Y, Zhou T, Yu X, Xue Z, Shen N. The role of long non-coding RNAs in rheumatic diseases. *Nat Rev Rheumatol*. 2017;13(11):657-669. doi:10.1038/nrrheum.2017.162
168. Katayama S, Tomaru Y, Kasukawa T, et al. Antisense Transcription in the Mammalian Transcriptome. *Science*. 2005;309(5740):1564-1566. doi:10.1126/science.1112009
169. Khorkova O, Myers AJ, Hsiao J, Wahlestedt C. Natural antisense transcripts. *Hum Mol Genet*. 2014;23(R1):R54-R63. doi:10.1093/hmg/ddu207
170. Patrucco L, Chiesa A, Soluri MF, et al. Engineering mammalian cell factories with SINEUP noncoding RNAs to improve translation of secreted proteins. *Gene*. 2015;569(2):287-293. doi:10.1016/j.gene.2015.05.070
171. Zucchelli S, Cotella D, Takahashi H, et al. SINEUPs: A new class of natural and synthetic antisense long non-coding RNAs that activate translation. *RNA Biol*. 2015;12(8):771-779. doi:10.1080/15476286.2015.1060395
172. Lepoivre C, Belhocine M, Bergon A, et al. Divergent transcription is associated with promoters of transcriptional regulators. *BMC Genomics*. 2013;14:914. doi:10.1186/1471-2164-14-914
173. Kermezli Y, Saadi W, Belhocine M, et al. A comprehensive catalog of LncRNAs expressed in T-cell acute lymphoblastic leukemia. *Leuk Lymphoma*. 2019;60(8):2002-2014. doi:10.1080/10428194.2018.1551534
174. Zhang Y, Zhang X-O, Chen T, et al. Circular Intronic Long Noncoding RNAs. *Mol Cell*. 2013;51(6):792-806. doi:10.1016/j.molcel.2013.08.017
175. Du WW, Yang W, Liu E, Yang Z, Dhaliwal P, Yang BB. Foxo3 circular RNA retards cell cycle progression via forming ternary complexes with p21 and CDK2. *Nucleic Acids Res*. 2016;44(6):2846-2858. doi:10.1093/nar/gkw027
176. Lu M. Circular RNA: functions, applications and prospects. *ExRNA*. 2020;2(1):1. doi:10.1186/s41544-019-0046-5
177. Panda A, Grammatikakis I, Munk R, Gorospe M, Abdelmohsen K. Emerging roles and context of circular RNAs. *Wiley Interdiscip Rev RNA*. 2016;8. doi:10.1002/wrna.1386
178. Schlackow M, Nojima T, Gomes T, Dhir A, Carmo-Fonseca M, Proudfoot NJ. Distinctive Patterns of Transcription and RNA Processing for Human lincRNAs. *Mol Cell*. 2017;65(1):25-38. doi:10.1016/j.molcel.2016.11.029
179. Levine M, Tjian R. Transcription regulation and animal diversity. *Nature*. 2003;424(6945):147-151. doi:10.1038/nature01763
180. Hsin J-P, Manley JL. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev*. 2012;26(19):2119-2137. doi:10.1101/gad.200303.112

181. Orphanides G, Reinberg D. A Unified Theory of Gene Expression. *Cell*. 2002;108(4):439-451. doi:10.1016/S0092-8674(02)00655-4
182. Gressel S, Schwalb B, Cramer P. The pause-initiation limit restricts transcription activation in human cells. *Nat Commun*. 2019;10. doi:10.1038/s41467-019-11536-8
183. Chen FX, Smith ER, Shilatifard A. Born to run: control of transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol*. 2018;19(7):464-478. doi:10.1038/s41580-018-0010-5
184. Uszczyńska-Ratajczak B, Lagarde J, Frankish A, Guigó R, Johnson R. Towards a complete map of the human long non-coding RNA transcriptome. *Nat Rev Genet*. 2018;19(9):535-548. doi:10.1038/s41576-018-0017-y
185. Kowalczyk MS, Hughes JR, Garrick D, et al. Intragenic Enhancers Act as Alternative Promoters. *Mol Cell*. 2012;45(4):447-458. doi:10.1016/j.molcel.2011.12.021
186. Quinn JJ, Chang HY. Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet*. 2016;17(1):47-62. doi:10.1038/nrg.2015.10
187. Orphanides G, Reinberg D. RNA polymerase II elongation through chromatin. *Nature*. 2000;407(6803):471-475. doi:10.1038/35035000
188. Volders P-J, Helsens K, Wang X, et al. LNCipedia: a database for annotated human lncRNA transcript sequences and structures. *Nucleic Acids Res*. 2013;41(D1):D246-D251. doi:10.1093/nar/gks915
189. Fazal FM, Han S, Parker KR, et al. Atlas of Subcellular RNA Localization Revealed by APEX-Seq. *Cell*. 2019;178(2):473-490.e26. doi:10.1016/j.cell.2019.05.027
190. Clemson CM, Hutchinson JN, Sara SA, et al. An Architectural Role for a Nuclear Non-coding RNA: NEAT1 RNA is Essential for the Structure of Paraspeckles. *Mol Cell*. 2009;33(6):717-726. doi:10.1016/j.molcel.2009.01.026
191. Gwizdek C, Ossareh-Nazari B, Brownawell AM, Evers S, Macara IG, Dargemont C. Minihelix-containing RNAs Mediate Exportin-5-dependent Nuclear Export of the Double-stranded RNA-binding Protein ILF3. *J Biol Chem*. 2004;279(2):884-891. doi:10.1074/jbc.M306808200
192. van Heesch S, van Iterson M, Jacobi J, et al. Extensive localization of long noncoding RNAs to the cytosol and mono- and polyribosomal complexes. *Genome Biol*. 2014;15(1):R6. doi:10.1186/gb-2014-15-1-r6
193. Couso J-P, Patraquim P. Classification and function of small open reading frames. *Nat Rev Mol Cell Biol*. 2017;18(9):575-589. doi:10.1038/nrm.2017.58
194. Housman G, Ulitsky I. Methods for distinguishing between protein-coding and long noncoding RNAs and the elusive biological purpose of translation of long noncoding RNAs. *Biochim Biophys Acta BBA - Gene Regul Mech*. 2016;1859(1):31-40. doi:10.1016/j.bbagr.2015.07.017
195. Wei L-H, Guo JU. Coding functions of "noncoding" RNAs. *Science*. 2020;367(6482):1074-1075. doi:10.1126/science.aba6117
196. Chen J, Brunner A-D, Cogan JZ, et al. Pervasive functional translation of noncanonical human open reading frames. *Science*. 2020;367(6482):1140-1146. doi:10.1126/science.aay0262
197. Houseley J, Tollervey D. The Many Pathways of RNA Degradation. *Cell*. 2009;136(4):763-776. doi:10.1016/j.cell.2009.01.019

198. Franks TM, Lykke-Andersen J. TTP and BRF proteins nucleate processing body formation to silence mRNAs with AU-rich elements. *Genes Dev.* 2007;21(6):719-735. doi:10.1101/gad.1494707
199. Izquierdo JM, Majós N, Bonnal S, et al. Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition. *Mol Cell.* 2005;19(4):475-484. doi:10.1016/j.molcel.2005.06.015
200. López de Silanes I, Zhan M, Lal A, Yang X, Gorospe M. Identification of a target RNA motif for RNA-binding protein HuR. *Proc Natl Acad Sci U S A.* 2004;101(9):2987-2992. doi:10.1073/pnas.0306453101
201. Kessler SH, Sachs AB. RNA recognition motif 2 of yeast Pab1p is required for its functional interaction with eukaryotic translation initiation factor 4G. *Mol Cell Biol.* 1998;18(1):51-57. doi:10.1128/mcb.18.1.51
202. Lal A, Mazan-Mamczarz K, Kawai T, Yang X, Martindale JL, Gorospe M. Concurrent versus individual binding of HuR and AUF1 to common labile target mRNAs. *EMBO J.* 2004;23(15):3092-3102. doi:10.1038/sj.emboj.7600305
203. Engreitz JM, Ollikainen N, Guttman M. Long non-coding RNAs: spatial amplifiers that control nuclear structure and gene expression. *Nat Rev Mol Cell Biol.* 2016;17(12):756-770. doi:10.1038/nrm.2016.126
204. Yao R-W, Wang Y, Chen L-L. Cellular functions of long noncoding RNAs. *Nat Cell Biol.* 2019;21(5):542-551. doi:10.1038/s41556-019-0311-8
205. Conrad T, Ørom UA. Cellular Fractionation and Isolation of Chromatin-Associated RNA. In: Ørom UA, ed. *Enhancer RNAs: Methods and Protocols.* Methods in Molecular Biology. Springer; 2017:1-9. doi:10.1007/978-1-4939-4035-6\_1
206. Chu C, Quinn J, Chang HY. Chromatin isolation by RNA purification (ChIRP). *J Vis Exp JoVE.* 2012;(61). doi:10.3791/3912
207. Quinn JJ, Chang HY. In situ dissection of RNA functional subunits by domain-specific chromatin isolation by RNA purification (dChIRP). *Methods Mol Biol Clifton NJ.* 2015;1262:199-213. doi:10.1007/978-1-4939-2253-6\_12
208. Simon MD. Capture Hybridization Analysis of RNA Targets (CHART). *Curr Protoc Mol Biol.* 2013;101(1):21.25.1-21.25.16. doi:10.1002/0471142727.mb2125s101
209. Li X, Zhou B, Chen L, Gou L-T, Li H, Fu X-D. GRID-seq reveals the global RNA–chromatin interactome. *Nat Biotechnol.* 2017;35(10):940-950. doi:10.1038/nbt.3968
210. Salviano-Silva A, Lobo-Alves SC, Almeida RC de, Malheiros D, Petzl-Erler ML. Besides Pathology: Long Non-Coding RNA in Cell and Tissue Homeostasis. *Non-Coding RNA.* 2018;4(1):3. doi:10.3390/ncrna4010003
211. Engreitz J, Lander ES, Guttman M. RNA antisense purification (RAP) for mapping RNA interactions with chromatin. *Methods Mol Biol Clifton NJ.* 2015;1262:183-197. doi:10.1007/978-1-4939-2253-6\_11
212. Sharma E, Sterne-Weiler T, O'Hanlon D, Blencowe BJ. Global Mapping of Human RNA-RNA Interactions. *Mol Cell.* 2016;62(4):618-626. doi:10.1016/j.molcel.2016.04.030
213. Xing Z, Lin C, Yang L. LncRNA Pulldown Combined with Mass Spectrometry to Identify the Novel LncRNA-Associated Proteins. *Methods Mol Biol Clifton NJ.* 2016;1402:1-9. doi:10.1007/978-1-4939-3378-5\_1

