



HAL
open science

Nouvelles signalisations impliquant les lysines-méthyltransférases SMYD2 et SMYD3 dans le cancer

Gaël Stéphane Wenceslas Roth

► **To cite this version:**

Gaël Stéphane Wenceslas Roth. Nouvelles signalisations impliquant les lysines-méthyltransférases SMYD2 et SMYD3 dans le cancer. Biologie du développement. Université Grenoble Alpes [2020-..], 2020. Français. NNT : 2020GRALV025 . tel-03505886

HAL Id: tel-03505886

<https://theses.hal.science/tel-03505886v1>

Submitted on 1 Jan 2022

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

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

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : **Biologie du développement - Oncogenèse**

Arrêté ministériel : 25 mai 2016

Présentée par

Gaël ROTH

Thèse dirigée par le **Pr Pierre HAINAUT, PhD**,
Co-encadrée par le **Dr Nicolas REYNOIRD, PhD**

Préparée au sein de l'Institut pour l'Avancée des Biosciences (IAB) INSERM
U1209 / CNRS UMR 5309 / Université de Grenoble-Alpes

dans l'**École Doctorale Chimie et Sciences du Vivant**

Nouvelles signalisations impliquant les lysines-méthyltransférases SMYD2 et SMYD3 dans le cancer

Thèse soutenue publiquement le **10 juillet 2020**
devant le jury composé de :

M. Denis MORO-SIBILOT, Professeur des universités - Praticien
Hospitalier, Université Grenoble-Alpes (Examineur, Président du jury)

Mme Muriel LEROMANCER, Directrice de recherche, Centre de
Recherche en Cancérologie de Lyon (Rapportrice)

M. Eric JULIEN, Directeur de recherche, Institut de Recherche en
Cancérologie de Montpellier (Rapporteur)

Mme Souhila MEDJKANE, Maître de conférences, Université Paris Diderot
(Examinatrice)

M. Pierre HAINAUT, Professeur des universités, Université Grenoble Alpes,
Directeur de l'IAB (Directeur de thèse)

M. Nicolas REYNOIRD, Chargé de recherche, IAB (Invité)



TABLE DES MATIÈRES

TABLE DES MATIÈRES.....	3
LISTE DES ABRÉVIATIONS.....	7
RÉSUMÉ DE LA THÈSE	11
THESIS SUMMARY	12
INTRODUCTION GÉNÉRALE	13
I. ÉPIGÉNÉTIQUE ET CONTRÔLE DE L'EXPRESSION GÉNIQUE	14
A. RÉGULATION AU NIVEAU DE LA CHROMATINE	14
1. La chromatine.....	14
2. Les modifications des histones	16
3. La méthylation de l'ADN.....	20
B. RÉGULATION POST-TRANSCRIPTIONNELLE	24
1. La méthylation de l'ARN.....	24
2. Épissage alternatif.....	25
3. Les ARNs non-codants	28
C. LES MODIFICATIONS POST-TRADUCTIONNELLES.....	30
II. LA MÉTHYLATION DES PROTÉINES	35
A. GÉNÉRALITÉS	35
B. LA MÉTHYLATION DES ARGININES	37
C. LA MÉTHYLATION DES LYSINES	40
1. La méthylation des lysines : une signalisation dynamique et réversible	40
2. Les principales signalisations régulées par la méthylation des lysines.....	48
3. Méthylation des lysines et cancérogenèse.....	60
4. La signalisation par méthylation des lysines : nouvelles cibles thérapeutiques	64
III. LYSINE-MÉTHYLTRANSFÉRASES SET AND MYND DANS LE CANCER.....	67
A. LYSINE-MÉTHYLTRANSFÉRASES SET AND MYND : GÉNÉRALITÉS	67
B. IMPLICATIONS DE SMYD2 DANS LE CANCER	72
1. Expression dans les cancers.....	72
2. Les différentes signalisations de SMYD2 dans la carcinogenèse.....	74
3. Inhibition pharmacologique de SMYD2.....	78

C. IMPLICATIONS DE SMYD3 DANS LE CANCER	79
1. Expression dans les cancers.....	79
2. SMYD3 : marques de méthylation impliquées dans la carcinogenèse	83
3. Inhibition pharmacologique de SMYD3	86
IV. CONCLUSION.....	89
V. RÉFÉRENCES.....	91

PARTIE I : IDENTIFICATION DE NOUVELLES SIGNALISATIONS DE SMYD2 ET SMYD3 DANS LE CANCER 105

I. RÉSUMÉ DE LA PARTIE I	106
II. ARTICLE.....	108
Abstract	109
Introduction.....	110
Results	114
Discussion and Perspectives	121
Material and Method.....	124
References	129
Tables.....	135
Figures	137
Supplementary Data.....	146

PARTIE II : LA TRIMÉTHYLATION DE RNF113A PAR SMYD3 INDUIT UNE AUGMENTATION DE SON ACTIVITÉ E3 UBIQUITINE LIGASE ET PARTICIPE À LA CHIMIORÉSISTANCE AUX AGENTS ALKYLANTS 157

I. RÉSUMÉ DE LA PARTIE II	158
II. ARTICLE.....	159
Abstract	160
Introduction.....	161
Results	164
Discussion	171
References	174
Figures	178
Material and Method.....	187
Supplementary Figures	194

DISCUSSION GÉNÉRALE 201

Identification de nouveaux substrats : contraintes techniques et critères de choix..... 202

Du candidat présomptif au substrat validé : stratégies et priorités..... 205

Caractérisation des conséquences de la triméthylation de RNF113A par SMYD3 208

Conclusions : de la découverte à la perspective clinique..... 214

Références 217

ANNEXE..... 219

LISTE DES ABRÉVIATIONS

Abréviation	Définition
53BP1	Tumor suppressor p53-binding protein 1
ADN	Acide désoxyribonucléique
AHNAK	Neuroblast differentiation-associated protein
AKT1	Serine/threonine-protein kinase 1
ALKBH3	Alkylated DNA Repair Protein alkb Homolog 3
ARN	Acide ribonucléique
ASCC	Activating signal cointegrator complex
ATF7IP	Activating transcription factor 7-interacting protein 1
BCAR3	Breast cancer anti-estrogen resistance 3
BER	Base excision repair
BMPR2	Bone morphogenetic protein receptor type 2
BTF3	Basic transcription factor 3
CARM1	Coactivator associated arginine methyltransferase 1
CCAR2	Cell cycle and apoptosis regulator 2
CDK	Cyclin E/cyclin-dependent kinase
Cdx2	Caudal type homeobox 2
CpG	Cytosine–phosphate–guanine
CRC	Colorectal cancer
CTD	C-terminal domain
CUL4	Cullin-4A
DCAF1	DDB1 and CUL4 associated factor 1
DDB1	DNA damage-binding protein 1
DNMT	ADN méthyle transférase
DOT1L	Disruptor of telomeric silencing 1-like
EEF2	Eukaryotic translation elongation factor 2
EHMT	Euchromatic histone lysine methyltransferase
ERα	Estrogen receptor alpha
ESCC	Esophageal squamous cell carcinoma
EZH2	Enhancer of zeste homolog 2
EZR	Ezrin
FEN1	Flap structure-specific endonuclease 1
Fn1	Fibronectin 1
GATA4	GATA binding protein 4
H	Histone
HAT	Histone acétyltransférase
HCC	Hepatocellular carcinoma
HDAC	Histone-déacétylase
Her2	Human epidermal growth factor receptor 2
HIF-1α	Hypoxia-induced factor 1 α
His	Histidin
HSP	Heat shock protein

IHC	Immunohistochimie
IL	Interleukine
JAK	Janus kinase
JARID	Jumonji/ARID Domain-Containing Protein
JmJD	Domaine Jumonji
JNK	Jun amino- terminal kinases
K	Lysine
KDM	Lysine deméthylase
KMT	Lysine méthyltransférase
L3MBTL	Lethal(3)malignant brain tumor-like protein
LC-MS/MS	Liquid chromatography and mass spectrometry analysis
LIG1	DNA ligase 1
lincRNA	Long ARN non codant intergénique
LOXL2	Lysyl Oxidase Like 2
LSD1	Lysine-specific histone demethylase 1A
m6A	N6-méthyladénosine
MAP3K2	Mitogen-activated protein kinase kinase kinase 2
MAPKAPK3	Mapk activated protein kinase 3
MBD	Methyl-binding domain
MDM2	Mouse double minute 2 homolog
MecP2	Methyl cpg binding protein 2
METTL14	Methyltransferase like 14
MGMT	O-6-methylguanine-DNA methyltransferase
miRNA	Microarn
MLL	Mixed lineage leukemia
Mmp9	Matrix Metallopeptidase 9
MMR	Mismatch repair
MMS	Methyl methanesulfonate
MPP8	M-phase phosphoprotein 8
MT	Méthyltransférase
MYND	Myeloid- nervy-deaf1
MYPT1	Myosin phosphatase target subunit 1
N6-CMT	N6-adenosine-methyltransferase
Nanog	Nanog Homeobox
NCOA3	Nuclear receptor coactivator 3
ncRNA	ARN non codant
NER	Nucleotide excision repair
NF-κB	Nuclear factor-kappa B
NLS	Nuclear localization signal
NMD	Non-sense mediated decay
NSD	Nuclear receptor binding SET domain protein
NuRD	Mi-2/nucleosome remodeling and histone deacetylase
Oct4	Octamer-binding transcription factor 4
PAD	Peptidyl-arginines déiminases
PAK4	P21-Activated kinase 4
PARP1	Poly(ADP-ribose) polymérase 1

PCNA	Proliferating cell nuclear antigen
PDAP1	Pdgfa associated protein 1
PDGFRa	Platelet-derived growth factor receptor A
PGM	Proline, glycine et méthionine
PH	Pleckstrin homology
PHD	Plant homeodomain
PHF20L1	Phd finger protein 20 like 1
PI3K	Phosphatidylinositol 3-kinase
piRNA	Petit ARN associé aux protéines piwi
PP2a	Protein phosphatase 2a
PP4	Protein phosphatase 4
PRC2	Polycomb repressive complex 2
PRMT	Arginine-méthyl transférases
PTEN	Phosphatase and tensin homolog
PTM	Modification post traductionnelle
Rb	Retinoblastoma
RIOK1	RIO kinase 1
RISC	RNA induced silencing complex
RIZ	Retinoblastoma-interacting zinc-finger
RNAi	ARN interférent
RNF113A	RING finger protein 113A
RORα	RAR-related orphan receptor alpha
rRNA	ARN ribosomal
RRP44	Exosome complex exonuclease
RTF1	Paf1/RNA polymerase II complex component
SAM	S-adénosyl méthionine
SAPK	Stress-activated protein kinases
SARG	Specifically androgen-regulated gene protein
SCC	Lung squamous carcinoma
SCLC	Small cell lung cancer
SET	Su(var)3-9, Enhancer-of-zeste and Trithorax
SET-I	SET insertion
SETD	SET domain containing
shARN	Short hairpin ARN
SILAC	Stable Isotope Labeling of Amino Acids in Cell Culture
SMYD	SET and MYND
snoRNA	Petit ARN nucléolaire
snRNA	Petit ARN nucléaire
Snai1	Snail Family Transcriptional Repressor 1
Sox2	Sex determining region Y)-box 2
SRA	SET and RING finger-associated
STAT3	Signal transducer and activator of transcription 3
STUB1	STIP1 Homology and U-Box Containing Protein 1
SUMO	Small Ubiquitin-like Modifier
SUV39H1	Suppressor of variegation 3-9 homolog 1
TEM	Temozolomide

TET	Ten-eleven translocation
TKO	Inducible triple knock out of Rb1, Rbl2 and p53
TNFα	Tumor necrosis factor alpha
tRNA	ARN de transfert
TUBE	Tandem ubiquitin binding element
UHRF	Ubiquitin-like containing PHD and RING finger domains protein
UTP14A	U3 small nucleolar RNA-associated protein 14 homolog A
VEGFR1	Vascular endothelial growth factor receptor 1
WNT10B	Wnt Family Member 10B
ZBTB	Btb domain containing

RÉSUMÉ DE LA THÈSE

Les modifications post-traductionnelles des protéines sont impliquées dans un grand nombre de voies de signalisation physiologiques et sont essentielles au fonctionnement normal cellulaire. Leur dérégulation est impliquée dans de multiples processus pathologiques, et notamment dans la carcinogénèse. Les lysine-méthyltransférases SMYD2 et SMYD3 appartiennent à la famille des *SET and MYND domain-containing protein* (SMYD) et sont toutes deux surexprimées dans de nombreux cancers. Elles participent en effet à la régulation de multiples signalisations oncogéniques canoniques telles que la méthylation de p53 par SMYD2, ou la méthylation de VEGFR1 (*vascular endothelial growth factor receptor 1*), ou MAP3K2 (*Mitogen-activated protein kinase kinase kinase 2*) au sein de la voie Ras/Raf/Mek/Erk par SMYD3.

Dans ce travail de thèse, nous avons identifié un certain nombre de substrats potentiels de SMYD2 et SMYD3 à partir de données issues d'approches protéomiques telles que le ProtoArray® et le *SILAC-3xMBT pulldown* couplé à la spectrométrie de masse. La première partie de cette thèse a consisté à produire et valider la méthylation des différents substrats candidats les plus prometteurs. Nous avons ainsi confirmé *in vitro* puis en cellules d'adénocarcinome pancréatique MIA PaCa-2, les événements de monométhylation par SMYD2 de la protéine BCAR3 (*Breast cancer anti-estrogen resistance 3*) au niveau de sa lysine K334 -protéine impliquée dans l'agressivité et l'invasivité tumorale-, et de RTF1 (*RNA polymerase-associated protein*) au niveau de K587 -protéine appartenant au complexe PAF1-. De même, nous avons démontré la triméthylation de RNF113A (Ring Finger Protein 113A) K20 par SMYD3. La seconde partie de cette thèse a porté sur la caractérisation des conséquences phénotypiques de la triméthylation de RNF113A K20 dans un modèle de cancer pulmonaire à petites cellules. Nous avons démontré que la triméthylation de RNF113A bloque son interaction avec la phosphatase PP4. Cette perte d'interaction aboutit à une augmentation de son niveau de phosphorylation, entraînant une stimulation de son activité de E3-ubiquitine ligase impliquée dans la réponse aux dommages alkylants. La triméthylation de RNF113A par SMYD3 pourrait donc participer à la chimiorésistance des cellules tumorales aux agents alkylants.

Ainsi, ce travail de thèse a permis d'identifier des nouvelles signalisations prometteuses impliquant SMYD2 et SMYD3, et plus particulièrement de mettre en évidence le rôle de SMYD3 dans la réparation de l'ADN suite à l'exposition aux agents alkylants par le biais de la méthylation de RNF113A.

THESIS SUMMARY

Post-translational modifications are involved in a large number of physiological signaling pathways and are essential for normal cellular functioning. Their deregulation is involved in multiple pathological processes, and particularly in carcinogenesis. The lysine methyltransferases SMYD2 and SMYD3 belong to the family of SET and MYND domain-containing proteins (SMYD) and are both overexpressed in many cancers. They participate in the regulation of multiple canonical oncogenic pathways via the methylation of substrates such as p53 by SMYD2, or VEGFR1 (vascular endothelial growth factor receptor 1), and MAP3K2 (Mitogen -activated protein kinase kinase kinase 2) within the Ras / Raf / Mek / Erk pathway by SMYD3.

In this thesis work, using data from proteomic approaches without a priori such as ProtoArray® and SILAC-3xMBT pulldown coupled with mass spectrometry, we identified potential substrates of SMYD2 and SMYD3. The first part of this thesis consisted in the production of the most promising candidate substrates and the validation of their methylation. We confirmed *in vitro* and in MIA PaCa-2 pancreatic adenocarcinoma cells, the monomethylation by SMYD2 of the protein BCAR3 (Breast cancer anti-estrogen resistance 3) on its lysine K334 -protein involved in aggressiveness and tumor invasiveness-, and RTF1 (RNA polymerase-associated protein) on K587 -protein belonging to the PAF1- complex. Similarly, we demonstrated the trimethylation of RNF113A (Ring Finger Protein 113A) K20 by SMYD3. The second part of this thesis focused on the characterization of phenotypic consequences of RNF113A K20me3 in a small cell lung cancer model. We demonstrated that trimethylation of RNF113A repels its phosphatase PP4, leading to an increase of RNF113A phosphorylation. This induces an increase of its E3-ubiquitin ligase activity involved in DNA dealkylation repair by interacting with the ASCC complex (activating signal cointegrator 1 complex). The trimethylation of RNF113A by SMYD3 therefore participates in the chemoresistance of tumor cells to alkylating agents.

This thesis work therefore made it possible to identify new promising signals involving SMYD2 and SMYD3, and to highlight the role of SMYD3 in DNA repair following exposure to alkylating agents through methylation of RNF113A.

INTRODUCTION GÉNÉRALE

I- EPIGENETIQUE ET CONTROLE DE L'EXPRESSION GENIQUE

L'épigénétique est un concept décrit pour la première fois en 1942 par Waddington, comme étant l'étude des processus par lesquels le génotype, l'ensemble des gènes, engendre le phénotype, les caractéristiques de l'organisme. L'épigénétique a longtemps été considérée comme tout ce qui ne relève pas de la génétique mais, depuis les années 1990, une définition plus précise a émergé et on considère maintenant l'épigénétique comme l'étude des changements dans l'expression des gènes qui sont héréditaires lors de la mitose et/ou de la

méiose, et qui ne résultent pas de modifications de la séquence de l'ADN ¹.

La régulation de l'expression des gènes est le fruit d'un ensemble de mécanismes complexes, finement régulés et interconnectés. Dans cette introduction, je propose de distinguer deux niveaux de régulation : le niveau chromatinien, qui englobe l'organisation structurelle et la dynamique de la chromatine, l'expression et la modification des histones et la méthylation de l'ADN, et le niveau post-chromatinien, qui inclut la méthylation et l'épissage des ARNs messagers et les ARNs non codants. Cette distinction est cependant formelle, car des ARNs non-codants jouent un rôle essentiel dans l'organisation structurale de la chromatine.

A. RÉGULATION AU NIVEAU DE LA CHROMATINE

1. La chromatine

Dans les cellules eucaryotes, l'ADN est complexé à un ensemble de protéines et cette association ADN/protéines constitue la chromatine. Cette dernière est extrêmement organisée et permet de contenir l'ADN, structure longue d'environ 2 m, dans le noyau cellulaire, en le compactant et en conservant sa fonctionnalité. Parmi ces protéines, on retrouve les histones H2A, H2B, H3, H4, H1, qui sont étroitement associées à l'ADN et qui

vont réguler le niveau de compaction de l'ADN, et donc sa transcription. L'association de l'ADN aux histones se fait sous forme d'unités structurales, appelées nucléosomes, qui correspondent à 147 paires de bases d'ADN enroulées autour d'un octamère d'histones composé de deux exemplaires de chacune des histones suivantes : H2A, H2B, H3 et H4. Les nucléosomes correspondent à un premier niveau de compactage de l'ADN, notamment permis par le fait que les histones ont une charge positive alors que l'ADN est chargé négativement. Les histones H1 ne font pas partie du nucléosome proprement dit mais renforcent sa structure en maintenant en place l'ADN enroulé. L'empilement des nucléosomes en boucles d'ordre supérieur permet la compaction de la chromatine pour former les chromosomes ².

Les histones précédemment citées sont considérées comme des histones canoniques, auxquels on oppose les variants d'histone qui correspondent à des isoformes non alléliques, aux caractéristiques structurales et fonctionnelles différentes des histones conventionnelles. Les variants d'histone sont présents chez tous les organismes eucaryotes. Ils possèdent un domaine de repliement du même type que les histones canoniques avec trois hélices α et deux boucles. Ils s'incorporent dans une particule nucléosomale et remplacent l'histone canonique, conférant ainsi une fonctionnalité différente au nucléosome concerné avec des conséquences différentes sur la modulation de la chromatine et l'activité transcriptionnelle ³.

La chromatine est une structure très dynamique avec un niveau de densité variable pour permettre l'accès de l'ADN à la machinerie transcriptionnelle répliquative ou de réparation. Le contrôle de cette dynamique constitue donc un premier niveau essentiel de régulation de l'expression génique. La chromatine peut être classée en deux catégories selon son niveau de compactage : (i) l'euchromatine, qui est la forme la plus accessible et ouverte, et qui favorise l'activité transcriptionnelle des régions concernées, notamment par le recrutement de l'ARN polymérase ; (ii) l'hétérochromatine qui constitue un environnement réprimant l'activité transcriptionnelle. L'hétérochromatine se divise elle-même en deux catégories. La première est l'hétérochromatine constitutive, qui correspond à des régions pauvres en gènes, telles que les centromères et les télomères, avec un niveau de compaction élevé et

quasi-permanent. La seconde est l'hétérochromatine facultative, qui concerne des régions du génome contenant des gènes exprimés de manière différentielle selon le niveau de développement et de différenciation cellulaire et qui sont réprimés de manière différentielle au cours de la vie cellulaire ⁴.

2. Les modifications des histones

Les histones sont la cible de multiples modifications post-traductionnelles (PTMs, *post-translational modifications*), notamment au niveau de leur queue, riche en résidus sérine, lysine et arginine. Cette zone est située en extrémité N-terminale et est exposée à l'extérieur du nucléosome. Les PTMs les plus fréquentes sont l'acétylation, la méthylation, la phosphorylation ou l'ubiquitination (Table 1). Ces PTMs modulent l'activité des histones et interagissent entre elles de manière compétitive pour certaines et synergique pour d'autres. Cet ensemble de modifications est classiquement appelé « code histone » ⁵. Les PTMs induisent des modifications conformationnelles des histones qui influencent la structure de la chromatine, soit vers l'état d'euchromatine favorisant la transcription des gènes des régions concernées soit vers l'état d'hétérochromatine ⁴. Ces modifications peuvent provoquer des modifications physico-chimiques, telles que des modifications de charge des nucléosomes, ou induire des modifications d'interaction entre les histones et d'autres protéines impliquées dans la régulation de la chromatine et de l'expression génique. Ce lien entre PTMs des histones et conformation de la chromatine est connu depuis le début des années 1990 ⁶, et répond au concept de « *writers, erasers and readers* » (écrivains, effaceurs, lecteurs) comme illustré par la Figure 1 ⁷. Les enzymes catalysant une PTMs sont les *writers*. Il peut s'agir par exemple des histone-acétylases ou des histone-méthyltransférases. Les *erasers* sont les enzymes retirant ces modifications, telles que les histone-déacétylases (HDAC) ou les déméthylases. Les *readers* regroupent différents interacteurs qui se lient spécifiquement à la forme modifiée de l'histone via un domaine spécifique et enclenchent un effet signalétique en aval.

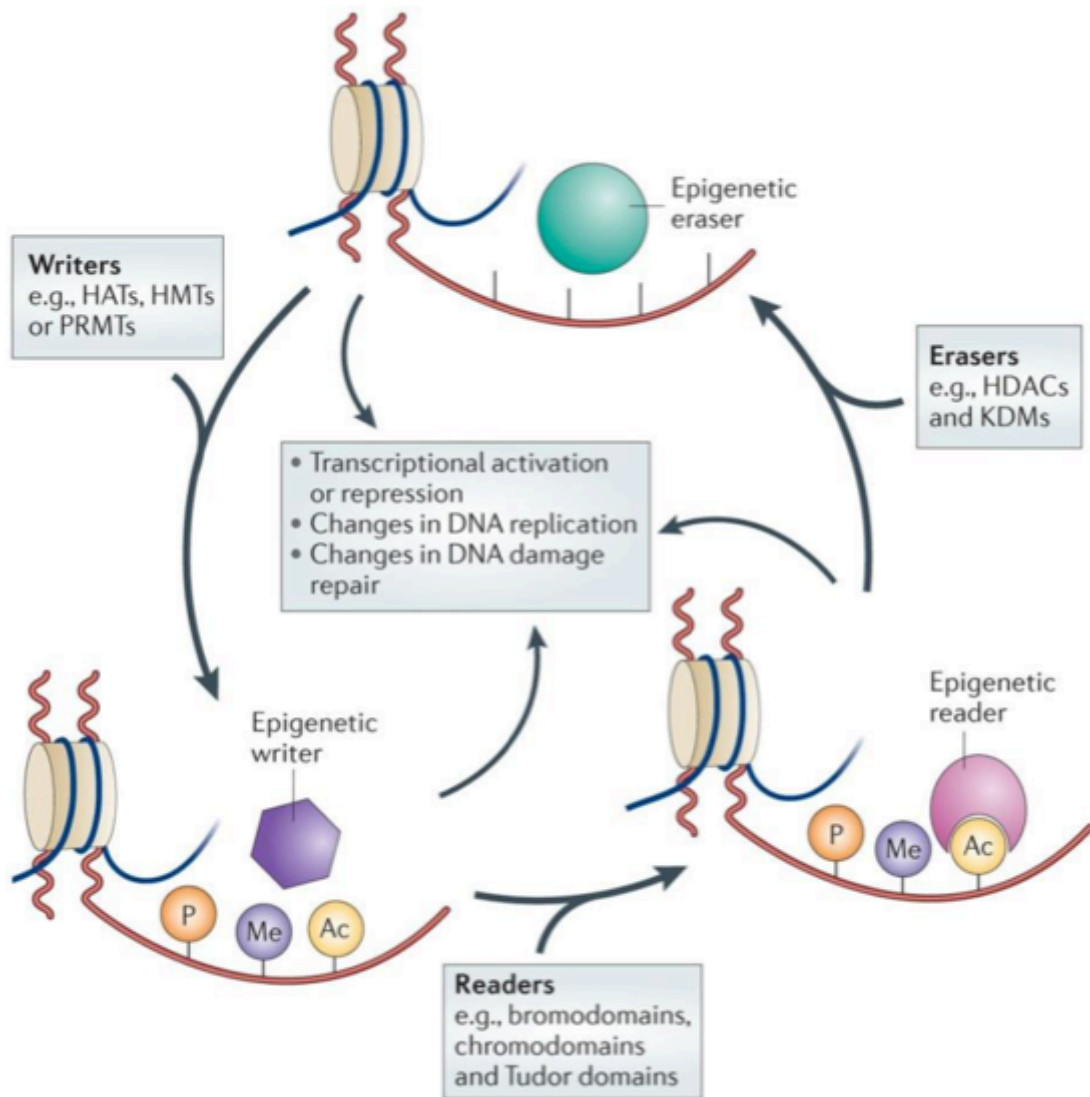


Figure 1: Concept de *writers*, *readers* and *erasers* in histones PTMs (Falkenberg and Johnstone. Nature Review 2014). La compaction de la chromatine et la régulation de l'expression génique dépendent de multiples PTMs orchestrées par des mécanismes de régulation complexes. Les PTMs sont catalysées par des enzymes ou *writers*, tels que les acétylases ou les méthyltransférases, et elles sont retirées par les *erasers*, tels que les déacétylases ou les déméthylases. Les *readers* correspondent aux protéines capables de lier et d'interagir de manière spécifique à la forme modifiée ou non.

Les premières PTMs décrites ont été l'acétylation et la méthylation des histones par Alfrey et al dans les années 1960. Dès lors, l'acétylation des histones semblait associée à l'augmentation de la transcription des régions de l'ADN en contact avec les histones acétylées⁸. Depuis, de nombreuses marques d'acétylation ont été décrites, telles que les

lysines H3K9, H3K14, H3K18, H3K23, H4K5, H4K8, H4K12 et H4K16, associées à une augmentation de l'activité transcriptionnelle ⁹. De la même manière que l'acétylation, la phosphorylation induit une modification structurale des histones en diminuant leur charge positive, ce qui réduit leur force d'interaction avec l'ADN ⁴. La phosphorylation peut donc influencer l'état de compaction de la chromatine et donc la transcription. Elle semble également faciliter la réparation de l'ADN en créant une ouverture en regard des lésions ADN qui permet l'accès aux différents systèmes de réparation de l'ADN (notamment dans le cas de la phosphorylation de la variante X de l'histone H2A) ¹⁰. D'autres PTMs induisent des changements conformationnels des nucléosomes de par leur masse qui induit un encombrement stérique qui modifie l'organisation structurale des histones. C'est le cas de l'ubiquitinylation par exemple, qui a lieu au niveau des lysines et qui consiste en la fixation covalente d'une ubiquitine, polypeptide de 76 acides aminés, ou de la SUMOylation (protéines de 100 aminés) ^{11,12}. Enfin, la méthylation des lysines et des arginines au niveau des histones ne modifie pas la charge ou la masse des histones, mais participe activement à la régulation de leur activité en modifiant notamment les interactions protéine-protéine. Les principales signalisations de méthylation des histones sont détaillées dans le chapitre « II-Méthylation des protéines ».

Comme mentionné précédemment, ces différentes PTMs influencent les interactions des histones avec d'autres protéines qui se lient spécifiquement à la forme modifiée ou non de l'histone et qui correspondent aux *readers*. Ces derniers possèdent généralement un domaine de liaison aux histones, également appelé domaine de lecture, spécifique d'une PTM. Un certain nombre de domaines différents sont décrits à ce jour, tels que les bromodomains, spécifiques des lysines acétylées. Ce sont de loin les lysines méthylées qui interagissent avec la plus grande variété de domaines différents, tels que les domaines Tudor, les chromodomains, les MBD (*methyl-binding domains*), ou les PHD (*plant homeodomain*) *fingers* ⁴. Si la notion de code histone a émergé, c'est principalement lié aux différents niveaux de régulation de l'activité des histones que sont les pertes et gains d'interactions avec d'autres protéines ou *readers*, couplés aux phénomènes

d'interdépendance, de synergie, de compétition ou d'antagonisme qui peuvent avoir lieu entre plusieurs PTMs dans une même région voire sur un même résidu ¹³. Cette dernière notion est illustrée par de multiples exemples tels que la phosphorylation de la sérine 10 de H3 (H3S10), inhibée en présence de la méthylation de H3K9 par SUV39H1 ¹⁴, la compétition entre l'acétylation et la méthylation de H3K9 au cours du cycle cellulaire ou encore la méthylation de H3K4 par scCOMPASS (homologue du complexe MLL (*mixed lineage leukemia*) chez l'homme) et de H3K79 par scDot1 dépendantes de l'ubiquitinylation de H2BK123 par scRad6/Bre1 chez *Saccharomyces cerevisiae* ¹⁵ (Figure 2). Les modifications des histones semblent également interagir avec la méthylation de l'ADN, un autre mode de

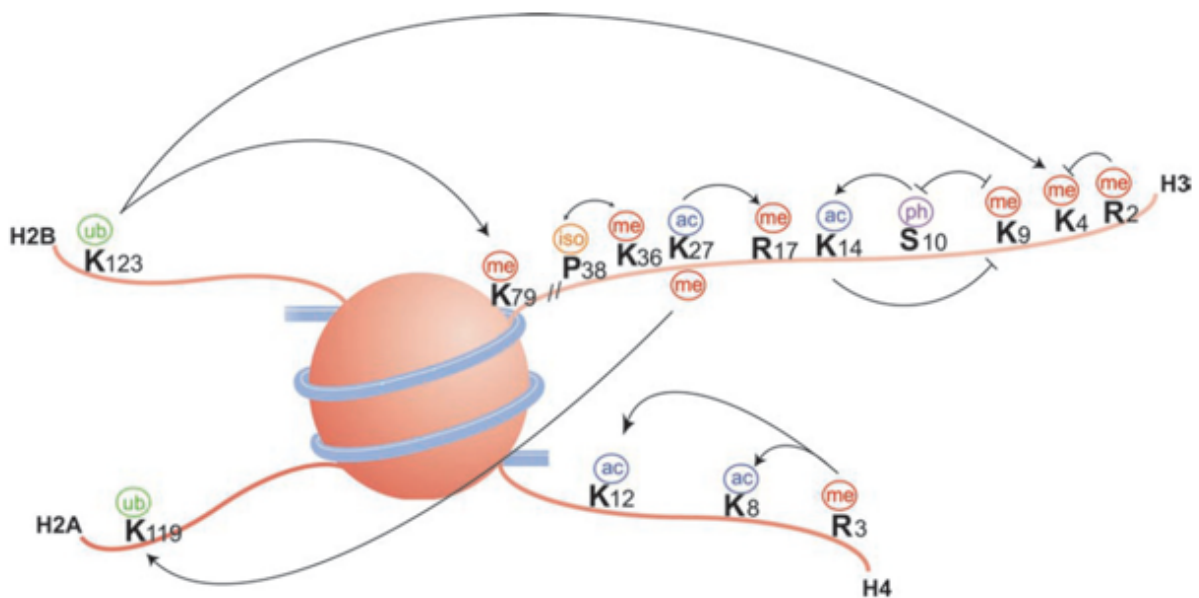


Figure 2 : Régulation des histones et interconnexions entre les différentes modifications post traductionnelles (PTMs) (Bannister and Kouzarides. Cell Research 2011). Les PTMs peuvent se potentialiser ou s'antagoniser selon les sites modifiés.

régulation de l'expression génique au niveau transcriptionnel. Par exemple, la protéine UHRF1 (*Ubiquitin-like, containing PHD and RING finger domains protein 1*) ne fixe que les nucléosomes qui portent la marque H3K9me3, qui semble elle-même corrélée à la méthylation de l'ADN nucléosomal au niveau des îlots CpG¹⁶. A l'inverse, la méthylation de l'ADN peut inhiber certaines PTMs, comme par exemple la déméthylase KDM2A qui ne peut fixer les nucléosomes portant la marque H3K9me3 si l'ADN nucléosomal n'est pas méthylé.