214. Bae JW, Kwon SC, Na Y, Kim VN, Kim J-S. Chemical RNA digestion enables robust RNA-binding site mapping at single amino acid resolution. *Nat Struct Mol Biol.* 2020;27(7):678-682. doi:10.1038/s41594-020-0436-2
215. Gagliardi M, Matarazzo MR. RIP: RNA Immunoprecipitation. *Methods Mol Biol Clifton NJ.* 2016;1480:73-86. doi:10.1007/978-1-4939-6380-5\_7
216. Darnell R. CLIP (cross-linking and immunoprecipitation) identification of RNAs bound by a specific protein. *Cold Spring Harb Protoc.* 2012;2012(11):1146-1160. doi:10.1101/pdb.prot072132
217. Wang F, Flanagan J, Su N, et al. RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J Mol Diagn JMD.* 2012;14(1):22-29. doi:10.1016/j.jmoldx.2011.08.002
218. Loda A, Heard E. Xist RNA in action: Past, present, and future. *PLOS Genet.* 2019;15(9):e1008333. doi:10.1371/journal.pgen.1008333
219. Gendrel A-V, Heard E. Fifty years of X-inactivation research. *Development.* 2011;138(23):5049-5055. doi:10.1242/dev.068320
220. Borsani G, Tonlorenzi R, Simmler MC, et al. Characterization of a murine gene expressed from the inactive X chromosome. *Nature.* 1991;351(6324):325-329. doi:10.1038/351325a0
221. Rinn JL, Kertesz M, Wang JK, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell.* 2007;129(7):1311-1323. doi:10.1016/j.cell.2007.05.022
222. Portoso M, Ragazzini R, Brenčić Ž, et al. PRC2 is dispensable for HOTAIR-mediated transcriptional repression. *EMBO J.* 2017;36(8):981-994. doi:10.15252/embj.201695335
223. Blanco MR, Guttman M. Re-evaluating the foundations of lncRNA–Polycomb function. *EMBO J.* 2017;36(8):964-966. doi:10.15252/embj.201796796
224. Wang X, Arai S, Song X, et al. Induced ncRNAs Allosterically Modify RNA Binding Proteins in cis to Inhibit Transcription. *Nature.* 2008;454(7200):126-130. doi:10.1038/nature06992
225. Li Z, Chao T-C, Chang K-Y, et al. The long noncoding RNA THRIL regulates TNF $\alpha$  expression through its interaction with hnRNPL. *Proc Natl Acad Sci U S A.* 2014;111(3):1002-1007. doi:10.1073/pnas.1313768111
226. Hung T, Wang Y, Lin MF, et al. Extensive and coordinated transcription of noncoding RNAs within cell cycle promoters. *Nat Genet.* 2011;43(7):621-629. doi:10.1038/ng.848
227. Yang F, Huo X, Yuan S, et al. Repression of the long noncoding RNA-LET by histone deacetylase 3 contributes to hypoxia-mediated metastasis. *Mol Cell.* 2013;49(6):1083-1096. doi:10.1016/j.molcel.2013.01.010
228. Miao Y, Ajami NE, Huang T-S, et al. Enhancer-associated long non-coding RNA LEENE regulates endothelial nitric oxide synthase and endothelial function. *Nat Commun.* 2018;9(1):292. doi:10.1038/s41467-017-02113-y
229. Marahrens Y, Panning B, Dausman J, Strauss W, Jaenisch R. Xist-deficient mice are defective in dosage compensation but not spermatogenesis. *Genes Dev.* 1997;11(2):156-166. doi:10.1101/gad.11.2.156
230. Paralkar VR, Taborda CC, Huang P, et al. Unlinking an lncRNA from Its Associated cis Element. *Mol Cell.* 2016;62(1):104-110. doi:10.1016/j.molcel.2016.02.029

231. Carpenter S, Atianand M, Aiello D, et al. A long noncoding RNA induced by TLRs mediates both activation and repression of immune response genes. *Science*. 2013;341(6147):789-792. doi:10.1126/science.1240925
232. Elling R, Robinson EK, Shapleigh B, et al. Genetic Models Reveal cis and trans Immune-Regulatory Activities for lincRNA-Cox2. *Cell Rep*. 2018;25(6):1511-1524.e6. doi:10.1016/j.celrep.2018.10.027
233. Redis RS, Vela LE, Lu W, et al. Allele-Specific Reprogramming of Cancer Metabolism by the Long Non-coding RNA CCAT2. *Mol Cell*. 2016;61(4):520-534. doi:10.1016/j.molcel.2016.01.015
234. Ashwal-Fluss R, Meyer M, Pamudurti NR, et al. circRNA Biogenesis Competes with Pre-mRNA Splicing. *Mol Cell*. 2014;56(1):55-66. doi:10.1016/j.molcel.2014.08.019
235. West JA, Davis CP, Sunwoo H, et al. The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Mol Cell*. 2014;55(5):791-802. doi:10.1016/j.molcel.2014.07.012
236. Imamura K, Imamachi N, Akizuki G, et al. Long noncoding RNA NEAT1-dependent SFPQ relocation from promoter region to paraspeckle mediates IL8 expression upon immune stimuli. *Mol Cell*. 2014;53(3):393-406. doi:10.1016/j.molcel.2014.01.009
237. Hall LL, Smith KP, Byron M, Lawrence JB. THE MOLECULAR ANATOMY OF A SPECKLE. *Anat Rec A Discov Mol Cell Evol Biol*. 2006;288(7):664-675. doi:10.1002/ar.a.20336
238. Tripathi V, Ellis JD, Shen Z, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell*. 2010;39(6):925-938. doi:10.1016/j.molcel.2010.08.011
239. Bernard D, Prasanth KV, Tripathi V, et al. A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. *EMBO J*. 2010;29(18):3082-3093. doi:10.1038/emboj.2010.199
240. Kleaveland B, Shi CY, Stefano J, Bartel DP. A Network of Noncoding Regulatory RNAs Acts in the Mammalian Brain. *Cell*. 2018;174(2):350-362.e17. doi:10.1016/j.cell.2018.05.022
241. Piwecka M, Glažar P, Hernandez-Miranda LR, et al. Loss of a mammalian circular RNA locus causes miRNA deregulation and affects brain function. *Science*. 2017;357(6357). doi:10.1126/science.aam8526
242. Hansen TB, Kjems J, Damgaard CK. Circular RNA and miR-7 in Cancer. *Cancer Res*. Published online 6 September 2013. doi:10.1158/0008-5472.CAN-13-1568
243. Milligan MJ, Lipovich L. Pseudogene-derived lincRNAs: emerging regulators of gene expression. *Front Genet*. 2015;5. doi:10.3389/fgene.2014.00476
244. Tam OH, Aravin AA, Stein P, et al. Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature*. 2008;453(7194):534-538. doi:10.1038/nature06904
245. Watanabe T, Totoki Y, Toyoda A, et al. Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature*. 2008;453(7194):539-543. doi:10.1038/nature06908
246. Denzler R, Agarwal V, Stefano J, Bartel DP, Stoffel M. Assessing the ceRNA Hypothesis with Quantitative Measurements of miRNA and Target Abundance. *Mol Cell*. 2014;54(5):766-776. doi:10.1016/j.molcel.2014.03.045

247. Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature*. 2010;465(7301):1033-1038. doi:10.1038/nature09144
248. Kretz M, Siprashvili Z, Chu C, et al. Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature*. 2013;493(7431):231-235. doi:10.1038/nature11661
249. Gong C, Maquat LE. lncRNAs transactivate Staufen1-mediated mRNA decay by duplexing with 3'UTRs via Alu elements. *Nature*. 2011;470(7333):284-288. doi:10.1038/nature09701
250. Gong C, Li Z, Ramanujan K, et al. A Long Non-coding RNA, LncMyoD, Regulates Skeletal Muscle Differentiation by Blocking IMP2-Mediated mRNA Translation. *Dev Cell*. 2015;34(2):181-191. doi:10.1016/j.devcel.2015.05.009
251. Pang W-J, Lin L-G, Xiong Y, et al. Knockdown of PU.1 AS lncRNA inhibits adipogenesis through enhancing PU.1 mRNA translation. *J Cell Biochem*. 2013;114(11):2500-2512. doi:10.1002/jcb.24595
252. Fasolo F, Patrucco L, Volpe M, et al. The RNA-binding protein ILF3 binds to transposable element sequences in SINEUP lncRNAs. *FASEB J*. 2019;33(12):13572-13589. doi:10.1096/fj.201901618RR
253. Lyu Q, Xu S, Lyu Y, et al. SENCER stabilizes vascular endothelial cell adherens junctions through interaction with CKAP4. *Proc Natl Acad Sci U S A*. 2019;116(2):546-555. doi:10.1073/pnas.1810729116
254. Kopp F, Elguindy MM, Yalvac ME, et al. PUMILIO hyperactivity drives premature aging of Norad-deficient mice. Ponting CP, Manley JL, Ponting CP, Ulitsky I, eds. *eLife*. 2019;8:e42650. doi:10.7554/eLife.42650
255. Yu T, Zhao Y, Hu Z, et al. MetaLnc9 Facilitates Lung Cancer Metastasis via a PGK1-Activated AKT/mTOR Pathway. *Cancer Res*. 2017;77(21):5782-5794. doi:10.1158/0008-5472.CAN-17-0671
256. Wang P, Xue Y, Han Y, et al. The STAT3-Binding Long Noncoding RNA Inc-DC Controls Human Dendritic Cell Differentiation. *Science*. 2014;344(6181):310-313. doi:10.1126/science.1251456
257. Zhang W, Guo Q, Liu G, et al. NKILA represses nasopharyngeal carcinoma carcinogenesis and metastasis by NF- $\kappa$ B pathway inhibition. *PLoS Genet*. 2019;15(8). doi:10.1371/journal.pgen.1008325
258. Anderson P, Kedersha N. Stress granules: the Tao of RNA triage. *Trends Biochem Sci*. 2008;33(3):141-150. doi:10.1016/j.tibs.2007.12.003
259. Hubstenberger A, Courel M, Bénard M, et al. P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. *Mol Cell*. 2017;68(1):144-157.e5. doi:10.1016/j.molcel.2017.09.003
260. Pitchiaya S, Mourao MDA, Jalihal AP, et al. Dynamic Recruitment of Single RNAs to Processing Bodies Depends on RNA Functionality. *Mol Cell*. 2019;74(3):521-533.e6. doi:10.1016/j.molcel.2019.03.001
261. TEIXEIRA D, SHETH U, VALENCIA-SANCHEZ MA, BRENGUES M, PARKER R. Processing bodies require RNA for assembly and contain nontranslating mRNAs. *RNA*. 2005;11(4):371-382. doi:10.1261/rna.7258505

262. Hosono Y, Niknafs YS, Prensner JR, et al. Oncogenic Role of THOR, a Conserved Cancer/Testis Long Noncoding RNA. *Cell*. 2017;171(7):1559-1572.e20. doi:10.1016/j.cell.2017.11.040
263. Kedersha N, Stoecklin G, Ayodele M, et al. Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J Cell Biol*. 2005;169(6):871-884. doi:10.1083/jcb.200502088
264. Lee S, Kopp F, Chang T-C, et al. Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. *Cell*. 2016;164(1):69-80. doi:10.1016/j.cell.2015.12.017
265. Namkoong S, Ho A, Woo YM, Kwak H, Lee JH. Systematic Characterization of Stress-Induced RNA Granulation. *Mol Cell*. 2018;70(1):175-187.e8. doi:10.1016/j.molcel.2018.02.025
266. Anderson P, Kedersha N. RNA granules. *J Cell Biol*. 2006;172(6):803-808. doi:10.1083/jcb.200512082
267. Valadi H, Ekström K, Bossios A, Sjöstrand M, Lee JJ, Lötvall JO. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat Cell Biol*. 2007;9(6):654-659. doi:10.1038/ncb1596
268. Skog J, Würdinger T, Rijn S van, et al. Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. *Nat Cell Biol*. 2008;10(12):1470-1476. doi:10.1038/ncb1800
269. Amorim MG, Valieris R, Drummond RD, et al. A total transcriptome profiling method for plasma-derived extracellular vesicles: applications for liquid biopsies. *Sci Rep*. 2017;7(1):1-11. doi:10.1038/s41598-017-14264-5
270. Gezer U, Özgür E, Cetinkaya M, Isin M, Dalay N. Long non-coding RNAs with low expression levels in cells are enriched in secreted exosomes. *Cell Biol Int*. 2014;38(9):1076-1079. doi:10.1002/cbin.10301
271. Hewson C, Capraro D, Burdach J, Whitaker N, Morris KV. Extracellular vesicle associated long non-coding RNAs functionally enhance cell viability. *Non-Coding RNA Res*. 2016;1(1):3-11. doi:10.1016/j.ncrna.2016.06.001
272. Lee JT, Strauss WM, Dausman JA, Jaenisch R. A 450 kb Transgene Displays Properties of the Mammalian X-Inactivation Center. *Cell*. 1996;86(1):83-94. doi:10.1016/S0092-8674(00)80079-3
273. Yang L, Kirby JE, Sunwoo H, Lee JT. Female mice lacking Xist RNA show partial dosage compensation and survive to term. *Genes Dev*. 2016;30(15):1747-1760. doi:10.1101/gad.281162.116
274. Gabory A, Jammes H, Dandolo L. The H19 locus: Role of an imprinted non-coding RNA in growth and development. *BioEssays*. 2010;32(6):473-480. doi:10.1002/bies.200900170
275. Wang KC, Yang YW, Liu B, et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature*. 2011;472(7341):120-124. doi:10.1038/nature09819
276. Sauvageau M, Goff LA, Lodato S, et al. Multiple knockout mouse models reveal lincRNAs are required for life and brain development. Reinberg D, ed. *eLife*. 2013;2:e01749. doi:10.7554/eLife.01749
277. Lai K-MV, Gong G, Atanasio A, et al. Diverse Phenotypes and Specific Transcription Patterns in Twenty Mouse Lines with Ablated LincRNAs. *PLoS ONE*. 2015;10(4). doi:10.1371/journal.pone.0125522



278. Nakagawa S, Naganuma T, Shioi G, Hirose T. Paraspeckles are subpopulation-specific nuclear bodies that are not essential in mice. *J Cell Biol.* 2011;193(1):31-39. doi:10.1083/jcb.201011110
279. Perry RB-T, Ulitsky I. The functions of long noncoding RNAs in development and stem cells. *Development.* 2016;143(21):3882-3894. doi:10.1242/dev.140962
280. Lin N, Chang K-Y, Li Z, et al. An evolutionarily conserved long noncoding RNA TUNA controls pluripotency and neural lineage commitment. *Mol Cell.* 2014;53(6):1005-1019. doi:10.1016/j.molcel.2014.01.021
281. Ng S-Y, Bogu GK, Soh BS, Stanton LW. The long noncoding RNA RMST interacts with SOX2 to regulate neurogenesis. *Mol Cell.* 2013;51(3):349-359. doi:10.1016/j.molcel.2013.07.017
282. Kretz M, Webster DE, Flockhart RJ, et al. Suppression of progenitor differentiation requires the long noncoding RNA ANCR. *Genes Dev.* 2012;26(4):338-343. doi:10.1101/gad.182121.111
283. Luo S, Lu JY, Liu L, et al. Divergent lncRNAs Regulate Gene Expression and Lineage Differentiation in Pluripotent Cells. *Cell Stem Cell.* 2016;18(5):637-652. doi:10.1016/j.stem.2016.01.024
284. Jiang W, Liu Y, Liu R, Zhang K, Zhang Y. The lncRNA DEANR1 facilitates human endoderm differentiation by activating FOXA2 expression. *Cell Rep.* 2015;11(1):137-148. doi:10.1016/j.celrep.2015.03.008
285. Herriges MJ, Swarr DT, Morley MP, et al. Long noncoding RNAs are spatially correlated with transcription factors and regulate lung development. *Genes Dev.* 2014;28(12):1363-1379. doi:10.1101/gad.238782.114
286. Deng C, Li Y, Zhou L, et al. HoxBln1 RNA Recruits Set1/MLL Complexes to Activate Hox Gene Expression Patterns and Mesoderm Lineage Development. *Cell Rep.* 2016;14(1):103-114. doi:10.1016/j.celrep.2015.12.007
287. Klattenhoff CA, Scheuermann JC, Surface LE, et al. Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. *Cell.* 2013;152(3):570-583. doi:10.1016/j.cell.2013.01.003
288. Grote P, Wittler L, Währisch S, et al. The tissue-specific lncRNA Fendrr is an essential regulator of heart and body wall development in the mouse. *Dev Cell.* 2013;24(2):206-214. doi:10.1016/j.devcel.2012.12.012
289. Boulberdaa M, Scott E, Ballantyne M, et al. A Role for the Long Noncoding RNA SENCRC in Commitment and Function of Endothelial Cells. *Mol Ther.* 2016;24(5):978-990. doi:10.1038/mt.2016.41
290. Yang L, Li Y, Gong R, et al. The Long Non-coding RNA-ORLNC1 Regulates Bone Mass by Directing Mesenchymal Stem Cell Fate. *Mol Ther.* 2019;27(2):394-410. doi:10.1016/j.ymthe.2018.11.019
291. Sun L, Goff LA, Trapnell C, et al. Long noncoding RNAs regulate adipogenesis. *Proc Natl Acad Sci.* 2013;110(9):3387-3392. doi:10.1073/pnas.1222643110
292. Mowel WK, McCright SJ, Kotzin JJ, et al. Group 1 Innate lymphoid cell lineage identity is determined by a cis-regulatory element marked by a long non-coding RNA. *Immunity.* 2017;47(3):435-449.e8. doi:10.1016/j.immuni.2017.08.012
293. Kurian L, Aguirre A, Sancho-Martinez I, et al. Identification of Novel Long Noncoding RNAs Underlying Vertebrate Cardiovascular Development. *Circulation.* 2015;131(14):1278-1290. doi:10.1161/CIRCULATIONAHA.114.013303