Indépendamment de ces interactions avec les modifications des histones, la méthylation de l'ADN est un mode à part entière de régulation de l'expression génique, qui résulte d'une régulation complexe.

Modification	Acides aminés modifiés	Principales fonctions régulées
Acétylation	Lysine-ac	Transcription, réparation de l'ADN, réplication, compaction de la chromatine
Méthylation des lysines	K-me1, K-me2, K-me3	Transcription (promotion, inhibition), réparation de l'ADN, régulation de la compaction de la chromatine, méthylation de l'ADN, régulation du cycle cellulaire, mitose.
Méthylation des arginines	R-me1, R-me2a, K-mes	Transcription
Phosphorylation	S-ph, T-ph	Transcription, réparation de l'ADN, compaction de la chromatine
Ubiquitinylation	K-ub	Transcription, réparation de l'ADN
SUMOylation	S-su	Transcription
ADP ribosylation	E-ar	Transcription
Déamination	R > Cit	Transcription
Isomérisation de la proline	P-cis > P-trans	Transcription

Table 1 : Résumé des principales modifications des histones. Descriptions des modifications décrites, des acides aminés concernés par ces modifications, et des fonctions biologiques régulées par ces modifications. K : lysine, R : arginine, S : sérine, T : tyrosine, E : glutamine, P : proline.

3. La méthylation de l'ADN

La méthylation de l'ADN constitue l'une des principales modifications épigénétiques étudiées et son rôle comme facteur de régulation transcriptionnelle est décrit depuis les années 1970¹⁷. Cette marque épigénétique a lieu au niveau des bases cytosines lorsque ces dernières sont au sein d'un dinucléotide cytosine–phosphate–guanine (CpG). Elle consiste en l'ajout d'un groupement méthyle (–CH₃) à la place d'un atome d'hydrogène au niveau d'un azote ou

d'un carbone de la molécule d'ADN. Le produit de cette méthylation est la 5-méthylcytosine (Figure 3).

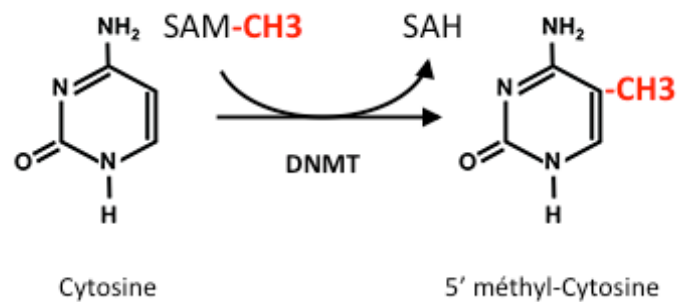


Figure 3 : méthylation de l'ADN au niveau d'un acide nucléique cytosine par transfert d'un groupe méthyle issu du cofacteur SAM, par l'action d'une ADN méthyle transférase (DNMT)

Dans l'organisme, environ 60 à 80% des cytosines des CpG sont méthylées. Les doublets CpG sont particulièrement fréquents dans certaines régions, appelées îlots CpG, où ils représentent environ 10% des nucléotides contre 1% dans les autres régions de l'ADN. Ces îlots sont principalement situés au niveau de l'extrémité 5' des gènes, riche en promoteurs. La méthylation des îlots CpG est souvent indicative d'une configuration de la chromatine qui réprime l'activité transcriptionnelle ¹⁸. Elle inhibe l'association de certains facteurs de transcription à leur site de liaison à l'ADN, ce qui diminue le recrutement et la formation de complexes protéiques impliqués dans la transcription, tels que les complexes histones/nucléosomes.

Ce mécanisme est très ancien et conservé. Il est impliqué dans différents phénomènes physiologiques tels que le vieillissement cellulaire ou la lutte contre les organismes étrangers en bloquant l'expression du génome viral ou bactérien. La méthylation peut être modifiée par des facteurs environnementaux (agents alkylants, perturbateurs du métabolisme) et certaines de ces modifications acquises peuvent être passées d'une génération à l'autre, comme le montre le cas des souris Agouti dont les progénitures présentent un phénotype

altéré, selon que les mères ont été exposées ou non à un régime hypervitaminique, malgré un génome identique d'un groupe de descendants à l'autre ¹⁹.

Dans les cellules mammifères, on retrouve deux types de méthylation : la méthylation dite « de maintenance » et la méthylation *de novo*. Cette modification est catalysée par les ADN méthyltransférases (DNMTs), groupe hétérogène d'enzymes composé de 4 membres : DNMTs 1, 2, 3a et 3b. La structure des DNMTs comprend une partie N-terminale qui permet la liaison à l'ADN et aux autres protéines, le site catalytique en C-terminal, et une séquence NLS (*nuclear localization signal*) qui permet le passage de l'enzyme dans le noyau de la cellule afin de méthyler l'ADN ^{20,21}.

La DNMT 1 intervient dans la méthylation de maintenance qui consiste en la reproduction du profil de méthylation du brin d'ADN néo-synthétisé au cours de la réplication de l'ADN qui a lieu pendant la division cellulaire. C'est une modification post-réplicationnelle.

La DNMT 2 a un rôle inconnu et a été identifiée du fait de son homologie de séquence avec la DNMT1.

Les DNMT 3a et 3b sont impliquées dans la méthylation *de novo* car elles ajoutent des nouveaux groupes méthyles sur les brins d'ADN. DNMT 3a est capable de méthyler les promoteurs de l'expression des gènes et DNMT 3b catalyse principalement des réactions de méthylation au niveau des séquences proches des centromères.

La méthylation de l'ADN est réversible, soit passivement par perte des marques de méthylation au cours des divisions cellulaires successives soit activement par l'action des ADN déméthylases, telles que les protéines TET (*ten-eleven translocation*). Cette famille de protéines comprend 3 membres et est impliquée dans le développement embryonnaire, la méiose, la maintenance de profils de méthylation au cours des divisions cellulaires ainsi que dans la reprogrammation, ou pluripotence induite, des cellules souches. Ces enzymes sont capables de convertir les 5-méthylcytosine en 5-hydroxyméthylcytosine ²². Le niveau d'hydroxyméthylation des cytosines semble corrélé à l'activité transcriptionnelle, ce qui suggère que ces enzymes participent à la régulation positive de l'expression génique. Leur activité semble influencée par les modifications des histones et les hybridations ADN/ARN.

Leur perte d'activité a, à l'inverse, été corrélée à l'augmentation de la méthylation dans les régions amplificatrices, ou « *enhancer regions* », au niveau de cellules souches, avec des modifications de l'expression des gènes associés ²⁰.

De multiples autres acteurs entrent en jeu dans la régulation de l'expression génique, en lien avec la méthylation de l'ADN. En effet, la méthylation des promoteurs favorise la liaison de complexes protéiques spécifiques de l'ADN méthylé, telles que les HDAC et induit la déacétylation de l'ADN en regard des zones méthylées, ce qui renforce l'inhibition de l'activité transcriptionnelle ²³. La famille des protéines à *methyl CpG binding proteins*, dont MecP2 (methyl CpG binding protein 2), interagissent également spécifiquement avec l'ADN méthylé. MecP2 contrôle la transcription de gènes principalement exprimés au niveau du système nerveux central, via le complexe répresseur Sin3A ^{24,25}. La protéine MBD1 promeut la méthylation de H3K9, réprimant ainsi indirectement la transcription. MBD2 semble impliquée dans la tumorigénèse ²⁶ et dans la différenciation des lymphocytes T helper lors de leur activation par les cellules dendritiques ²⁷.

MBD3 est un élément du complexe NuRD (*Mi-2/nucleosome remodeling and histone deacetylase*) et favorise la déméthylation de l'ADN ²⁸. MBD4, possède une activité ADN glycosylase et est également impliquée dans la déméthylation active de l'ADN ²⁹. MBD5 et MBD6 semblent être associées à l'hétérochromatine mais ne sont pas capables de se lier à l'ADN méthylé ³⁰.

Deux autres classes de protéines sont capables de s'associer à l'ADN méthylé : (i) les protéines UHRF1 et UHRF2, grâce à leur domaine SRA (*SET and RING finger-associated*) et (ii) les protéines Kaiso, ZBTB 4 (Zinc Finger and BTB Domain Containing 4), et ZBTB 38, contenant trois domaines à doigt de zinc. UHRF1 interagit avec les DNMTs 3a et 3b et participe ainsi à la méthylation *de novo* et de maintien de l'ADN ³¹. Kaiso, ZBTB4 et ZBTB38 sont associées à une répression de la transcription ³².

La dérégulation de la méthylation de l'ADN est impliquée dans la genèse de nombreuses pathologies notamment cancéreuses. Les principaux mécanismes sont l'hyperméthylation des promoteurs des gènes suppresseurs de tumeurs, qui survient dans un environnement

globalement hypométhylé, ou l'hypométhylation des promoteurs des oncogènes³³. En effet, des mutations de la DNMT 3a sont retrouvées dans 25% des cas de leucémies aigües myéloïdes, et la mutation inactivatrice de TET2 est impliquée dans plusieurs syndromes myéloprolifératifs^{34,35}. De même, des mutations de la protéine MBD4 sont présentes dans 40% des cancer colorectaux avec instabilité microsatellitaire par déficience du *mismatch repair* sporadique³⁶.

Ainsi, la méthylation de l'ADN est un des mécanismes de régulation de l'expression génique qui participe à l'homéostasie cellulaire et au maintien de l'information épigénétique au fil des divisions cellulaires et des générations successives, avec une implication potentielle dans de nombreuses pathologies. Cependant, l'ADN n'est pas le seul acide nucléique dont la méthylation participe à la régulation de l'expression génique. L'ARN peut également être méthylé, avec des conséquences sur sa stabilité et son activité, ce qui constitue un des différents mécanismes de régulation post-chromatiniens.

B. RÉGULATION POST-TRANSCRIPTIONNELLE

1. La méthylation de l'ARN

La méthylation de l'ARN est la modification post-transcriptionnelle la plus fréquente chez les eucaryotes. Elle survient principalement au niveau des nucléosides, avec plus de 70 types de modification décrits. La plus fréquente est la transformation de l'adénosine en N6-méthyladénosine (m6A)³⁷. La méthylation de l'ARN semble impliquée dans de nombreux processus physiologiques et pathologiques mais les différents mécanismes qui la régulent restent encore mal caractérisés. Un article récent a montré que la triméthylation de l'histone H3 au niveau de sa lysine 36 (H3K36me3) favorise la méthylation m6A de l'ARN. En effet, les cellules avec une forte concentration de H3K36me3 présentent un taux élevé de m6A alors qu'en cas de déplétion en H3K36me3, des concentrations faibles de m6A sont mesurées. H3K36me3 est reconnue et liée par METTL14 (*Methyltransferase Like 14*), la

sous-unité catalytique du complexe N6-adenosine-méthyltransferase (N6-CMT). Cette interaction facilite la liaison du N6-CMT à l'ARN polymérase II, permettant ainsi une méthylation m6A des nouveaux ARN transcrits pendant les phases d'élongation et de transcription³⁸. La découverte de l'interaction entre les histones méthylées et la méthylation de l'ARN représente un nouveau mécanisme de régulation et un niveau supplémentaire de complexité dans le contrôle de l'expression des gènes.

La méthylation de l'ARNm et de l'ARN long non-codant sont les phénomènes les mieux décrits, mais les modifications de l'ARN ribosomal, de l'ARN de transfert et d'autres ARNs non-codants sont également étudiées de manière active.

Des marques de méthylation au niveau de la queue poly (A) en 3' et de la coiffe en 5' des ARNm ont aussi été décrites et participent à la régulation du métabolisme des ARNm. La modification de ces structures présentes sur la grande majorité des molécules d'ARN, non codées par l'ADN, va influencer la stabilisation des ARNm, leur migration du noyau vers le cytoplasme et l'initiation de la traduction. Ainsi, la coiffe 5' peut être méthylée par des méthyltransférases nucléaires spécifiques au niveau de la position N-7, ce qui renforce la stabilité de l'ARN dans les cellules eucaryotes et les virus. Elle peut également être méthylée en 2'-O, ce qui semble impliqué dans les mécanismes d'échappement des virus face à l'activité antivirale de l'immunité innée de l'hôte³⁹.

2. Épissage alternatif

Au début des années 2000, avant la publication du génome humain complet, on estimait le nombre de gènes autour de 300 000, alors qu'il est d'environ 22 000. Il n'est donc pas beaucoup plus riche en gènes que celui d'organismes plus simples tels que le nématode, le poisson-zèbre ou la souris. La taille du génome n'est donc pas un bon reflet de la complexité biologique d'un être vivant. L'épissage alternatif des ARN pré-messagers représente une étape fondamentale de l'expression des gènes et participe à l'amplification de la complexité et de la variabilité du transcriptome, en augmentant le nombre de transcrits possible par

gène grâce à l'utilisation alternative d'exons ou d'introns, mais aussi de signaux de polyadénylation situés dans des exons alternatifs. Différents mécanismes d'épissage alternatif sont décrits tels que l'exon-cassette, les exons mutuellement exclusifs, la rétention d'intron, les sites d'épissage alternatif en 5' (site donneur d'épissage) ou 3' (site accepteur d'épissage) ⁴⁰.

L'épissage alternatif est une source majeure de diversité de transcrits et du protéome. C'est donc un mécanisme de régulation à la fois qualitative et quantitative de l'expression des gènes. D'autre part, l'épissage peut aboutir à un variant d'ARN « non-sens » reconnu et dégradé par le système NMD (*non-sense mediated decay*) (Figure 4). Le NMD est responsable du contrôle qualité des ARNm, afin que soient traduits uniquement les ARNm correctement épissés et notamment dépourvus d'introns. La présence de ces derniers, de par leur longueur, induirait des modifications structurelles et fonctionnelles majeures des protéines résultant de la traduction de l'ARNm concerné. En effet, lors de l'épissage, le complexe de jonction exonique reste associé au niveau des jonctions exon-exon, ce qui marque l'absence d'intron. Lors de la traduction, le ribosome va dans un premier temps enlever ces complexes de jonction exonique et s'arrêter naturellement au codon STOP final. Au cours de l'épissage alternatif d'un transcrit, un codon STOP peut apparaître prématurément. Ce dernier sera alors reconnu par le NMD, entraînant ainsi la dégradation de l'ARNm par des exonucléases. Ainsi l'inclusion ou l'exclusion d'un exon peut entraîner la dégradation d'un ARN entier.

Un gène ayant en moyenne 10 exons qui peuvent chacun être épissés de multiples façons différentes, l'épissage alternatif amplifie donc considérablement la diversité du transcriptome de part la variabilité combinatoire qu'il induit. Les récentes études de transcriptomique montrent que des cellules issues d'une même population cellulaire expriment les mêmes gènes, mais avec des variants différents d'épissage et donc un protéome différent. Ces transcriptomes seront néanmoins plus proches qu'entre des cellules issues de populations différentes. On parle de plasticité du transcriptome, associée à la plasticité phénotypique des

cellules, Cette notion suggère donc qu'aucune cellule d'un même organisme n'exprime le même transcriptome et protéome.

L'épissage alternatif est impliqué dans le développement et la différenciation des différents tissus chez les organismes eucaryotes et sa dérégulation est impliquée dans de multiples pathologies suite à des troubles du développement, telles que des pathologies neuromusculaires, neuropsychiatriques ou cardiovasculaires ⁴¹. L'épissage alternatif est également impliqué dans la carcinogenèse par divers mécanismes qui augmentent l'expression des oncogènes ou inhibent l'expression de gènes suppresseurs de tumeur ⁴²⁻⁴⁴. Par exemple, dans le carcinome rénal à cellules claires, une augmentation de la rétention d'introns est observée au niveau du gène SETD2 (SET domain containing 2), diminuant ainsi la production de la lysine-méthyltransférase et induisant une diminution de la triméthylation de H3K36 dont elle est responsable. Cela aboutit à une augmentation de l'accessibilité de la chromatine et de la transcription de l'expression des gènes des régions régulées par l'histone 3 ⁴⁵.

L'épissage alternatif est donc un mécanisme essentiel de régulation de l'expression génique, permettant une très grande variabilité phénotypique et impliqué dans de nombreux processus physiologiques et pathologiques. La production des ARN messagers et leur traduction sont également contrôlées par d'autres mécanismes complexes tels que la production d'ARNs non-codants.

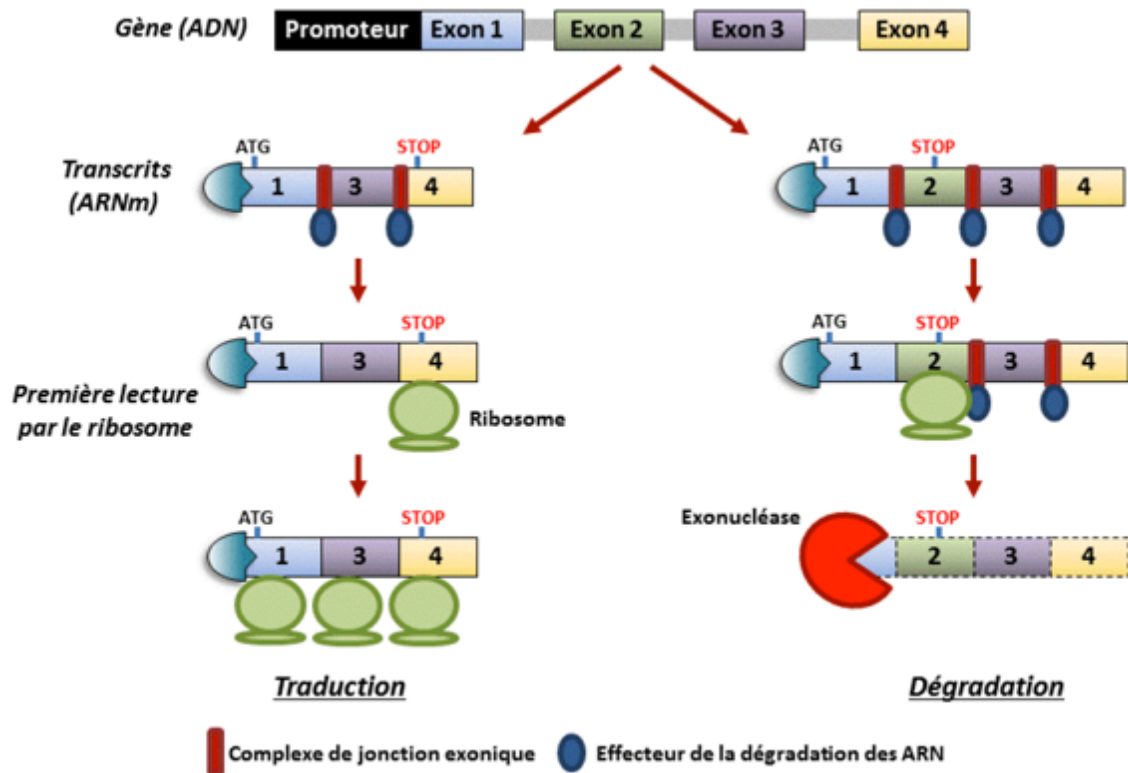


Figure 4 : Epissage alternatif et dégradation (Samaan et al. Pour la Science 2013).

3. Les ARNs non-codants

Les ARNs non-codants (ncRNA) correspondent aux molécules d'ARN qui ne sont pas traduites en protéines. Les types d'ARN non-codant abondants et importants sur le plan fonctionnel comprennent les ARNs de transfert (tRNA) et les ARNs ribosomiaux (rRNA), ainsi que les petits ARNs tels que les microARNs (miRNA), les petits ARNs associés aux protéines piwi (piRNA), spécifiques des cellules germinales, les petits ARNs nucléolaires (snoRNA), les petits ARNs nucléaires (snRNA), et les longs ARNs non codants intergéniques (lincRNA).

L'importance des ARNs non-codants a longtemps été sous-estimée mais on sait à l'heure actuelle qu'il en existe des milliers, notamment grâce aux études transcriptomiques et bioinformatiques récentes. De nombreux ncRNAs ont une fonction encore inconnue et il est probable qu'une proportion importante ne soit pas fonctionnelle ⁴⁶.

Le mécanisme de répression de la transcription par ARN implique la transformation d'ARNs

double brin aberrants en petits ARNs simple brin de 21 à 26 nucléotides de long. La répression post-transcriptionnelle des gènes se fait principalement dans le cytoplasme par un procédé relativement bien conservé. La ribonucléase DICER va générer des ARNs simple brin en coupant en petits fragments l'ARN double brin, puis ces ARNs simple brin vont ensuite être reconnus par les protéines Argonaute et intégrés au sein du complexe RISC (*RNA induced silencing complex*). Cet ARN simple brin va s'associer à un ARN messenger spécifique par complémentarité exacte de séquence et entraîner sa dégradation avant qu'il puisse être traduit en protéine ⁴⁷.

Les miRNAs utilisent aussi le système de répression DICER/RISC pour réprimer au niveau post-transcriptionnel un certain nombre de gènes afin de réguler leur expression. Au niveau du génome, ces miRNAs sont souvent regroupés et vont être exprimés sous forme de transcrits polycistroniques (pre-miRNA). Chaque miRNA va dans un premier temps être clivé par la ribonucléase Drosha au niveau du noyau, sous forme de structure en épingle, et va ensuite être intégré au système DICER/RISC. La complémentarité des miRNAs avec une cible n'est pas de 100%, du fait de leur structure en épingle, ce qui leur offre un plus large panel d'ARNm cibles. La reconnaissance se fait sur la région 3'-UTR de l'ARN messenger ce qui n'induit pas de clivage mais va inhiber sa traduction en bloquant le recrutement du ribosome 60S ⁴⁷.

Les ARN double brin semblent également pouvoir réprimer des gènes cibles en favorisant la méthylation de leur promoteur et ainsi agir au niveau transcriptionnel ⁴⁸. L'importance de ce phénomène a été démontrée à de nombreuses reprises chez la levure avec la création de mutants altérés pour un des composants de cette machinerie RNAi, chez qui l'on retrouvait, de manière systématique, une baisse des marques de méthylation de l'hétérochromatine et une hausse de la transcription de certains loci.

Les ncARNs sont également impliqués dans la régulation des gènes soumis à l'empreinte parentale. En effet, la méthylation de l'ADN d'un des deux allèles est corrélée à la présence de ncRNAs ciblés dans la même région. L'inhibition des transcrits des ncRNAs dans les régions d'empreinte parentale a un effet délétère et est impliquée dans certaines pathologies

Les piRNAs, longs de 29 nucléotides, sont importants dans le contrôle des transposons⁵⁰. Chez le zebrafish *Danio rerio*, les piRNAs ont été impliqués dans la maintenance des cellules germinales et beaucoup d'entre eux semblent ciblés sur les transposons, laissant suggérer un rôle de répression des éléments répétés chez les vertébrés⁵¹.

Enfin, les longs ARNs non-codants agissent à travers plusieurs mécanismes pour réguler l'expression des gènes codants. Ils peuvent influencer la distribution des modificateurs de la chromatine vers des gènes cibles spécifiques, ils entrent en compétition avec des microARNs et protègent les ARN messagers cibles de la dégradation, ils modulent des événements d'épissage et interfèrent avec la traduction, ils promeuvent l'association des séquences amplificatrices et promotrices pour initier la transcription⁵².

L'ensemble de ces mécanismes épigénétiques aboutit à une très grande variété de protéines traduites, avec parfois de nombreuses isoformes possédant des fonctions ou des niveaux d'activité différents. Ensuite, au cours de leur maturation, ces protéines vont subir des modifications supplémentaires, appelées modifications post-traductionnelles (PTMs), qui vont influencer leur stabilité, leur activité et leurs interactions avec d'autres protéines, ou parfois provoquer leur dégradation.

C. LES MODIFICATIONS POST-TRADUCTIONNELLES

Les PTMs constituent un troisième niveau de régulation de l'expression phénotypique des gènes. Elles correspondent à l'ensemble des modifications chimiques au niveau des protéines survenant au sein des cellules après la traduction, qui consiste en la polymérisation d'acides aminés par les ribosomes à partir des ARNs messagers (ARNm). Ces PTMs interviennent dans la maturation des protéines et ont un impact majeur sur leur stabilité, leur localisation et la régulation de leur activité. Les PTMs les plus fréquemment observées sont la phosphorylation, l'acétylation et la N-glycosylation. Néanmoins de nombreuses autres PTMs sont décrites telles que l'amidation, l'hydroxylation, la méthylation,

l'ubiquitylation, etc.⁵³ (Figure 5). Parmi les 20 acides aminés, la lysine constitue un exemple d'acide aminé capable d'accueillir une grande variété de PTMs. Ces dernières peuvent être divisées en 3 catégories : (i) l'acétylation⁵⁴ et les PTMs apparentées telles que la crotonylation⁵⁵, la propionylation, la butyrylation⁵⁶, la succinylation⁵⁷, la malonylation⁵⁸, la formylation⁵⁹ ; (ii) l'ubiquitinylation et les PTMs dites "*ubiquitinyl-like*" telles que la SUMOylation, l'ISGylation et la NEDDylation⁶⁰ ; (iii) d'autres modifications telles que l'hydroxylation⁶¹, la glycation⁶² et la méthylation^{63,64} et la phosphorylation.

De nombreuses études démontrent l'importance grandissante des modifications post-traductionnelles propres aux signalisations épigénétiques et cellulaires. Les PTMs interviennent dans les principales étapes de la physiologie cellulaire et sont une source majeure de diversité fonctionnelle du protéome des mammifères^{4,65}. De nombreux processus de signalisation cellulaire reposent sur les PTMs et leur caractère souvent réversible, qui permet une régulation très dynamique des différents processus cellulaires physiologiques ou pathologiques.

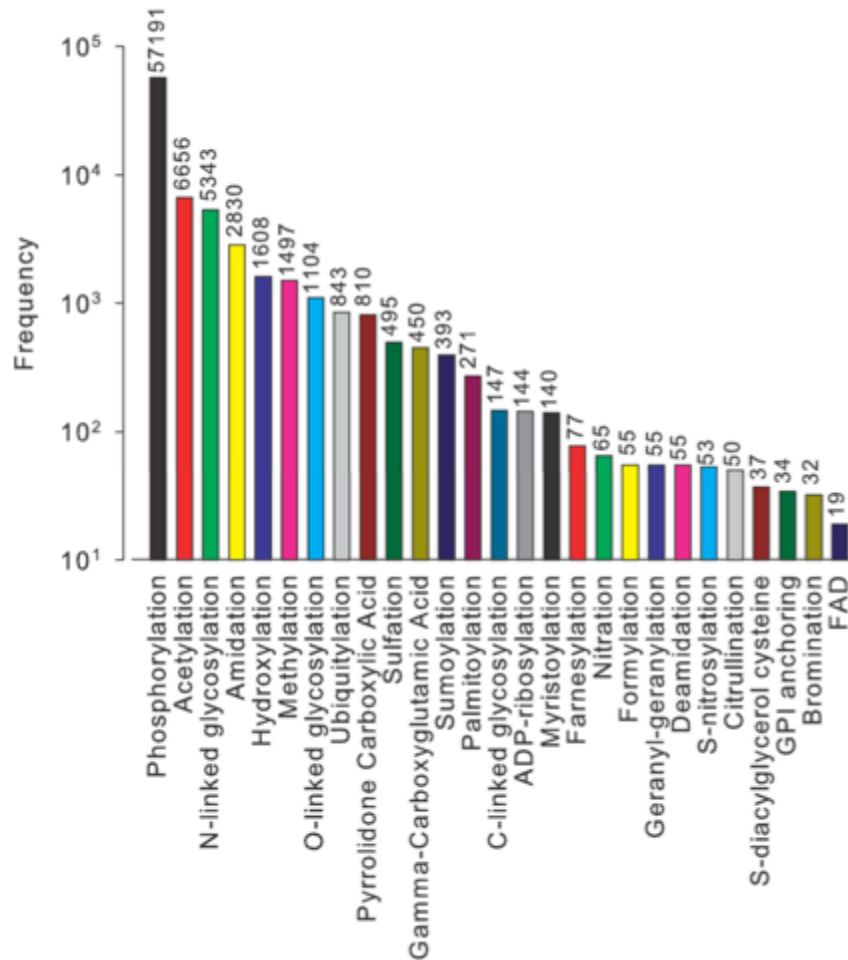


Figure 5 : PTMs observées expérimentalement à partir de la base de données Swiss Prot. Khoury *et al.* Sci Report 2011.

L'ubiquitylation, parmi tant d'autres exemples, illustre parfaitement le niveau de précision avec lequel ces PTMs doivent être régulées et les multiples conséquences biologiques qui découlent des différentes marques possibles sur une même protéine. En effet, il peut s'agir de polyubiquitylation, dont le principal rôle est d'induire la dégradation protéolytique, ou de monoubiquitylation, qui semble plutôt stabiliser les protéines et est impliquée dans la régulation de l'histone, l'endocytose et la pénétration de la membrane cytoplasmique par les rétrovirus. En effet, la monoubiquitylation des protéines au niveau d'un groupe Lys-e-NH₂ par une ubiquitine-ligase E3 peut initier la relocalisation des protéines des récepteurs transmembranaires de la membrane cytoplasmique vers la face *trans* du réticulum de Golgi

⁶⁶. Elle semble impliquée dans le tri des protéines qui sont ensuite être liées, via leur ubiquitine, par différentes protéines partenaires. Ces partenaires vont alors agir comme des protéines chaperonnes en masquant cette marque d'ubiquitylation et en augmentant ainsi la stabilité de la protéine qui sera dégradée dans le lysosome. À l'inverse, la polyubiquitylation, notamment en cas de fixation de 4 ubiquitines sur la chaîne latérale d'une lysine, induit la liaison à des protéines chaperonnes qui escortent la protéine modifiée vers le protéasome pour une dégradation rapide ⁶⁷.