294. Stojic L, Lun ATL, Mangei J, et al. Specificity of RNAi, LNA and CRISPRi as loss-of-function methods in transcriptional analysis. *Nucleic Acids Res.* 2018;46(12):5950-5966. doi:10.1093/nar/gky437
295. LIU T, HUANG Y, CHEN J, et al. Attenuated ability of BACE1 to cleave the amyloid precursor protein via silencing long noncoding RNA BACE1-AS expression. *Mol Med Rep.* 2014;10(3):1275-1281. doi:10.3892/mmr.2014.2351
296. Aryal B, Suárez Y. Non-coding RNA regulation of endothelial and macrophage functions during atherosclerosis. *Vascul Pharmacol.* 2019;114:64-75. doi:10.1016/j.vph.2018.03.001
297. Michelini F, Pitchiaya S, Vitelli V, et al. Damage-induced lncRNAs control the DNA damage response through interaction with DDRNAs at individual double-strand breaks. *Nat Cell Biol.* 2017;19(12):1400-1411. doi:10.1038/ncb3643
298. Aguado J, Sola-Carvajal A, Cancila V, et al. Inhibition of DNA damage response at telomeres improves the detrimental phenotypes of Hutchinson–Gilford Progeria Syndrome. *Nat Commun.* 2019;10(1):1-11. doi:10.1038/s41467-019-13018-3
299. Yan X, Hu Z, Feng Y, et al. Comprehensive Genomic Characterization of Long Non-Coding RNAs across Human Cancers. *Cancer Cell.* 2015;28(4):529-540. doi:10.1016/j.ccell.2015.09.006
300. Schmitt AM, Chang HY. Long Noncoding RNAs in Cancer Pathways. *Cancer Cell.* 2016;29(4):452-463. doi:10.1016/j.ccell.2016.03.010
301. Rippe K, Luke B. TERRA and the state of the telomere. *Nat Struct Mol Biol.* 2015;22(11):853-858. doi:10.1038/nsmb.3078
302. Deng Z, Wang Z, Stong N, et al. A role for CTCF and cohesin in subtelomere chromatin organization, TERRA transcription, and telomere end protection. *EMBO J.* 2012;31(21):4165-4178. doi:10.1038/emboj.2012.266
303. Wang Y, He L, Du Y, et al. The Long Noncoding RNA lncTCF7 Promotes Self-Renewal of Human Liver Cancer Stem Cells through Activation of Wnt Signaling. *Cell Stem Cell.* 2015;16(4):413-425. doi:10.1016/j.stem.2015.03.003
304. Trimarchi T, Bilal E, Ntziachristos P, et al. Genome-wide mapping and characterization of Notch-regulated long noncoding RNAs in acute leukemia. *Cell.* 2014;158(3):593-606. doi:10.1016/j.cell.2014.05.049
305. Hung C-L, Wang L-Y, Yu Y-L, et al. A long noncoding RNA connects c-Myc to tumor metabolism. *Proc Natl Acad Sci U S A.* 2014;111(52):18697-18702. doi:10.1073/pnas.1415669112
306. Kajino T, Shimamura T, Gong S, et al. Divergent lncRNA MYMLR regulates MYC by eliciting DNA looping and promoter-enhancer interaction. *EMBO J.* 2019;38(17):e98441. doi:10.15252/emboj.201798441
307. Kong R, Zhang E, Yin D, et al. Long noncoding RNA PVT1 indicates a poor prognosis of gastric cancer and promotes cell proliferation through epigenetically regulating p15 and p16. *Mol Cancer.* 2015;14:82. doi:10.1186/s12943-015-0355-8
308. Tseng Y-Y, Moriarity BS, Gong W, et al. PVT1 dependence in cancer with MYC copy-number increase. *Nature.* 2014;512(7512):82-86. doi:10.1038/nature13311
309. Aguilo F, Zhou M-M, Walsh MJ. Long Noncoding RNA, Polycomb, and the Ghosts Haunting INK4b-ARF-INK4a Expression. *Cancer Res.* 2011;71(16):5365-5369. doi:10.1158/0008-5472.CAN-10-4379

310. Lan W-G, Xu D-H, Xu C, et al. Silencing of long non-coding RNA ANRIL inhibits the development of multidrug resistance in gastric cancer cells. *Oncol Rep.* 2016;36(1):263-270. doi:10.3892/or.2016.4771
311. Sun L-Y, Li X-J, Sun Y-M, et al. LncRNA ANRIL regulates AML development through modulating the glucose metabolism pathway of AdipoR1/AMPK/SIRT1. *Mol Cancer.* 2018;17(1):127. doi:10.1186/s12943-018-0879-9
312. Kotake Y, Nakagawa T, Kitagawa K, et al. Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15 INK4B tumor suppressor gene. *Oncogene.* 2011;30(16):1956-1962. doi:10.1038/onc.2010.568
313. Richards EJ, Zhang G, Li Z-P, et al. Long non-coding RNAs (LncRNA) regulated by transforming growth factor (TGF)  $\beta$ : LncRNA-hit-mediated TGF $\beta$ -induced epithelial to mesenchymal transition in mammary epithelia. *J Biol Chem.* 2015;290(11):6857-6867. doi:10.1074/jbc.M114.610915
314. Papoutsoglou P, Tsubakihara Y, Caja L, et al. The TGFB2-AS1 lncRNA Regulates TGF- $\beta$  Signaling by Modulating Corepressor Activity. *Cell Rep.* 2019;28(12):3182-3198.e11. doi:10.1016/j.celrep.2019.08.028
315. Sun M, Liu X-H, Wang K-M, et al. Downregulation of BRAF activated non-coding RNA is associated with poor prognosis for non-small cell lung cancer and promotes metastasis by affecting epithelial-mesenchymal transition. *Mol Cancer.* 2014;13:68. doi:10.1186/1476-4598-13-68
316. Wang D, Wang D, Wang N, Long Z, Ren X. Long Non-Coding RNA BANCR Promotes Endometrial Cancer Cell Proliferation and Invasion by Regulating MMP2 and MMP1 via ERK/MAPK Signaling Pathway. *Cell Physiol Biochem Int J Exp Cell Physiol Biochem Pharmacol.* 2016;40(3-4):644-656. doi:10.1159/000452577
317. Alves CP, Fonseca AS, Muys BR, et al. Brief Report: The lincRNA Hotair Is Required for Epithelial-to-Mesenchymal Transition and Stemness Maintenance of Cancer Cell Lines. *STEM CELLS.* 2013;31(12):2827-2832. doi:10.1002/stem.1547
318. Yu L, Fang F, Lu S, Li X, Yang Y, Wang Z. lncRNA-HIT promotes cell proliferation of non-small cell lung cancer by association with E2F1. *Cancer Gene Ther.* 2017;24(5):221-226. doi:10.1038/cgt.2017.10
319. Yuan J, Yang F, Wang F, et al. A Long Noncoding RNA Activated by TGF- $\beta$  Promotes the Invasion-Metastasis Cascade in Hepatocellular Carcinoma. *Cancer Cell.* 2014;25(5):666-681. doi:10.1016/j.ccr.2014.03.010
320. Ji P, Diederichs S, Wang W, et al. MALAT-1, a novel noncoding RNA, and thymosin  $\beta$  4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene.* 2003;22(39):8031-8041. doi:10.1038/sj.onc.1206928
321. Gutschner T, Hämmerle M, Eißmann M, et al. The non-coding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res.* 2013;73(3):1180-1189. doi:10.1158/0008-5472.CAN-12-2850
322. Gupta RA, Shah N, Wang KC, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature.* 2010;464(7291):1071-1076. doi:10.1038/nature08975
323. Li W, Kang Y. A New Lnc in Metastasis: Long Non-Coding RNA Mediates the Pro-metastatic Functions of TGF- $\beta$ . *Cancer Cell.* 2014;25(5):557-559. doi:10.1016/j.ccr.2014.04.014