Les PTMs se produisent généralement dans des régions protéiques dites «régulatrices» où les acides aminés sont plus accessibles et donc plus facilement modifiables par des enzymes. On parle de régions de désorganisation intrinsèque ou *intrinsically disordered regions* (IDR) des protéines. Ces zones sont considérées comme flexibles car souvent dépourvues de structures secondaires, à distance du noyau hydrophobe des protéines. Un nombre accru de PTMs se situent donc dans ces zones, comme démontré par un certain nombre de travaux utilisant notamment la spectroscopie par résonance magnétique nucléaire ⁶⁸. Ceci s'explique peut-être par le fait que les signalisations de réponse cellulaire nécessitent des enzymes de modification qui doivent pouvoir accéder rapidement aux sites des PTMs. Certaines PTMs telles que la phosphorylation sont capables d'impacter cet état de désorganisation en provoquant une transition de l'état désorganisé vers un état organisé (*disorder-to-order transition*) et ainsi influencer la survenue d'autres PTMs à proximité ⁶⁹. Par ailleurs, les zones désorganisées sont particulièrement riches en sérines et thréonines potentiellement phosphorylables ⁷⁰. La phosphorylation est la modification post traductionnelle la plus abondante dans les cellules et représente une étape clé dans la transduction des signaux eucaryotes. Cette PTM est critique dans de multiples processus tels que le métabolisme, le cycle cellulaire, le développement, la signalisation hormonale et la perception des agents pathogènes. Ainsi, outre l'impact des PTMs sur les interactions protéine-protéine par l'induction de nouvelles interactions ou l'inhibition d'interactions existantes, la phosphorylation illustre le fait que les PTMs peuvent également générer une

gamme de modifications structurelles qui régulent la fonctionnalité des protéines et leurs interactions ⁷¹.

Avec l'épissage alternatif, les PTMs renforcent la diversité biologique issue du protéome dans de multiples domaines tels que la communication inter-cellules, la croissance et la différenciation cellulaires, la détection des différents états métaboliques, la médiation du transport intracellulaire et l'initiation de la mort cellulaire programmée ⁷². Les dérégulations de ces PTMs ou de leur lecture par les interacteurs spécifiques, qu'elles soient héréditaires ou secondaires à des facteurs environnementaux, représentent une importante source de mécanismes physiopathologiques impliqués dans de nombreuses pathologies humaines telles que les maladies cardiovasculaires, neurologiques, endocriniennes et métaboliques, ainsi que dans le cancer.

Dans cette introduction, nous aborderons de manière plus détaillée la méthylation et son rôle dans la régulation des fonctions protéiques et des interactions protéine-protéine, ainsi que ses implications dans les phénomènes de carcinogenèse.

II- LA METHYLATION DES PROTEINES

A. GÉNÉRALITÉS

La méthylation correspond à l'ajout par liaison covalente d'un groupe méthyle à un substrat en présence du cofacteur S-adénosyl méthionine (SAM), un des cofacteurs les plus abondants dans l'organisme. Le SAM est synthétisé dans le cytoplasme de toutes les cellules et le foie constitue un acteur essentiel dans l'homéostasie de ce dernier, en étant un important site de synthèse et de dégradation du SAM. Le SAM est impliqué dans 3 principales réactions : la trans-méthylation, la trans-sulfuration et la synthèse des polyamines⁷³. Comme précédemment mentionné, la méthylation est un mécanisme dynamique et réversible, finement régulé, qui répond au concept de « *writers, erasers and readers* ». Elle est catalysée par les méthyltransférases, et concerne différents substrats tels que l'ADN, l'ARN et les protéines.

La méthylation des protéines est l'une des PTMs les plus diversifiées du point de vue fonctionnel. Elle consiste en l'adjonction d'un groupe méthyle par liaison covalente sur un acide aminé. De nombreux acides aminés peuvent être méthylés, à savoir : les acides aspartique ou glutamique, l'alanine, l'arginine, l'asparagine, l'histidine, la lysine et la méthionine. Les protéines peuvent également être méthylées au niveau de leur terminaison amine (N-Ter) ou carboxyle (C-Ter). Néanmoins, les deux principaux acides aminés concernés sont la lysine et l'arginine. Cette modification est réalisée par les lysine-méthyltransférases (KMTs) ou les arginine-méthyltransférases (PRMTs). Plusieurs niveaux de méthylation sont possibles, à savoir la mono-, di-, et tri-méthylation au niveau des lysines⁷⁴, et la mono- et di-méthylation (symétrique ou asymétrique) au niveau des arginines selon les capacités de la méthyl-transférase impliquée et le site de méthylation. Ce niveau de méthylation aura des conséquences sur l'activité du substrat, notamment via la modification de ses interactions avec d'autres protéines. Enfin, la méthylation des protéines est un

phénomène dynamique et réversible grâce à l'action des lysine-déméthylases (KDMs) ou des arginine-déméthylases.

Cette PTM a d'abord été décrite dans la régulation de l'activité des histones comme mécanisme de régulation de la dynamique de la chromatine et de l'expression génique. Ce domaine de recherche a largement dominé la littérature relative à l'étude de la méthylation des protéines pendant de nombreuses années ⁷⁵. Plus récemment, de nombreuses cibles indépendantes de la chromatine ont été décrites comme méthylées, notamment par des MTs initialement décrites comme des histones méthyltransférases, ce qui rend de plus en plus difficile la distinction entre les MTs méthylant exclusivement les histones dites histone-méthyltransférases, et celles capables de méthyler des cibles non-histones. Ainsi, dans la réponse aux lésions de l'ADN par exemple, la méthylation est impliquée dans l'activité de protéines non-histones telles que p53, ou PARP1 (Poly(ADP-ribose) polymérase 1) ⁷⁶.

En 2011, Petrossian et Clarke ont proposé une classification des méthyltransférases humaines en utilisant les structures enzymatiques disponibles. Grâce à des outils bioinformatiques, ils ont identifié 208 MTs dépendantes du cofacteur SAM, dont 30% étaient déjà identifiées comme impliquées dans diverses pathologies et cancers ⁷⁷. Selon leur étude, la majorité des MTs humaines possèdent l'un des trois domaines catalytiques suivants : un domaine *7β-strand* (131 MTs), un domaine SET (57 MTs) ou un domaine SPOUT (8 MTs). Treize MTs possédant un domaine n'appartenant pas à ces 3 familles ont également été décrites.

La méthylation des protéines peut avoir différentes fonctions, que ce soit en tant que PTM indépendante ou en interagissant avec d'autres PTMs ⁷⁸. Par ailleurs, la méthylation ne modifie pas profondément la structure de la protéine car elle représente un faible encombrement stérique et n'affecte pas la charge de l'acide aminé modifié. Toutefois, comme les arginines ou lysines peuvent être méthylées jusqu'à deux ou trois fois, cet événement peut affecter leur capacité à agir comme donneur de liaison hydrogène [6]. Ainsi, la méthylation des protéines est un régulateur de la conformation et des interactions

protéine-protéine, ce qui a des conséquences sur la stabilité des protéines, leur localisation cellulaire et leur activité, soit en tant qu'acteur isolé soit via un complexe spécifique.

B. LA MÉTHYLATION DES ARGININES

La méthylation des arginines intervient dans les signalisations cellulaires et épigénétiques telles que la transcription, l'épissage de l'ARN, la traduction, les voies de signalisation activées par les lésions de l'ADN et la détermination du destin cellulaire ⁷⁹.

Les arginines ont une charge positive et possèdent une chaîne latérale avec un groupe guanidine plan qui comporte 5 sites potentiels donneurs de liaisons hydrogène capables de lier des molécules chargées négativement, tels que les acides nucléiques. L'ajout d'un groupe méthyle ne modifie pas la charge globale de l'arginine, mais sa présence sur l'un des sites de liaison modifie les interactions de l'arginine avec ses partenaires habituels.

La famille des arginine-méthyltransférases (PRMTs) est constituée de 9 membres ⁸⁰. Toutes les PRMTs sont capables de catalyser la monométhylation des arginines, permettant ainsi l'obtention d'une omega-N-méthylarginine. La grande majorité des PRMTs génèrent ensuite une diméthylation, à l'exception de PRMT7 qui est la seule arginine mono-méthyltransférase exclusive connue à ce jour ⁸¹. La mono-méthylation des arginines est, chez les mammifères, un événement aussi fréquent que l'ubiquitinylation et la phosphorylation ⁸². On distingue donc trois catégories de PRMTs: les PRMTs de type 1 générant une di-méthyl-arginine asymétrique (PRMT1, PRMT4 aussi appelée CARM1 (*coactivator-associated arginine methyltransferase 1*), PRMT6, et PRMT8), et les PRMTs de type 2 générant une di-méthyl-arginine symétrique (PRMT5, PRMT9). PRMT7 est la seule monométhyltransférase ou PRMT de type 3 (Figure 6).

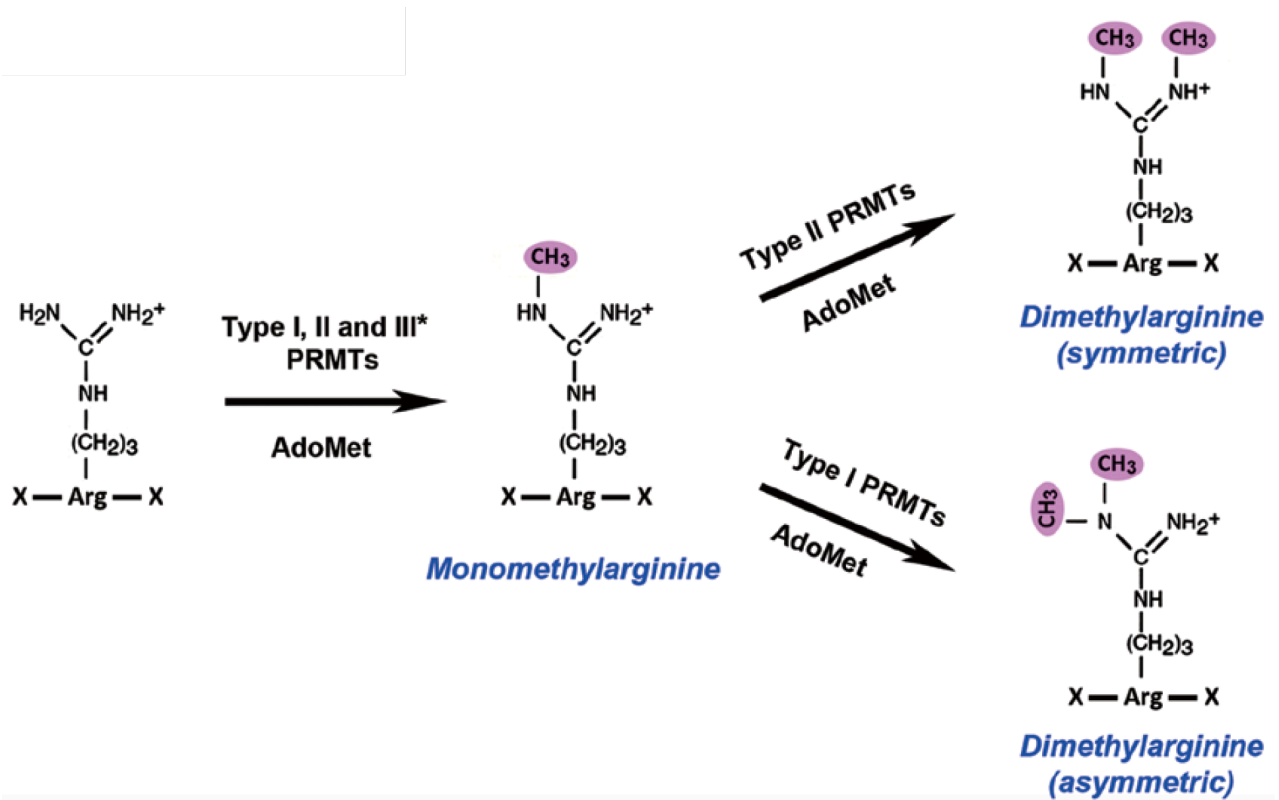


Figure 6 : Différents états de méthylation des arginines. (Wang *et al.* Cell Research 2014). La méthylation passe dans un premier temps par un état monométhylé avant d'aboutir à un état diméthylé soit symétrique, en cas d'ajout d'un groupe méthyle sur le même atome d'azote, soit asymétrique lorsque le groupement méthyle est ajouté sur l'azote opposé.

Cette PTM est particulièrement énergivore, avec l'utilisation de plus de 12 molécules d'ATP pour une méthylation. Elle a néanmoins été conservée au cours de l'évolution ce qui confirme son caractère essentiel ⁸³.

La méthylation des arginines survient sur des sites préférentiels, correspondant à des motifs bien particuliers. En effet, il a été démontré que les marques de méthylation des résidus arginines survenaient principalement au niveau des arginines proches de résidus glycine, qui permettent une conformation particulière favorisant l'événement de méthylation. Cette connivence est notamment présente au niveau des motifs RG et RGG qui, par définition, sont riches en arginines et en glycines ⁸⁴. Ces motifs sont décrits dans plus de 2000 protéines et sont impliqués dans la liaison à l'ADN et dans les interactions protéine-protéine. D'autres motifs favorisent l'action de certaines PRMTs, tels que les motifs PGM (proline,

glycine et méthionine) pour CARM1 (Coactivator Associated Arginine Methyltransferase 1)/PRMT4⁸⁵, ou encore des motifs RxR au sein des régions riches en lysines pour PRMT7⁸⁶.

La principale conséquence de la méthylation des arginines est la perte d'interaction du substrat méthylé avec ses interacteurs secondaires à une modification conformationnelle due à l'effet stérique de l'ajout du groupe méthyle. Ce changement de conformation interfère avec la liaison de l'arginine à ses interacteurs via des ponts hydrogène⁸⁷. Outre cet effet inhibiteur sur les interactions protéine-protéine, la marque de méthylation peut également favoriser les interactions des substrats méthylés avec les protéines possédant un domaine Tudor. Ce domaine a été décrit sur seulement 36 protéines alors que plus de 2000 protéines présentent un site de méthylation d'arginine potentiel. Il est donc probable que d'autres domaines de liaison aux arginines méthylées restent à identifier⁷⁹.

L'importance de la méthylation des arginines dans la régulation de l'expression génique via la régulation des histones est bien connue. Ainsi, un certain nombre de marques répressives, telles que H3R2me2a, H3R8me2a, H3R8me2s, H4R3me2s et activatrices, telles que H4R3me2a, H3R2me2s, H3R17me2a, H3R26me2a, ont été décrites⁸⁰.

La méthylation d'autres protéines non-histones est impliquée dans la régulation de la transcription, telles que RelA R30, un substrat de PRMT1 qui appartient à la voie TNF α /NF- κ B⁸⁸, ou l'acétyltransférase p300, diméthylée par CARM1⁸⁹. La méthylation au niveau des arginines des protéines non-histone est aussi impliquée dans d'autres voies de signalisation cellulaire et dans un certain nombre de pathologies⁹⁰.

La méthylation des arginines est réversible, mais ce phénomène est moins dynamique et moins bien démontré que la déméthylation des lysines⁸⁵. En effet, les peptidyl-arginines déiminases (PAD) sont capables d'hydrolyser le groupe guanidine méthylé au niveau de la chaîne latérale des arginines, mais il ne s'agit pas à proprement parler d'une réversibilité de la méthylation.

La déméthylation des histones est catalysée par un groupe spécifique d'enzymes appelées histone déméthylases à domaine Jumonji (JmJD), qui ont une activité oxygénase et qui sont

capables de déméthylater les arginines et les lysines⁹¹. Néanmoins, concernant les arginines, ce phénomène reste encore peu caractérisé. JmjD6, par exemple, était initialement considérée comme une arginine déméthylase, puis a été décrite comme une lysine hydroxylase impliquée dans l'épissage de l'ARN⁹². Aussi, certaines lysines déméthylases semblent avoir une activité d'arginine déméthylation telles que KDM3A, KDM4A ou KDM5C⁹³.

C. LA MÉTHYLATION DES LYSINES

La méthylation des lysines est connue depuis les années 1960⁹⁴. A l'opposé de la méthylation des arginines qui comprend un nombre limité d'acteurs décrits, la méthylation des lysines fait intervenir de nombreuses enzymes, avec plus de 60 KMTs potentielles et 30 KDMs rapportées à ce jour.

1. La méthylation des lysines : une signalisation dynamique et réversible

a) Les lysines méthyltransférases (KMTs)

Les lysine-méthyltransférases sont des enzymes capables d'ajouter un, deux ou trois groupements méthyle au niveau des acides aminés lysines (Figure 7). L'ajout d'un groupe méthyle augmente le caractère hydrophobe et basique du résidu concerné⁷⁸, mais il induit un changement de poids moléculaire et de charge très modeste, ce qui rend par ailleurs cette modification difficile à détecter par les méthodes physico-chimiques classiques. L'état mono-, di- ou tri-méthylé des lysines impacte les interactions du substrat avec les autres protéines, selon que la présence d'un ou plusieurs groupe méthyles favorise ou inhibe la liaison du substrat à ses partenaires potentiels.

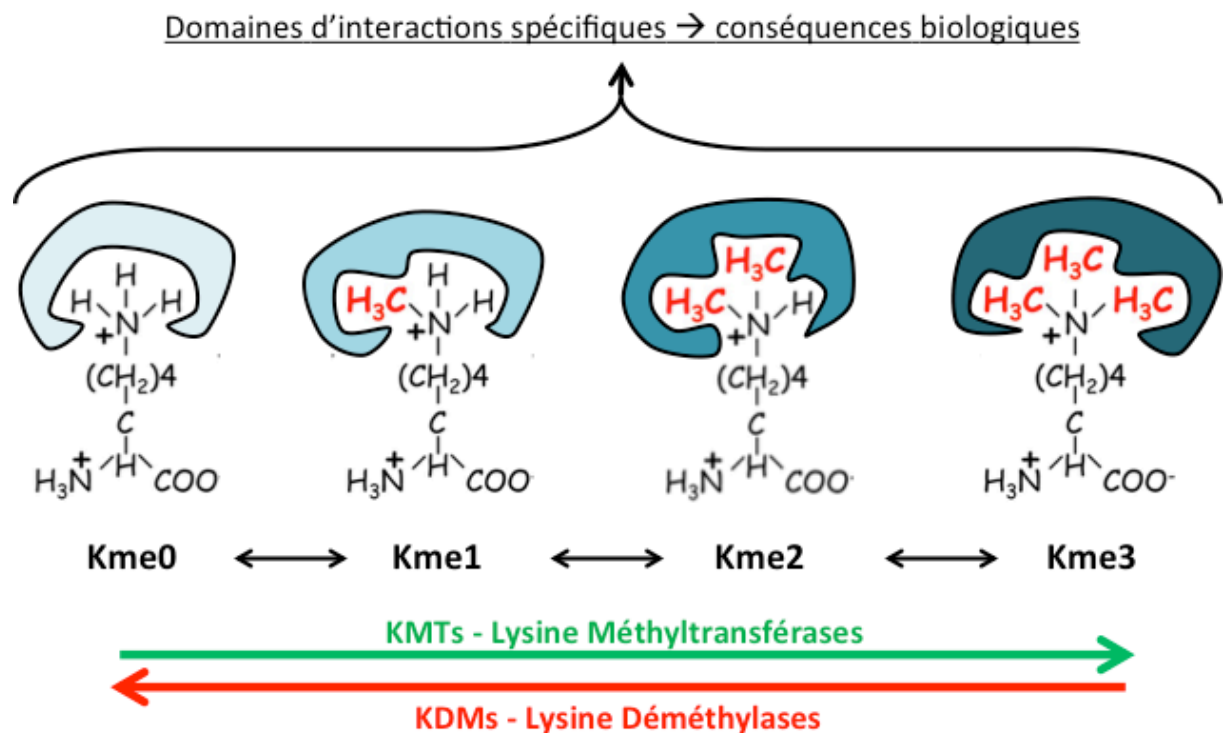


Figure 7 : Acteurs de la signalisation par méthylation des lysines

Le rôle des KMTs en tant que régulateurs de l'activité fonctionnelle de la chromatine par la méthylation des histones est bien établi ⁹⁵, mais ce processus concerne également des protéines non-histones clairement impliquées dans les cancers comme p53 ⁹⁶. De ce fait, le dysfonctionnement des KMTs est incriminé dans de nombreux phénomènes pathologiques et notamment dans la carcinogenèse de nombreux cancers ⁷⁸.

On distingue deux grands groupes de lysine méthyltransférases, selon leur type de domaine catalytique : les KMTs contenant un domaine SET et celles porteuses d'un domaine 7β-strand (7BS).

- Les KMTs à domaine SET

Les KMTs à domaine SET sont les mieux caractérisées et les plus nombreuses, avec plus de 50 méthyltransférases contenant un domaine SET pour SU(var), Enhancer of Zeste and Trithorax, correspondant aux noms des 3 premières enzymes décrites chez la drosophile contenant ce domaine ⁹⁷. Le domaine SET correspond à une séquence de 130 acides aminés et est composé de 3 régions pré-SET (ou N-SET), SET et post-SET (ou C-SET) qui

forment un nœud de feuillets beta (Figure 8) contenant 4 motifs très conservés GXG, (X correspondant le plus souvent à un résidu volumineux et hydrophobe), YXG, NHXCXPN et ELXFDY⁹⁸. Schématiquement, le SAM et le substrat se placent de part et d'autre d'un canal appelé le *Target lysine access channel*, ou canal d'accès à la lysine cible qui correspond à une structure pseudo-nodale ou *knot-like structure* (Figure 8). Ce dernier est composé d'une tyrosine très conservée, qui semble essentielle au transfert du groupe méthyle, puisque sa substitution par une phénylalanine abolit l'activité de méthylation de KMTs à domaine SET⁹⁹. Il est également le siège d'un réseau de résidus aromatiques et de liaisons hydrogène qui influencent l'orientation du substrat afin de rendre la lysine accessible à la méthylation^{63,100}. En effet, les interactions entre le site catalytique SET, le SAM et le résidu lysine du substrat sont complexes. Le motif GXG et le motif NHXCXPN via son asparagine et son histidine lient le cofacteur SAM par le biais de liaisons hydrogène et d'interactions de van der Waals. De plus, un résidu de charge positive, très conservé, au niveau du site de liaison au cofacteur (*cofactor-binding site*) forme un pont salin avec le groupe carboxyle du SAM. Enfin, la chaîne latérale d'un résidu aromatique des régions SET ou post-SET interagit avec la portion adénine du SAM. L'ensemble de ces interactions induisent une modification conformationnelle du SAM dite « en fer à cheval », positionnant son groupe méthyle labile dans le *Target lysine access channel* afin de méthyler le substrat.

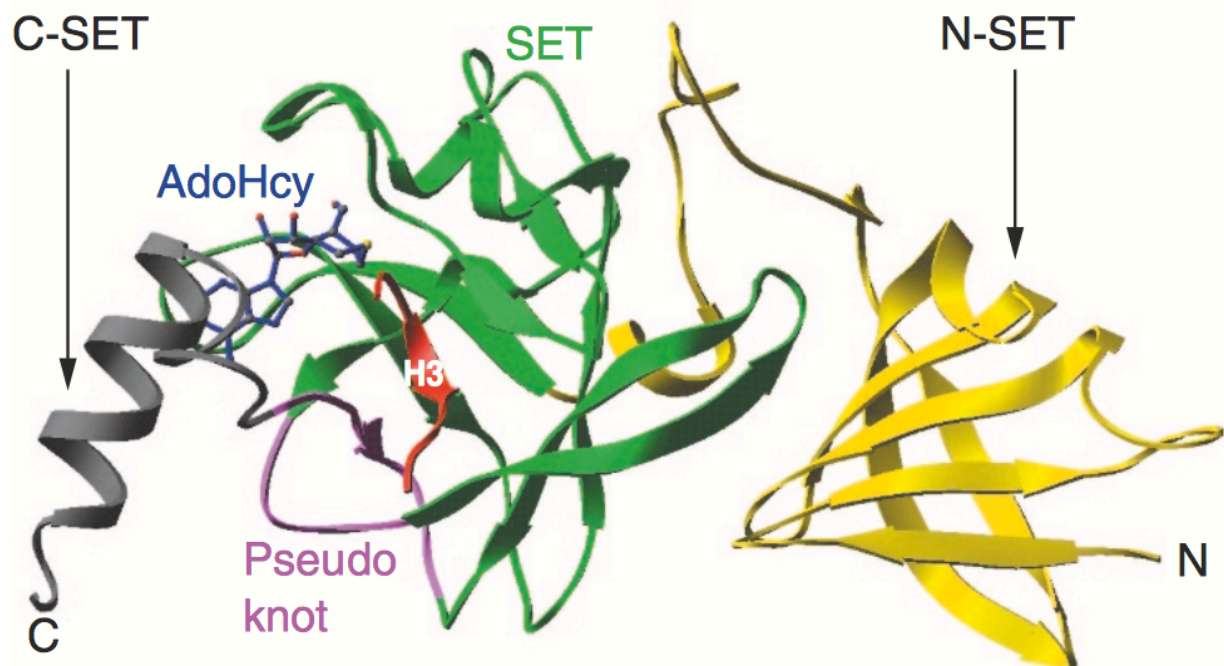


Figure 8 : Exemple de la structure de la KMT SET7/9 humaine méthylant l'histone 3 (Dillon et al. Genome Biology 2005). Les domaines N-SET, SET et C-SET dans SET7/9 sont indiqués. Le pseudo noeud et le peptide histone H3 lié sont également illustrés. Le produit AdoHcy (S-Adenosyl-I-homocysteine) issu de la déméthylation du SAM est représenté sous forme de batonnets bleu foncé. Les ions zinc sont représentés sous forme de billes rouges. N : terminaison amine; C, extrémité carboxyle.

Les KMTs à domaine SET sont réparties en 7 principales familles : SUV3-9 (suppressor of variegation 3-9), SETD1, SETD2, EZ (enhancer of zeste), RIZ (Retinoblastoma-interacting zinc-finger), SMYD (SET and MYND (Myeloid-Nervy-DEAF1) domain containing) et SUV4-20. D'autres enzymes inclassables, dites orphelines, telles que SET7/9 et SET8 aussi appelée PR-SET7, sont également décrites. Les membres d'une même famille partagent un haut niveau d'homologie au niveau du domaine SET, et présentent des motifs similaires de part et d'autre du site catalytique ⁹⁸. Les KMTs appartenant aux familles SUV4-20 et RIZ semblent capables de ne méthyler que les histones, alors que les méthyltransférases appartenant aux familles SUV3-9 (G9a, GLP et SETDB1), SETD1, SETD2, SMYD (SMYD2 et SMYD3) et EZ (EZH2) sont capables de méthyler des histones et des protéines non-

histones⁶³. La méthylation des lysines via le domaine SET se fait préférentiellement au niveau de la queue des histones, zone particulièrement riche en lysine et plus accessible de par sa flexibilité.

Un nombre croissant de substrats non-histones a été identifié ces dernières années^{78,90,101,102}. Carlson & Gozani ont classé les enzymes du domaine SET en trois groupes : les enzymes méthylant les histones, les enzymes modifiant les protéines non-histones et les enzymes n'ayant pas encore d'activité enzymatique claire, appelées «orphelines» (Figure 9)¹⁰¹. Toutefois, la méthylation des lysines n'est pas limitée aux méthyltransférases à domaine SET et de plus en plus de MTs à domaine 7β -strand sont décrites comme des KMTs.

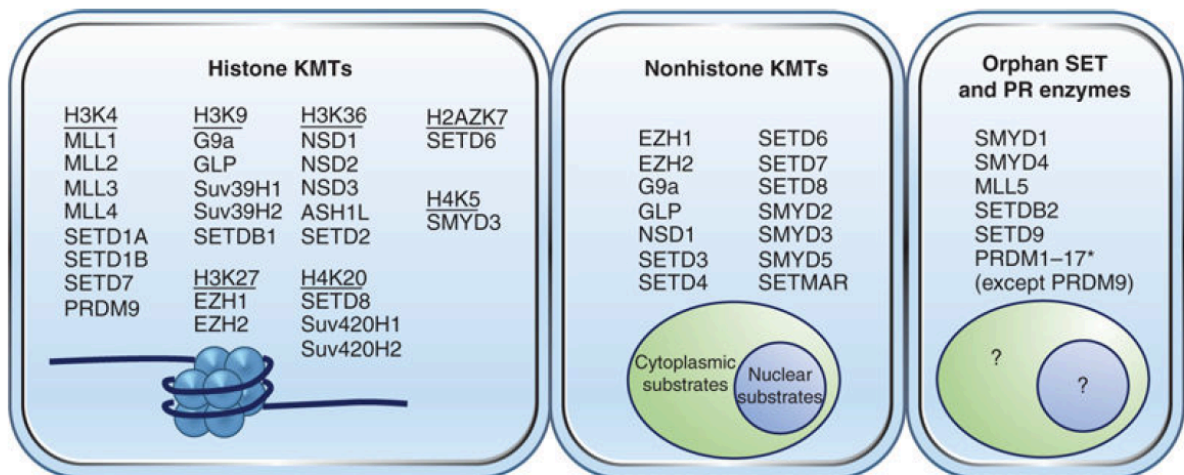


Figure 9 : Classification des KMTs à domaine SET. Les KMTs SET sont classées en 3 catégories selon qu'elles méthylent des histones, des protéines non-histone ou qu'il s'agit d'enzymes orphelines sans activité enzymatique bien établie. De nombreuses enzymes décrites comme histone-KMTs ont néanmoins des substrats non-histone décrit à l'heure actuelle. En outre, il existe des protéines qui sont principalement des KMTs non-histone, avec des rapports d'activité de méthylation des histones.

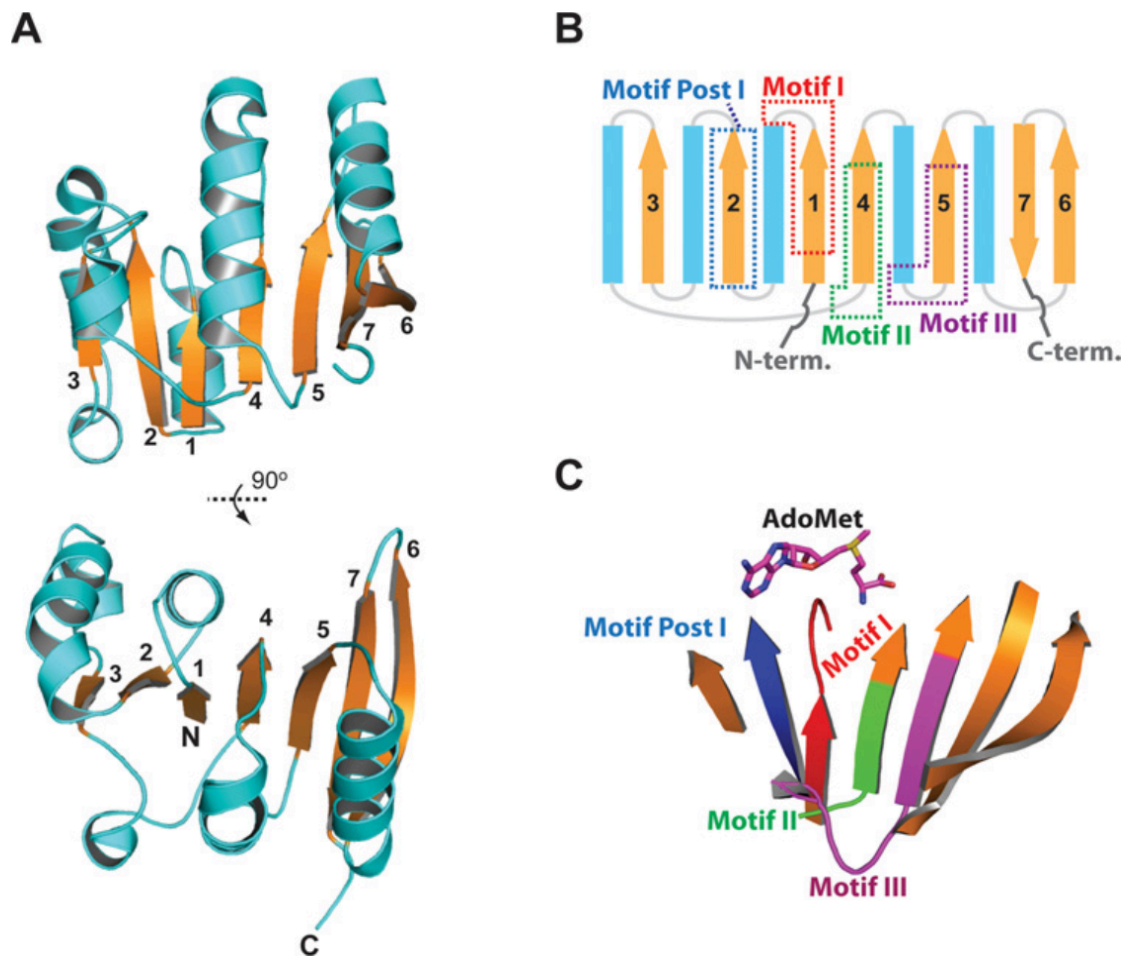


Figure 10 : Structure des KMTs à domaine 7β-strand. (Falnes *et al.* Biochem. J. 2016)

(A) Structure d'une KMT 7BS typique illustrée par l'exemple de METTL21A. Les hélices α et les boucles sont représentées en bleu cyan, les brins du feuillet β en orange. Les brins β sont numérotés dans le sens N-ter > C-ter. (B) Représentation schématique de la topologie des KMTs 7BS. (C) Emplacement des motifs caractéristiques permettant le positionnement du cofacteur SAM au niveau du site catalytique des KMTs 7BS.