324. Fan Y, Shen B, Tan M, et al. Long non-coding RNA UCA1 increases chemoresistance of bladder cancer cells by regulating Wnt signaling. *FEBS J.* 2014;281(7):1750-1758. doi:10.1111/febs.12737
325. Li Z, Zhao X, Zhou Y, et al. The long non-coding RNA HOTTIP promotes progression and gemcitabine resistance by regulating HOXA13 in pancreatic cancer. *J Transl Med.* 2015;13(1):84. doi:10.1186/s12967-015-0442-z
326. Lu C, Wei Y, Wang X, et al. DNA-methylation-mediated activating of lncRNA SNHG12 promotes temozolomide resistance in glioblastoma. *Mol Cancer.* 2020;19. doi:10.1186/s12943-020-1137-5
327. Niu Y, Ma F, Huang W, et al. Long non-coding RNA TUG1 is involved in cell growth and chemoresistance of small cell lung cancer by regulating LIMK2b via EZH2. *Mol Cancer.* 2017;16(1):5. doi:10.1186/s12943-016-0575-6
328. Lakin ND, Jackson SP. Regulation of p53 in response to DNA damage. *Oncogene.* 1999;18(53):7644-7655. doi:10.1038/sj.onc.1203015
329. Huarte M, Guttman M, Feldser D, et al. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell.* 2010;142(3):409-419. doi:10.1016/j.cell.2010.06.040
330. Dimitrova N, Zamudio JR, Jong RM, et al. LincRNA-p21 Activates p21 In cis to Promote Polycomb Target Gene Expression and to Enforce the G1/S Checkpoint. *Mol Cell.* 2014;54(5):777-790. doi:10.1016/j.molcel.2014.04.025
331. Léveillé N, Melo CA, Rooijers K, et al. Genome-wide profiling of p53-regulated enhancer RNAs uncovers a subset of enhancers controlled by a lncRNA. *Nat Commun.* 2015;6(1):1-12. doi:10.1038/ncomms7520
332. Lu K, Li W, Liu X, et al. Long non-coding RNA MEG3 inhibits NSCLC cells proliferation and induces apoptosis by affecting p53 expression. *BMC Cancer.* 2013;13:461. doi:10.1186/1471-2407-13-461
333. Marín-Béjar O, Marchese FP, Athie A, et al. Pint lincRNA connects the p53 pathway with epigenetic silencing by the Polycomb repressive complex 2. *Genome Biol.* 2013;14(9):R104. doi:10.1186/gb-2013-14-9-r104
334. Pickard MR, Williams GT. Molecular and Cellular Mechanisms of Action of Tumour Suppressor GAS5 lncRNA. *Genes.* 2015;6(3):484-499. doi:10.3390/genes6030484
335. Chang Y-N, Zhang K, Hu Z-M, et al. Hypoxia-regulated lncRNAs in cancer. *Gene.* 2016;575(1):1-8. doi:10.1016/j.gene.2015.08.049
336. Choudhry H, Albukhari A, Morotti M, et al. Tumor hypoxia induces nuclear paraspeckle formation through HIF-2 $\alpha$  dependent transcriptional activation of NEAT1 leading to cancer cell survival. *Oncogene.* 2015;34(34):4482-4490. doi:10.1038/onc.2014.378
337. Xue M, Li X, Li Z, Chen W. Urothelial carcinoma associated 1 is a hypoxia-inducible factor-1 $\alpha$ -targeted long noncoding RNA that enhances hypoxic bladder cancer cell proliferation, migration, and invasion. *Tumour Biol J Int Soc Oncodevelopmental Biol Med.* 2014;35(7):6901-6912. doi:10.1007/s13277-014-1925-x
338. Wu W, Hu Q, Nie E, et al. Hypoxia induces H19 expression through direct and indirect Hif-1 $\alpha$  activity, promoting oncogenic effects in glioblastoma. *Sci Rep.* 2017;7:45029. doi:10.1038/srep45029

339. Ferdin J, Nishida N, Wu X, et al. HINCUTs in cancer: hypoxia-induced noncoding ultraconserved transcripts. *Cell Death Differ.* 2013;20(12):1675-1687. doi:10.1038/cdd.2013.119
340. Gómez-Maldonado L, Tiana M, Roche O, et al. EFNA3 long noncoding RNAs induced by hypoxia promote metastatic dissemination. *Oncogene.* 2015;34(20):2609-2620. doi:10.1038/onc.2014.200
341. Lin A, Li C, Xing Z, et al. The LINK-A lncRNA activates normoxic HIF1 $\alpha$  signalling in triple-negative breast cancer. *Nat Cell Biol.* 2016;18(2):213-224. doi:10.1038/ncb3295
342. Takahashi K, Yan IK, Haga H, Patel T. Modulation of hypoxia-signaling pathways by extracellular linc-RoR. *J Cell Sci.* 2014;127(7):1585-1594. doi:10.1242/jcs.141069
343. Yang F, Zhang H, Mei Y, Wu M. Reciprocal Regulation of HIF-1 $\alpha$  and LincRNA-p21 Modulates the Warburg Effect. *Mol Cell.* 2014;53(1):88-100. doi:10.1016/j.molcel.2013.11.004
344. Wang X, Li L, Zhao K, et al. A novel LncRNA HITT forms a regulatory loop with HIF-1 $\alpha$  to modulate angiogenesis and tumor growth. *Cell Death Differ.* 2020;27(4):1431-1446. doi:10.1038/s41418-019-0449-8
345. Gu L-Q, Xing X-L, Cai H, et al. Long non-coding RNA DILC suppresses cell proliferation and metastasis in colorectal cancer. *Gene.* 2018;666:18-26. doi:10.1016/j.gene.2018.03.100
346. Zhang J, Li Z, Liu L, et al. Long noncoding RNA TSLNC8 is a tumor suppressor that inactivates the interleukin-6/STAT3 signaling pathway. *Hepatol Baltim Md.* 2018;67(1):171-187. doi:10.1002/hep.29405
347. Liu B, Sun L, Liu Q, et al. A cytoplasmic NF- $\kappa$ B interacting long noncoding RNA blocks I $\kappa$ B phosphorylation and suppresses breast cancer metastasis. *Cancer Cell.* 2015;27(3):370-381. doi:10.1016/j.ccell.2015.02.004
348. Rapicavoli NA, Qu K, Zhang J, Mikhail M, Laberge R-M, Chang HY. A mammalian pseudogene lncRNA at the interface of inflammation and anti-inflammatory therapeutics. *Gingeras T, ed. eLife.* 2013;2:e00762. doi:10.7554/eLife.00762
349. Wu K, Zhao Z, Liu K, Zhang J, Li G, Wang L. Long noncoding RNA lnc-sox5 modulates CRC tumorigenesis by unbalancing tumor microenvironment. *Cell Cycle.* 2017;16(13):1295-1301. doi:10.1080/15384101.2017.1317416
350. Wang Y, Yang X, Sun X, et al. Bone marrow infiltrated lnc-INSR induced suppressive immune microenvironment in pediatric acute lymphoblastic leukemia. *Cell Death Dis.* 2018;9(10). doi:10.1038/s41419-018-1078-8
351. Kathuria H, Millien G, McNally L, et al. NKX2-1-AS1 negatively regulates CD274/PD-L1, cell-cell interaction genes, and limits human lung carcinoma cell migration. *Sci Rep.* 2018;8. doi:10.1038/s41598-018-32793-5
352. Lu Z, Xiao Z, Liu F, et al. Long non-coding RNA HULC promotes tumor angiogenesis in liver cancer by up-regulating sphingosine kinase 1 (SPHK1). *Oncotarget.* 2015;7(1):241-254. doi:10.18632/oncotarget.6280
353. Leisegang MS, Fork C, Josipovic I, et al. Long Noncoding RNA MANTIS Facilitates Endothelial Angiogenic Function. *Circulation.* 2017;136(1):65-79. doi:10.1161/CIRCULATIONAHA.116.026991
354. Man HSJ, Sukumar AN, Lam GC, et al. Angiogenic patterning by STEEL, an endothelial-enriched long noncoding RNA. *Proc Natl Acad Sci U S A.* 2018;115(10):2401-2406. doi:10.1073/pnas.1715182115

355. Zhou Q, Yu B, Anderson C, et al. LncEGFL7OS regulates human angiogenesis by interacting with MAX at the EGFL7/miR-126 locus. Fish J, Stainier DY, eds. *eLife*. 2019;8:e40470. doi:10.7554/eLife.40470
356. Li K, Blum Y, Verma A, et al. A noncoding antisense RNA in tie-1 locus regulates tie-1 function in vivo. *Blood*. 2010;115(1):133-139. doi:10.1182/blood-2009-09-242180
357. Yang J, Lin J, Liu T, et al. Analysis of lncRNA expression profiles in non-small cell lung cancers (NSCLC) and their clinical subtypes. *Lung Cancer*. 2014;85(2):110-115. doi:10.1016/j.lungcan.2014.05.011
358. Hosono Y, Niknafs YS, Prensner JR, et al. Oncogenic Role of THOR, a Conserved Cancer/Testis Long Non-coding RNA. *Cell*. 2017;171(7):1559-1572.e20. doi:10.1016/j.cell.2017.11.040
359. Qiu M, Xu Y, Wang J, et al. A novel lncRNA, LUADT1, promotes lung adenocarcinoma proliferation via the epigenetic suppression of p27. *Cell Death Dis*. 2015;6(8):e1858. doi:10.1038/cddis.2015.203
360. Zhou J, Xiao H, Yang X, et al. Long noncoding RNA CASC9.5 promotes the proliferation and metastasis of lung adenocarcinoma. *Sci Rep*. 2018;8(1):37. doi:10.1038/s41598-017-18280-3
361. Hou Z, Zhao W, Zhou J, et al. A long noncoding RNA Sox2ot regulates lung cancer cell proliferation and is a prognostic indicator of poor survival. *Int J Biochem Cell Biol*. 2014;53:380-388. doi:10.1016/j.biocel.2014.06.004
362. Wang Y, Wu N, Luo X, Zhang X, Liao Q, Wang J. SOX2OT, a novel tumor-related long non-coding RNA. *Biomed Pharmacother*. 2020;123:109725. doi:10.1016/j.biopha.2019.109725
363. Park S-M, Choi E-Y, Bae D-H, Sohn HA, Kim S-Y, Kim Y-J. The lncRNA EPEL Promotes Lung Cancer Cell Proliferation Through E2F Target Activation. *Cell Physiol Biochem*. 2018;45(3):1270-1283. doi:10.1159/000487460
364. Chen Z, Li J-L, Lin S, et al. cAMP/CREB-regulated LINC00473 marks LKB1-inactivated lung cancer and mediates tumor growth. *J Clin Invest*. 126(6):2267-2279. doi:10.1172/JCI85250
365. Yang J, Qiu Q, Qian X, et al. Long noncoding RNA LCAT1 functions as a ceRNA to regulate RAC1 function by sponging miR-4715-5p in lung cancer. *Mol Cancer*. 2019;18(1):171. doi:10.1186/s12943-019-1107-y
366. Han L, Zhang E -b, Yin D -d, et al. Low expression of long noncoding RNA PANDAR predicts a poor prognosis of non-small cell lung cancer and affects cell apoptosis by regulating Bcl-2. *Cell Death Dis*. 2015;6:e1665. doi:10.1038/cddis.2015.30
367. Seiler J, Breinig M, Caudron-Herger M, Polycarpou-Schwarz M, Boutros M, Diederichs S. The lncRNA VELUCT strongly regulates viability of lung cancer cells despite its extremely low abundance. *Nucleic Acids Res*. 2017;45(9):5458-5469. doi:10.1093/nar/gkx076
368. Kumar P, Khadiraikar S, Shukla SK. PILAR1, a novel prognostic lncRNA, reveals the presence of a unique subtype of lung adenocarcinoma patients with KEAP1 mutations. *Gene*. 2019;691:167-175. doi:10.1016/j.gene.2018.12.060
369. Athie A, Marchese FP, González J, et al. Analysis of copy number alterations reveals the lncRNA ALAL-1 as a regulator of lung cancer immune evasion. *J Cell Biol*. 2020;219(9). doi:10.1083/jcb.201908078

370. Jia X, Wang Z, Qiu L, et al. Upregulation of LncRNA-HIT promotes migration and invasion of non-small cell lung cancer cells by association with ZEB1. *Cancer Med.* 2016;5(12):3555-3563. doi:10.1002/cam4.948
371. Sakai S, Ohhata T, Kitagawa K, et al. Long Noncoding RNA ELIT-1 Acts as a Smad3 Cofactor to Facilitate TGF $\beta$ /Smad Signaling and Promote Epithelial–Mesenchymal Transition. *Cancer Res.* 2019;79(11):2821-2838. doi:10.1158/0008-5472.CAN-18-3210
372. Zhao W, Zhang L-N, Wang X-L, Zhang J, Yu H-X. Long noncoding RNA NSCLCAT1 increases non–small cell lung cancer cell invasion and migration through the Hippo signaling pathway by interacting with CDH1. *FASEB J.* 2019;33(1):1151-1166. doi:10.1096/fj.201800408R
373. Yuan S, Liu Q, Hu Z, et al. Long non-coding RNA MUC5B-AS1 promotes metastasis through mutually regulating MUC5B expression in lung adenocarcinoma. *Cell Death Dis.* 2018;9(5). doi:10.1038/s41419-018-0472-6
374. Wang C, Gu Y, Zhang E, et al. A cancer-testis non-coding RNA LIN28B-AS1 activates driver gene LIN28B by interacting with IGF2BP1 in lung adenocarcinoma. *Oncogene.* 2019;38(10):1611-1624. doi:10.1038/s41388-018-0548-x
375. Li P, Zhang G, Li J, et al. Long Noncoding RNA RGMB-AS1 Indicates a Poor Prognosis and Modulates Cell Proliferation, Migration and Invasion in Lung Adenocarcinoma. *PLoS ONE.* 2016;11(3). doi:10.1371/journal.pone.0150790
376. Grelet S, Link LA, Howley B, et al. A regulated PNUTS mRNA to lncRNA splice switch mediates EMT and tumor progression. *Nat Cell Biol.* 2017;19(9):1105-1115. doi:10.1038/ncb3595
377. Guan H, Zhu T, Wu S, et al. Long noncoding RNA LINC00673-v4 promotes aggressiveness of lung adenocarcinoma via activating WNT/ $\beta$ -catenin signaling. *Proc Natl Acad Sci U S A.* 2019;116(28):14019-14028. doi:10.1073/pnas.1900997116
378. Liu D, Shi X. Long non-coding RNA NKILA inhibits proliferation and migration of lung cancer via IL-11/STAT3 signaling. *Int J Clin Exp Pathol.* 2019;12(7):2595-2603.
379. Gao Y-P, Li Y, Li H-J, Zhao B. LncRNA NBR2 inhibits EMT progression by regulating Notch1 pathway in NSCLC. *Eur Rev Med Pharmacol Sci.* 2019;23(18):7950-7958. doi:10.26355/eurrev\_201909\_19011
380. Cao Y, Lin M, Bu Y, et al. p53-inducible long non-coding RNA PICART1 mediates cancer cell proliferation and migration. *Int J Oncol.* 2017;50(5):1671-1682. doi:10.3892/ijo.2017.3918
381. Zhao JM, Cheng W, He XG, Liu YL, Wang FF, Gao YF. Long non-coding RNA PICART1 suppresses proliferation and promotes apoptosis in lung cancer cells by inhibiting JAK2/STAT3 signaling. *Neoplasma.* 2018;65(5):779-789. doi:10.4149/neo\_2018\_171130N778
382. Olivero CE, Martínez-Terroba E, Zimmer J, et al. p53 Activates the Long Noncoding RNA Pvt1b to Inhibit Myc and Suppress Tumorigenesis. *Mol Cell.* 2020;77(4):761-774.e8. doi:10.1016/j.molcel.2019.12.014
383. Arab K, Park YJ, Lindroth AM, et al. Long noncoding RNA TARID directs demethylation and activation of the tumor suppressor TCF21 via GADD45A. *Mol Cell.* 2014;55(4):604-614. doi:10.1016/j.molcel.2014.06.031
384. Denzler R, Agarwal V, Stefano J, Bartel DP, Stoffel M. Assessing the ceRNA Hypothesis with Quantitative Measurements of miRNA and Target Abundance. *Mol Cell.* 2014;54(5):766-776. doi:10.1016/j.molcel.2014.03.045