- Les KMTs à domaine 7β-strand

La structure du site catalytique des KMTs à domaine 7BS est formée par l'alternance de brin beta ($\beta_1 - \beta_7$) et d'hélices alpha. Cette structure forme un feuillet β à sept brins avec trois hélices de chaque côté. L'ordre des brins est classiquement 3-2-1-4-5-7-6, avec le 7^{ème} brin antiparallèle aux autres. Les hélices Z, A et B sont situées d'un côté du feuillet β , et les hélices C, D et E de l'autre (Figure 10A)¹⁰³. Le feuillet beta contient 4 motifs catalytiques conservés, notés «I», «Post-I», «II» et «III» (Figure 10B), qui favorisent l'adhésion aux

substrats et au cofacteur SAM, notamment en ce qui concerne les motifs I et Post-I (Figure 10C). Ces motifs sont organisés entre eux et distants les uns des autres de manière variable d'une KMT 7BS à l'autre. Ces particularités structurales constituent par ailleurs, des paramètres essentiels dans la reconnaissance et la classification des KMTs 7BS ¹⁰³. Ces méthyltransférases appartiennent à une super famille présente dans de nombreuses espèces eucaryotes et procaryotes, et chez les archées. Les membres de cette famille sont capables de méthyler l'ADN, l'ARN, plusieurs acides aminés tels que l'arginine, la glutamine, l'aspartate et l'histidine. La première KMT 7BS décrite est la COMT (catechol-O-methyltransferase), impliquée dans la dégradation des neurotransmetteurs et qui constitue une cible pharmacologique d'intérêt. Depuis, de nombreuses autres KMT 7BS ont été identifiées ^{104,105}.

b) Les Lysines déméthylases (KDMs)

La méthylation des lysines fait l'objet d'une régulation fine et dynamique et son caractère réversible est mieux démontré que pour les arginines. On distingue plus d'une trentaine de KDMs avec deux grandes familles de lysine-déméthylases ¹⁰⁶.

La première famille regroupe les mono-oxydases ayant une activité dépendante de la flavine avec LSD1 (lysine-specific histone demethylase 1A) ou KDM1A en chef de file, première déméthylase décrite. Les KDMs de cette famille peuvent uniquement déméthyliser les lysines mono- ou di-méthylées. Elles ne sont pas capables de se lier aux lysines tri-méthylées. Ainsi, LSD1 déméthyle H3K4me1/2, et H3K9me1/2 selon le contexte. Cette dernière est donc clairement impliquée dans la régulation épigénétique via les modifications d'histones, et participe à la production de lncRNAs. Sa dérégulation est impliquée dans plusieurs mécanismes de carcinogenèse et elle est surexprimée dans de nombreux cancers ¹⁰⁷.

La seconde famille rassemble les histones lysines Jmj déméthylases précédemment décrites du fait de leur supposée activité arginine déméthylases. Elles sont capables de déméthyliser des résidus lysines à l'état mono-, di-, ou tri-méthylé (Figure 7). KDM5, par exemple, déméthyle H3K4me2/3 avec un effet répresseur sur la transcription de gènes suppresseurs

de tumeur, stimulant ainsi la croissance tumorale dans de nombreux cancers dont les cancers du sein, de la prostate, de l'estomac, du foie et du poumon ¹⁰⁸.

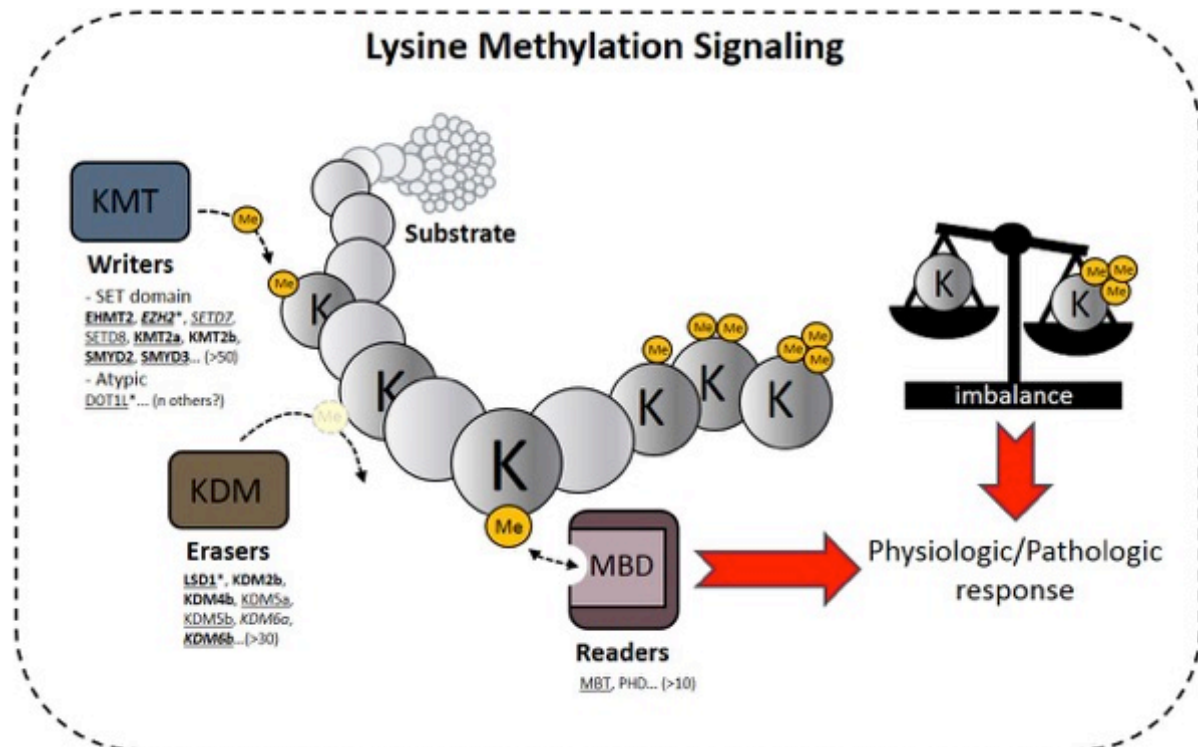


Figure 11 : Signalisation de la méthylation de la lysine et de ses acteurs dans le cancer du pancréas. (Roth et al. Current opinion in Oncology 2017) Représentation de la méthylation de la lysine et de ses acteurs, comprenant les lysine-méthyltransférases (KMT), les lysine-déméthylases (KDM) et les interacteurs liant le groupe méthyle (MBD). *En italique*, les enzymes intervenant dans la physiologie du pancréas ; **en gras**, les enzymes impliquées dans le cancer du pancréas ; souligné, les enzymes avec inhibiteurs existants ; avec *, les enzymes ayant un inhibiteur évalué en essai clinique.

La méthylation des lysines est impliquée dans de nombreux mécanismes physiologiques et pathologiques avec un nombre grandissant de substrats caractérisés. Nous allons aborder dans les parties suivantes les principales signalisations cellulaires impactées par cette PTM, en situation physiologique ou pathologique hors cancer, puis plus spécifiquement dans le cadre de la carcinogenèse. Les signalisations dépendantes des protéines famille SMYD seront détaillées dans la section "Les lysine méthyltransférases SET and MYND dans le cancer".

2. Les principales signalisations régulées par la méthylation des lysines

La méthylation des lysines est complexe et s'inscrit dans le concept épigénétique de *writers*, *erasers and readers*, au même titre que de nombreuses PTMs. La présence ou l'absence d'une marque de méthylation, de même que le niveau de méthylation du site concerné vont réguler les interactions protéine-protéine et impacter ainsi de nombreuses voies de signalisation cellulaires (Figure 11) ¹⁰⁹. Les histones sont les premières protéines dont la méthylation a été décrite, dès les années 1960 ^{8,110}. Il est aujourd'hui évident que les KMTs, notamment celles à domaine SET, jouent un rôle essentiel dans la régulation épigénétique.

KMTs (writers)	Pathologies dans lesquelles sont impliquées ces KMTs	KDMs (erasers)	Pathologies dans lesquelles sont impliquées ces KDMs
--------------------------	---	--------------------------	---

a) Signalisations impliquant la méthylation des histones

Une des premières signalisations qui a été démontrée est la méthylation de H3K9 par SUV39H1, découverte pionnière dans le domaine. Cette dernière interfère avec la phosphorylation de la sérine 10 de H3, entraînant des divisions mitotiques aberrantes ^{14,111}. De nombreuses autres KMTs ont ensuite été décrites comme catalyseur de modifications au niveau des histones, avec un effet répresser ou activateur sur l'expression des gènes selon les cibles ¹¹². Toutes les découvertes ultérieures et les événements identifiés ont mis en avant le concept du « code histone », selon lequel les empreintes post-traductionnelles des histones régulent la dynamique de la chromatine et l'expression génique, avec de multiples conséquences biologiques physiologiques ou pathologiques. De nombreuses recherches ont été menées dans ce sens, identifiant différents domaines de lecture, tels que les chromodomaines, les domaines Tudor et les domaines PHD ¹¹¹. Ainsi, à l'heure actuelle, les sites les plus étudiés sont H3K4, H3K9, H3K27, H3K36, H3K79 et H4K20, mais de nombreuses autres lysines méthylées ont

SET1A	Carcinome hépatocellulaire	LSD1	Adénocarcinome prostatique
SET1B		LSD2	Cancer du poumon
MLL1		RIOX1	Cancer du poumon
MLL2		JARID1A	Spondylarthrite ankylosante
MLL3	Carcinome urothélial	JARID1B	Cancer mammaire
MLL4	Lymphome	JARID1C	Retard mental
SMYD1	Rhabdomyosarcome	JARID1D	Adénocarcinome prostatique
SMYD2	Adénocarcinome pancréatique Adénocarcinome gastrique		
SETD7/9	Adénocarcinome gastrique		
PRDM9	Leucémie de l'enfant		

Tableau 3 : Résumé des différentes KMTs et KDMs modifiant H3K4 (Tableau adapté de Hyun *et al.* Experimental & Molecular Medicine 2017)

Abréviations : KMTs : lysine-méthyl transférases, KDMs : lysine déméthylases, SET1A,1B : SET Domain Containing 1A, 1B ; MLL 1, 2, 3, 4 : Mixed-Lineage Leukemia Protein 1, 2, 3, 4 ; SMYD1, 2 : SET and MYND domain containing protein 1, 2 ; SETD7/9 : SET domain containing protein 7/9 ; PRDM9 : PR/SET Domain containing protein 9 ; LSD 1, 2 : Lysine Demethylase 1, 2 ; RIOX1 : Ribosomal Oxygenase 1, JARID 1A, 1B, 1C, 1D : Jumonji/ARID Domain-Containing Protein 1A, 1B, 1C, 1D également été retrouvées sur H3 et H4, ainsi que H1, H2A, H2B ¹¹³. Les marques telles que H3K4, H3K36 et H3K79 sont préférentiellement associées à l'euchromatine et donc à l'activation des gènes, alors que les méthylations H3K9, H3K27 et H4K20 sont associées à l'hétérochromatine et donc à la répression génique ¹¹⁴. Les PTMs des histones surviennent surtout sur leur queue ou partie N-terminale. De nombreuses autres marques de méthylation ont ensuite été décrites, mais avec un niveau de démonstration insuffisant. Par exemple, H3K4 a été décrite comme méthylée par une dizaine de KMTs à ce jour, dont plusieurs membres de la famille MLL et de la famille SMYD (Tableau 2). La plupart de ces KMTs sont décrites comme impliquées dans des pathologies notamment cancéreuses, dont le rôle oncogénique pourrait passer en partie par la méthylation de H3K4. Néanmoins, la démonstration des conséquences phénotypiques précises de cette marque reste encore très lacunaire voire inexistante pour la majorité de ces KMTs. Il n'est ainsi pas exclu qu'il s'agisse

d'artefacts *in vitro*, non reproductibles en cellules ou en modèle *in vivo*. Par exemple, jusqu'à récemment, la protéine SMYD3 était aussi considérée comme méthylant H3K4, mais un travail de l'équipe du Dr Reynoird a montré que cette méthylation, de même que celle de H4K20 par SMYD3, ne sont pas reproductibles, et que SMYD3 méthyle en réalité l'histone H4K5 *in vitro* et de manière endogène ¹¹⁵ (ce point est détaillé dans le chapitre « SMYD3 et cancer »).

Par ailleurs, la déméthylation de H3K4 illustre parfaitement la dynamique et la mise en jeu d'acteurs multiples pour une même marque de méthylation. En effet, H3K4 peut être déméthylée par LSD1 aux états mono- et di-méthylé, et par JARID1B/KDM5B, une déméthylase à domaine Jmj, à l'état tri-méthylé ^{107,116}. Parmi ces signalisations intriquées, les paramètres qui privilégient l'une par rapport à l'autre sont complexes et semblent tributaires de facteurs spatiaux (localisation intracellulaire, localisation des cellules et microenvironnement péricellulaire) et temporels (cycle cellulaire, âge des cellules). Par exemple, il a été montré que des marques contradictoires telles que H3K4me3 (activatrice) et H3K27me3 (répressive) peuvent coexister. Cet état chromatinien associant des marques répressives et activatrices simultanées est appelé « *poised chromatin* » ou chromatine bivalente. Elle semble essentielle dans le fonctionnement des cellules germinales, notamment dans 3 domaines : la prévention de la méthylation de l'ADN, la conservation de l'identité cellulaire et la préparation de leur pluripotence ¹¹⁷. Ainsi, dans l'exemple de la coexistence de H3K4me3 et de H3K27me3 dans les cellules germinales, la présence de H3K4me3 est clairement associée à l'hypométhylation de l'ADN de certains promoteurs, donc à l'activation des gènes concernés. Néanmoins, les gènes régulateurs du développement somatique doivent rester silencieux tout au long de la vie au niveau des cellules germinales. Ainsi, ceci peut expliquer pourquoi une marque répressive telle que H3K27me3 coexiste au niveau des promoteurs des gènes de développement. H3K27me3 pourrait aider à empêcher leur activation, alors que H3K4me3 active la transcription d'autres gènes utiles au développement des cellules germinales. A l'inverse, au niveau des îlots CpG,

on retrouve la marque H3K4me3 seule ¹¹⁷.

De nombreuses marques de méthylation ont été décrites sur les histones. La recherche récente se concentre sur la caractérisation fine de marques déjà décrites afin de mieux comprendre les différents mécanismes de régulation de ces évènements. Par ailleurs, il a été démontré à de nombreuses reprises que certaines KMTs, initialement considérées comme des histones méthyltransférases, ont une activité cytoplasmique. Ceci a permis de mettre en évidence que ces enzymes ont des substrats non-histones, ouvrant un nouveau domaine de recherche très actif dans la compréhension de la signalisation cellulaire de la méthylation des lysines.

b) Signalisations indépendantes de la méthylation des histones

De nombreuses méthyltransférases ont été initialement décrites comme des histone-méthyltransférases, du fait d'une part de l'intérêt marqué de la communauté scientifique pour les histones avec une recherche plus active dans ce domaine, et d'autre part de l'abondance et de la richesse en lysine des histones, rendant la caractérisation des évènements de méthylation plus évidents. Ceci illustre les biais induits par les méthodes de détection de substrats méthylés utilisées à l'heure actuelle. La méthylation des protéines non-histones est maintenant très bien établie avec de nombreux exemples de signalisations essentielles régulées par cette PTM.