385. Marshall EA, Stewart GL, Sage AP, Lam WL, Brown CJ. Beyond sequence homology: Cellular biology limits the potential of XIST to act as a miRNA sponge. *PLOS ONE*. 2019;14(8):e0221371. doi:10.1371/journal.pone.0221371
386. Wen X, Han X-R, Wang Y-J, et al. Effects of long noncoding RNA SPRY4-IT1-mediated EZH2 on the invasion and migration of lung adenocarcinoma. *J Cell Biochem*. 2018;119(2):1827-1840. doi:10.1002/jcb.26344
387. Cao S, Lin L, Xia X, Wu H. lncRNA SPRY4-IT1 Regulates Cell Proliferation and Migration by Sponging miR-101-3p and Regulating AMPK Expression in Gastric Cancer. *Mol Ther Nucleic Acids*. 2019;17:455-464. doi:10.1016/j.omtn.2019.04.030
388. Shi X, Sun M, Liu H, et al. A critical role for the long non-coding RNA GAS5 in proliferation and apoptosis in non-small-cell lung cancer. *Mol Carcinog*. 2015;54 Suppl 1:E1-E12. doi:10.1002/mc.22120
389. Mao C, Wang X, Liu Y, et al. A G3BP1-Interacting lncRNA Promotes Ferroptosis and Apoptosis in Cancer via Nuclear Sequestration of p53. *Cancer Res*. 2018;78(13):3484-3496. doi:10.1158/0008-5472.CAN-17-3454
390. Su W, Feng S, Chen X, et al. Silencing of long non-coding RNA MIR22HG triggers cell survival/death signaling via oncogenes YBX1, MET, and p21 in lung cancer. *Cancer Res*. 2018;78(12):3207-3219. doi:10.1158/0008-5472.CAN-18-0222
391. Huang Z, Lei W, Tan J, Hu H-B. Long noncoding RNA LINC00961 inhibits cell proliferation and induces cell apoptosis in human non-small cell lung cancer. *J Cell Biochem*. 2018;119(11):9072-9080. doi:10.1002/jcb.27166
392. Tan B-S, Yang M-C, Singh S, et al. LncRNA NORAD is repressed by the YAP pathway and suppresses lung and breast cancer metastasis by sequestering S100P. *Oncogene*. 2019;38(28):5612-5626. doi:10.1038/s41388-019-0812-8
393. Yang R, Liu N, Chen L, et al. GIAT4RA functions as a tumor suppressor in non-small cell lung cancer by counteracting Uchl3-mediated deubiquitination of LSH. *Oncogene*. 2019;38(46):7133-7145. doi:10.1038/s41388-019-0909-0
394. Zhang C, Hao Y, Wang Y, Xu J, Teng Y, Yang X. TGF- $\beta$ /SMAD4-Regulated LncRNA-LINP1 Inhibits Epithelial-Mesenchymal Transition in Lung Cancer. *Int J Biol Sci*. 2018;14(12):1715-1723. doi:10.7150/ijbs.27197
395. Baud V, Karin M. Is NF-kappaB a good target for cancer therapy? Hopes and pitfalls. *Nat Rev Drug Discov*. 2009;8(1):33-40. doi:10.1038/nrd2781
396. Semenza GL. Targeting HIF-1 for cancer therapy. *Nat Rev Cancer*. 2003;3(10):721-732. doi:10.1038/nrc1187
397. Sporn MB, Liby KT. NRF2 and cancer: the good, the bad and the importance of context. *Nat Rev Cancer*. 2012;12(8):564-571. doi:10.1038/nrc3278
398. Conesa A, Madrigal P, Tarazona S, et al. A survey of best practices for RNA-seq data analysis. *Genome Biol*. 2016;17(1):13. doi:10.1186/s13059-016-0881-8
399. Thai P, Statt S, Chen CH, Liang E, Campbell C, Wu R. Characterization of a novel long noncoding RNA, SCAL1, induced by cigarette smoke and elevated in lung cancer cell lines. *Am J Respir Cell Mol Biol*. 2013;49(2):204-211. doi:10.1165/rcmb.2013-0159RC
400. Li P, Zhang N, Ping F, Gao Y, Cao L. lncRNA SCAL1 inhibits inducible nitric oxide synthase in lung cells under high-glucose conditions. *Exp Ther Med*. 2019;18(3):1831-1836. doi:10.3892/etm.2019.7729



401. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science*. 2014;346(6213):1258096. doi:10.1126/science.1258096
402. Stadtmauer EA, Fraietta JA, Davis MM, et al. CRISPR-engineered T cells in patients with refractory cancer. *Science*. 2020;367(6481). doi:10.1126/science.aba7365
403. Dickinson DJ, Ward JD, Reiner DJ, Goldstein B. Engineering the *Caenorhabditis elegans* genome using Cas9-triggered homologous recombination. *Nat Methods*. 2013;10(10):1028-1034. doi:10.1038/nmeth.2641
404. Huan L, Guo T, Wu Y, et al. Hypoxia induced LUCAT1/PTBP1 axis modulates cancer cell viability and chemotherapy response. *Mol Cancer*. 2020;19(1):11. doi:10.1186/s12943-019-1122-z
405. Sun Y, Jin S-D, Zhu Q, et al. Long non-coding RNA LUCAT1 is associated with poor prognosis in human non-small lung cancer and regulates cell proliferation via epigenetically repressing p21 and p57 expression. *Oncotarget*. 2017;8(17):28297-28311. doi:10.18632/oncotarget.16044
406. Y I, K T. Ligand-targeted delivery of therapeutic siRNA. *Pharm Res*. 2006;23(8):1631-1640. doi:10.1007/s11095-006-9001-x
407. Martinez J, Patkaniowska A, Urlaub H, Lührmann R, Tuschl T. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. *Cell*. 2002;110(5):563-574. doi:10.1016/s0092-8674(02)00908-x
408. Vickers TA, Koo S, Bennett CF, Crooke ST, Dean NM, Baker BF. Efficient reduction of target RNAs by small interfering RNA and RNase H-dependent antisense agents. A comparative analysis. *J Biol Chem*. 2003;278(9):7108-7118. doi:10.1074/jbc.M210326200
409. Zeng Y, Cullen BR. RNA interference in human cells is restricted to the cytoplasm. *RNA*. 2002;8(7):855-860.
410. Rüdell S, Flatley A, Weinmann L, Kremmer E, Meister G. A multifunctional human Argonaute2-specific monoclonal antibody. *RNA N Y N*. 2008;14(6):1244-1253. doi:10.1261/rna.973808
411. Watts JK, Corey DR. Gene silencing by siRNAs and antisense oligonucleotides in the laboratory and the clinic. *J Pathol*. 2012;226(2):365-379. doi:10.1002/path.2993
412. Gagnon KT, Corey DR. Argonaute and the Nuclear RNAs: New Pathways for RNA-Mediated Control of Gene Expression. *Nucleic Acid Ther*. 2012;22(1):3-16. doi:10.1089/nat.2011.0330
413. Haaland RE, Herrmann CH, Rice AP. siRNA depletion of 7SK snRNA induces apoptosis but does not affect expression of the HIV-1 LTR or P-TEFb-dependent cellular genes. *J Cell Physiol*. 2005;205(3):463-470. doi:10.1002/jcp.20528
414. Kim S-S, Harford JB, Moghe M, Rait A, Pirolo KF, Chang EH. Targeted nanocomplex carrying siRNA against MALAT1 sensitizes glioblastoma to temozolomide. *Nucleic Acids Res*. 2018;46(3):1424-1440. doi:10.1093/nar/gkx1221
415. Wang L, Zhu H. Long non-coding nuclear paraspeckle assembly transcript 1 acts as prognosis biomarker and increases cell growth and invasion in cervical cancer by sequestering microRNA-101. *Mol Med Rep*. 2018;17(2):2771-2777. doi:10.3892/mmr.2017.8186
416. Kampmann M. CRISPRi and CRISPRa Screens in Mammalian Cells for Precision Biology and Medicine. *ACS Chem Biol*. 2018;13(2):406-416. doi:10.1021/acscchembio.7b00657

417. Bester AC, Lee JD, Chavez A, et al. An Integrated Genome-wide CRISPRa Approach to Functionalize lncRNAs in Drug Resistance. *Cell*. 2018;173(3):649-664.e20. doi:10.1016/j.cell.2018.03.052
418. Mehta G, Hsiao AY, Ingram M, Luker GD, Takayama S. Opportunities and challenges for use of tumor spheroids as models to test drug delivery and efficacy. *J Control Release Off J Control Release Soc*. 2012;164(2):192-204. doi:10.1016/j.jconrel.2012.04.045
419. Liu SJ, Nowakowski TJ, Pollen AA, et al. Single-cell analysis of long non-coding RNAs in the developing human neocortex. *Genome Biol*. 2016;17:67. doi:10.1186/s13059-016-0932-1
420. West JA, Davis CP, Sunwoo H, et al. The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Mol Cell*. 2014;55(5):791-802. doi:10.1016/j.molcel.2014.07.012
421. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*. 2003;31(13):3406-3415. doi:10.1093/nar/gkg595
422. Zuker M, Jacobson AB. Using reliability information to annotate RNA secondary structures. *RNA*. 1998;4(6):669-679.
423. Moreno Leon L, Gautier M, Allan R, et al. The nuclear hypoxia-regulated NLUCAT1 long non-coding RNA contributes to an aggressive phenotype in lung adenocarcinoma through regulation of oxidative stress. *Oncogene*. 2019;38(46):7146-7165. doi:10.1038/s41388-019-0935-y
424. Buffa FM, Harris AL, West CM, Miller CJ. Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene. *Br J Cancer*. 2010;102(2):428-435. doi:10.1038/sj.bjc.6605450
425. Gastaldi C, Bertero T, Xu N, et al. miR-193b/365a cluster controls progression of epidermal squamous cell carcinoma. *Carcinogenesis*. 2014;35(5):1110-1120. doi:10.1093/carcin/bgt490
426. Ades EW, Candal FJ, Swerlick RA, et al. HMEC-1: Establishment of an Immortalized Human Microvascular Endothelial Cell Line. *J Invest Dermatol*. 1992;99(6):683-690. doi:10.1111/1523-1747.ep12613748
427. Turner CE. Paxillin and focal adhesion signalling. *Nat Cell Biol*. 2000;2(12):E231-E236. doi:10.1038/35046659
428. Estrach S, Cailleateau L, Franco CA, et al. Laminin-binding integrins induce Dll4 expression and Notch signaling in endothelial cells. *Circ Res*. 2011;109(2):172-182. doi:10.1161/CIRCRESAHA.111.240622
429. Watt FM, Estrach S, Ambler CA. Epidermal Notch signalling: differentiation, cancer and adhesion. *Curr Opin Cell Biol*. 2008;20(2):171-179. doi:10.1016/j.ceb.2008.01.010
430. Stenzel D, Franco CA, Estrach S, et al. Endothelial basement membrane limits tip cell formation by inducing Dll4/Notch signalling in vivo. *EMBO Rep*. 2011;12(11):1135-1143. doi:10.1038/embor.2011.194
431. Moreno Roig E, Yaromina A, Houben R, Groot AJ, Dubois L, Vooijs M. Prognostic Role of Hypoxia-Inducible Factor-2 $\alpha$  Tumor Cell Expression in Cancer Patients: A Meta-Analysis. *Front Oncol*. 2018;8. doi:10.3389/fonc.2018.00224
432. Ye J, Zhu J, Chen H, et al. A novel lncRNA-LINC01116 regulates tumorigenesis of glioma by targeting VEGFA. *Int J Cancer*. 2020;146(1):248-261. doi:10.1002/ijc.32483

433. Brodie S, Lee HK, Jiang W, et al. The novel long non-coding RNA TALNEC2, regulates tumor cell growth and the stemness and radiation response of glioma stem cells. *Oncotarget*. 2017;8(19):31785-31801. doi:10.18632/oncotarget.15991
434. Zhang Z-F, Xu H-H, Hu W-H, Hu T-Y, Wang X-B. LINC01116 promotes proliferation, invasion and migration of osteosarcoma cells by silencing p53 and EZH2. *Eur Rev Med Pharmacol Sci*. 2019;23(16):6813-6823. doi:10.26355/eurrev\_201908\_18720
435. Zhang B, Yu L, Han N, et al. LINC01116 targets miR-520a-3p and affects IL6R to promote the proliferation and migration of osteosarcoma cells through the Jak-stat signaling pathway. *Biomed Pharmacother Biomedecine Pharmacother*. 2018;107:270-282. doi:10.1016/j.biopha.2018.07.119
436. Wu J, Chen Z, Zhang L, et al. Knockdown of LINC01116 inhibits cell migration and invasion in head and neck squamous cell carcinoma through epithelial-mesenchymal transition pathway. *J Cell Biochem*. 2020;121(1):867-875. doi:10.1002/jcb.29331
437. Hu H-B, Chen Q, Ding S-Q. LncRNA LINC01116 competes with miR-145 for the regulation of ESR1 expression in breast cancer. *Eur Rev Med Pharmacol Sci*. 2018;22(7):1987-1993. doi:10.26355/eurrev\_201804\_14726
438. Su X, Zhang J, Luo X, et al. LncRNA LINC01116 Promotes Cancer Cell Proliferation, Migration And Invasion In Gastric Cancer By Positively Interacting With lncRNA CASC11. *OncoTargets Ther*. 2019;12:8117-8123. doi:10.2147/OTT.S208133
439. Shimojo M, Kasahara Y, Inoue M, et al. A gapmer antisense oligonucleotide targeting SRRM4 is a novel therapeutic medicine for lung cancer. *Sci Rep*. 2019;9(1):7618. doi:10.1038/s41598-019-43100-1
440. Lim KRQ, Maruyama R, Echigoya Y, et al. Inhibition of DUX4 expression with antisense LNA gapmers as a therapy for facioscapulohumeral muscular dystrophy. *Proc Natl Acad Sci*. 2020;117(28):16509-16515. doi:10.1073/pnas.1909649117
441. Savary G, Buscot M, Dewaeles E, et al. The DNM3OS lncRNA is a reservoir of fibromiRs with major functions in lung fibroblast response to TGF- $\beta$  and pulmonary fibrogenesis. *Eur Respir J*. 2017;50(suppl 61). doi:10.1183/1393003.congress-2017.OA2908
442. Wheeler TM, Leger AJ, Pandey SK, et al. Targeting nuclear RNA for in vivo correction of myotonic dystrophy. *Nature*. 2012;488(7409):111-115. doi:10.1038/nature11362
443. Watts JK, Corey DR. Silencing disease genes in the laboratory and the clinic. *J Pathol*. 2012;226(2):365-379. doi:10.1002/path.2993
444. Bertram JS, Janik P. Establishment of a cloned line of Lewis Lung Carcinoma cells adapted to cell culture. *Cancer Lett*. 1980;11(1):63-73. doi:10.1016/0304-3835(80)90130-5
445. Shoval H, Karsch-Bluman A, Brill-Karniely Y, et al. Tumor cells and their crosstalk with endothelial cells in 3D spheroids. *Sci Rep*. 2017;7(1):10428. doi:10.1038/s41598-017-10699-y
446. Jackson AL, Linsley PS. Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. *Nat Rev Drug Discov*. 2010;9(1):57-67. doi:10.1038/nrd3010
447. Jackson AL, Bartz SR, Schelter J, et al. Expression profiling reveals off-target gene regulation by RNAi. *Nat Biotechnol*. 2003;21(6):635-637. doi:10.1038/nbt831
448. Yu W, Hegarty JP, Berg A, et al. NKX2-3 Transcriptional Regulation of Endothelin-1 and VEGF Signaling in Human Intestinal Microvascular Endothelial Cells. *PLOS ONE*. 2011;6(5):e20454. doi:10.1371/journal.pone.0020454

449. Ramasamy SK, Kusumbe AP, Adams RH. Regulation of tissue morphogenesis by endothelial cell-derived signals. *Trends Cell Biol.* 2015;25(3):148-157. doi:10.1016/j.tcb.2014.11.007
450. Marty-Santos L, Cleaver O. Pdx1 regulates pancreas tubulogenesis and E-cadherin expression. *Dev Camb Engl.* 2016;143(1):101-112. doi:10.1242/dev.126755
451. Ischenko I, Petrenko O, Hayman MJ. Analysis of the tumor-initiating and metastatic capacity of PDX1-positive cells from the adult pancreas. *Proc Natl Acad Sci.* 2014;111(9):3466-3471. doi:10.1073/pnas.1319911111
452. Ma J, Wang B-B, Ma X-Y, Deng W-P, Xu L-S, Sha W-H. Potential involvement of heat shock proteins in pancreatic-duodenal homeobox-1-mediated effects on the genesis of gastric cancer: A 2D gel-based proteomic study. *World J Gastroenterol.* 2018;24(37):4263-4271. doi:10.3748/wjg.v24.i37.4263
453. Carlevaro-Fita J, Johnson R. Global Positioning System: Understanding Long Noncoding RNAs through Subcellular Localization. *Mol Cell.* 2019;73(5):869-883. doi:10.1016/j.molcel.2019.02.008
454. Vrakas CN, Herman AB, Ray M, Kelemen SE, Scalia R, Autieri MV. RNA stability protein ILF3 mediates cytokine-induced angiogenesis. *FASEB J Off Publ Fed Am Soc Exp Biol.* 2019;33(3):3304-3316. doi:10.1096/fj.201801315R
455. Yao P, Potdar AA, Ray PS, et al. The HILDA complex coordinates a conditional switch in the 3'-untranslated region of the VEGFA mRNA. *PLoS Biol.* 2013;11(8):e1001635. doi:10.1371/journal.pbio.1001635
456. Vumbaca F, Phoenix KN, Rodriguez-Pinto D, Han DK, Claffey KP. Double-stranded RNA-binding protein regulates vascular endothelial growth factor mRNA stability, translation, and breast cancer angiogenesis. *Mol Cell Biol.* 2008;28(2):772-783. doi:10.1128/MCB.02078-06
457. Afonina E, Stauber R, Pavlakis GN. The human poly(A)-binding protein 1 shuttles between the nucleus and the cytoplasm. *J Biol Chem.* 1998;273(21):13015-13021. doi:10.1074/jbc.273.21.13015
458. Böhm BB, Fehrl Y, Janczi T, Schneider N, Burkhardt H. Cell adhesion-induced transient interaction of ADAM15 with poly(A) binding protein at the cell membrane colocalizes with mRNA translation. *PLoS ONE.* 2018;13(9). doi:10.1371/journal.pone.0203847
459. Kourtidis A, Necela B, Lin W-H, et al. Cadherin complexes recruit mRNAs and RISC to regulate epithelial cell signaling. *J Cell Biol.* 2017;216(10):3073-3085. doi:10.1083/jcb.201612125