La méthylation des protéines non-histones regroupe un ensemble de processus biologiques qui ont un impact majeur sur la physiologie des cellules ainsi que dans la genèse de nombreuses pathologies telles que les cancers. L'ensemble de ces méthylations constitue le méthylome. Leur identification a notamment été permise par les avancées techniques de la protéomique, avec les progrès de la spectrométrie de masse et le développement des technologies de prédiction des sites de méthylation. Par exemple, Biggar et al. (2015) ont développé une approche prédictive pour l'identification des sites de méthylation en utilisant les sites disponibles identifiés dans la littérature et les bases de données accessibles,

permettant l'utilisation d'approches ciblées par spectrométrie de masse et une meilleure compréhension du méthylome ⁹⁰.

Tout d'abord, un certain nombre d'enzymes initialement décrites comme des histones-MTs sont en réalité capables de méthylater des substrats non-histones. Ces événements de méthylation concernent néanmoins souvent des fonctions de régulation de la chromatine et de l'expression génique. Par exemple, EZH2 (Enhancer of zeste homolog 2) peut modifier l'expression des gènes par modification de l'histone H3K27, en tant que marque répressive, ou par méthylation des facteurs de transcription ROR α (RAR-related orphan receptor alpha) et STAT3 (Signal transducer and activator of transcription 3) ¹¹⁸. Par ailleurs, une même KMT peut avoir un effet opposé sur la régulation des gènes, en fonction du contexte et des substrats affectés, comme l'illustre l'exemple de SETD1A. Cette dernière joue un rôle important dans le renouvellement et la différenciation des cellules souches embryonnaires par la régulation des gènes homéotiques, en méthylant l'histone H3 ¹¹⁹, mais elle peut également avoir un rôle oncogénique via la méthylation de la protéine YAP et sa séquestration dans le noyau ¹²⁰. Il est important de noter que lesdites histone-méthyltransférases peuvent affecter l'activité de substrat histones et non-histones en utilisant des mécanismes similaires. Par exemple, la diméthylation de H3K9 par la KMT G9a/GLP, aussi connue sous le nom de EHMT2/EHMT1 (Euchromatic Histone Lysine Methyltransferase 2/1), semble indispensable dans le développement de lignées cellulaires à partir des souches embryonnaires en réprimant l'expression des gènes impliqués dans la pluripotence ^{121,122}. G9a est aussi responsable de la triméthylation de la protéine ATF7IP (Activating transcription factor 7-interacting protein 1), associée au facteur de transcription ATF7 qui régule de nombreux gènes et réprime notamment l'expression des gènes spécifiques de l'immunité innée ¹²³. Les marques H3K9me2 et ATF7IPK16me3 semblent toutes deux favoriser le recrutement de la protéine à chromodomaine MPP8 (M-Phase Phosphoprotein 8), qui est elle-même impliquée dans la régulation de l'expression des gènes notamment via son interaction avec la méthyltransférase SETDB1 ¹²⁴ (Figure 12).

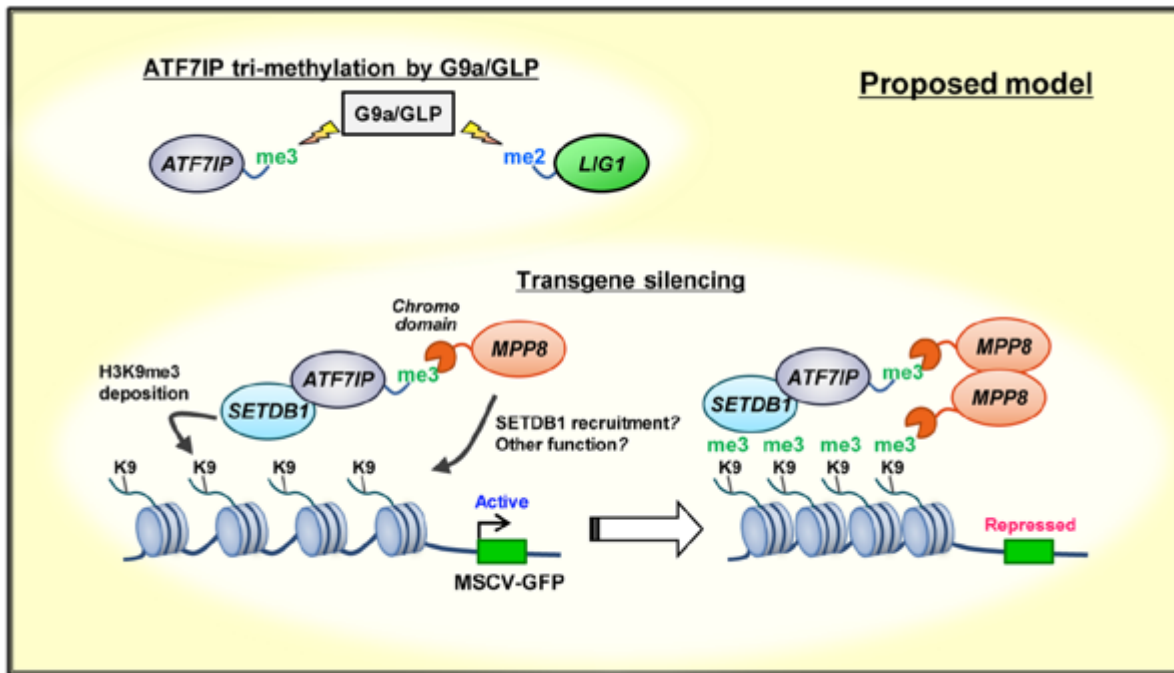


Figure 12 : Modèle proposé des conséquences moléculaires de la triméthylation de ATF7IP par G9a (Tsusaka et al. Epigenetics & Chromatin 2018).

Cependant l'activité enzymatique des KMT peut également être liée à différents processus. Ainsi, G9a/GLP peut avoir un rôle activateur de l'expression de gènes en méthylant des résidus de structures très proches des histones comme par exemple l'ADN ligase. Cet événement de méthylation non-histone favorise l'interaction de l'ADN ligase 1 avec UHRF1, qui favorise la réplication de l'ADN et intervient dans la régulation de la méthylation de l'ADN^{76,125}. Enfin, alors que la méthylation de H4K20 par SET8 (également appelée KMT5A) promeut la réplication de l'ADN et la réparation des dommages, la méthylation de p53 par SET8 entraîne une répression des gènes dépendant de l'activité de p53¹²⁶.

Des signalisations sont parfois décrites en l'absence d'activité biologique connue. Une longue liste de méthyltransférases à domaine SET manquent de démonstration claire d'une activité de méthylation. Elles sont qualifiées d'«orphelines»¹⁰¹ et il est possible que ces enzymes soient en réalité incapables de méthyler les histones, même de manière non-spécifique. Une autre possibilité est que de telles enzymes nécessitent des contextes spécifiques pour être actives (complexe protéique, modifications par des PTMs), ce qui rend

l'identification de leurs substrats extrêmement difficile. Certaines méthylations ont également été infirmées, démontrant l'absence de valeur biologique de certains événements retrouvés par les méthodes de détection à large échelle, comme l'illustre l'exemple des marques de méthylation de H3K4 ou H4K20 par SMYD3 qui ont été invalidées au profit de la méthylation de H4K5 ¹¹⁵. De plus, certaines études récentes indiquent une nouvelle fonction des enzymes contenant le domaine SET, suggérant que la lysine n'est pas le seul résidu méthylé de ce domaine. Il a en effet été récemment découvert que SETD3 agit comme une histidine-méthyltransférase, modifiant l'histidine 73 de l'actine. Cet événement de méthylation attribue un nouveau rôle aux KMTs à domaine SET dans la régulation de la stabilité de l'actine, qui semble liée à leur activité glycolytique ¹²⁷. Cet événement de méthylation est plus fréquent que la méthylation de H3K4 par SETD3, ce qui est probablement dû à la localisation principalement cytoplasmique de SETD3. Dans les souris n'exprimant plus SETD3 (Setd3 -/-), la méthylation de l'actine n'est plus possible et aucune autre marque de perte de méthylation n'est retrouvée en spectrométrie de masse. Cette signalisation joue probablement un rôle important dans la régulation de la contractilité des cellules musculaires et des muscles lisses car la mutation de l'actine H73 en modèle murin impacte fortement cette fonction ¹²⁸.

Ensuite, un nombre croissant de KMTs ne méthylant que des protéines non-histones sont décrites dans la littérature, confirmant le rôle primordial de la méthylation des lysines dans de nombreuses signalisations cellulaires. Une des premières protéines non-histones caractérisée comme substrat d'une KMT a été p53, protéine très largement étudiée du fait de son rôle essentiel dans la régulation des mécanismes de carcinogenèse ⁹⁵. p53 est massivement modifiée par diverses PTMs régulant son activité à la manière du code histone ⁹⁵. En effet, le domaine C-terminal (CTD) de p53 est riche en lysines et en arginines, et est donc sujet à de nombreuses PTMs telles que l'acétylation, la méthylation, la phosphorylation, l'ubiquitinylation, ce qui peut inciter à faire l'analogie avec la queue des histones. Au moins cinq KMT semblent méthyliser p53 et impacter ainsi son activité (ce point est détaillé dans la section "KMT et cancer").

Les paragraphes suivants offrent une synthèse des signalisations cellulaires physiologiques et parfois pathologiques, en cas de dérégulation, dans lesquelles la méthylation joue un rôle majeur.

- Régulation de l'activité transcriptionnelle

La méthylation des lysines participe à la régulation de l'expression de nombreux gènes en modifiant de multiples facteurs de transcription.

- GATA4 (GATA binding protein 4), est un facteur de transcription régulé par EZH2 dans le développement du muscle cardiaque chez la souris et chez l'homme. EZH2 a été initialement décrit comme appartenant au complexe PRC2 (*Polycomb Repressive Complex 2*), qui réprime l'expression des gènes par le biais de la méthylation de H3K27, mais il peut aussi inhiber l'activité de GATA4 en le méthylant sur sa lysine K299, ce qui bloque son interaction avec la protéine p300, qui ne peut plus l'activer en l'acétylant (Figure 13) ¹²⁹.

- STAT3, un autre substrat de EZH2 qui le méthyle sur sa lysine K180, appartient à la voie de signalisation JAK/STAT. Cette voie est impliquée de manière physiologique dans la transduction du signal et le maintien de l'homéostasie cellulaire. EZH2 est aussi capable de méthyle STAT3 sur sa lysine K49, lorsqu'elle est activée au préalable par l'interleukine IL-6, ce qui favorise l'expression d'un ensemble de gènes dépendants de la voie IL6-STAT3 ¹³⁰. Ces fonctions activatrices de la transcription de gènes contrastent avec le rôle répresseur de EZH2 au sein du complexe PRC2 précédemment décrit, ce qui constitue un des multiples exemples de la complexité de la signalisation de la méthylation des lysines et l'influence du contexte biologique. Enfin, STAT3 est également méthylé par la KMT SETD7, sur sa lysine K140, ce qui réprime son activité transcriptionnelle ¹³¹.

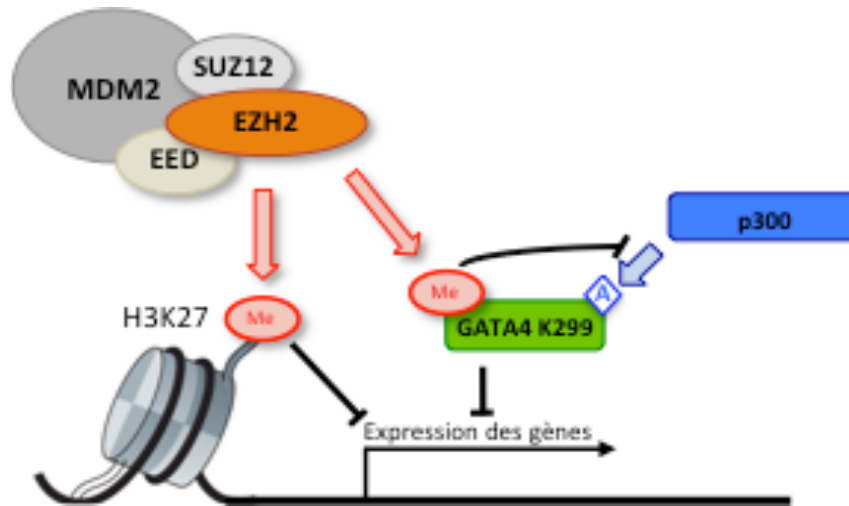


Figure 13 : Schéma de la régulation de l'expression génique par EZH2. EZH2 inhibe l'expression des gènes par le biais de la méthylation de l'histone H3 au niveau de sa lysine K27 et du facteur de transcription GATA4 au niveau de sa lysine K299, empêchant ainsi son acétylation par p300.

- NF- κ B (*nuclear factor-kappa B*), facteur de transcription qui joue un rôle important dans divers processus biologiques liés à l'immunité, l'inflammation et l'apoptose, est un substrat de SETD7. Il se compose de plusieurs sous-unités régulées par un panel de PTMs dépendantes de stimuli extracellulaires spécifiques, visant à adapter l'activité de NF- κ B et l'expression de ses gènes-cibles. L'activation constitutive du complexe NF- κ B est souvent retrouvée dans des situations pathologiques telles que l'inflammation ou le cancer, en participant à la néoangiogenèse et l'invasion tumorale. L'ubiquitination et la dégradation de la sous-unité NF- κ B RelA sont les processus habituels d'inhibition de la voie NF- κ B. Le processus de dégradation de RelA est notamment déclenché par la monométhylation des lysines K314 et K315 par SETD7¹³². La méthylation de RelA sur sa lysine K310 par SETD6 contrôle l'activité transcriptionnelle de ce dernier via le recrutement de la méthyltransférase G9A-like. En effet, cette dernière va se fixer à RelAK310me1, ce qui va favoriser la méthylation de H3K9 et réprimer ainsi de nombreux gènes dépendant de la voie NF- κ B¹³³. A l'inverse, la KMT NSD1 (*nuclear receptor binding SET domain protein 1*) active NF- κ B en méthylant sa sous-unité RelA au niveau de ses lysines K218 et K221^{134,135}.

- Les protéines chaperonnes

Les chaperonnes se lient aux protéines en cours de synthèse et contribuent à façonner leur structure 3D, les protégeant contre les différents systèmes de dégradation et contre les interactions protéine-protéine indésirables. Elles facilitent aussi le passage des protéines au travers des différentes membranes qui compartimentent la cellule, comme la membrane interne des mitochondries. Un groupe important de protéines chaperonnes sont les *Heat shock proteins* (HSP), avec HSP 90 et HSP 70, qui sont conservées au cours de l'évolution et régulées par la méthylation des lysines. HSP90 est monométhylée par SMYD2 (K616me1), une modification importante dans le développement musculaire par le biais de l'organisation du myofilament, ainsi que dans le cancer (ce point est détaillé dans la section « Implications de SMYD2 dans le cancer »). HSP70 est diméthylée sur sa lysine K561 par SETD1A, ce qui semble favoriser son passage du cytoplasme vers le noyau, lui permettant ainsi d'interagir avec la kinase Aurora B avec un impact sur le cycle cellulaire ¹³⁶. Cet événement de méthylation est néanmoins sujet à controverse car la KMT METTL21A, une protéine à domaine 7 β -strand, est également décrite comme méthylant la lysine K561 de HSP70 lysine. De

plus, la déplétion en METTL21 entraîne une perte complète de la méthylation de HSP70 ¹³⁷.

- La méthylation de l'ADN

DNMT1, impliquée dans la maintenance des marques de méthylation comme précédemment décrit, est monométhylée au niveau de sa lysine 142 par SETD7, ce qui augmente sa stabilité et provoque des conséquences différentes selon le partenaire qui se fixera sélectivement à cette marque de méthylation. En effet, la protéine PHF20L1 (PHD Finger Protein 20 Like 1), en se fixant à DNMT1 K142me1, inhibe sa polyubiquitinylation et donc sa dégradation, alors que L3MBTL3 (*Lethal(3)malignant brain tumor-like protein 3*) favorise son ubiquitinylation ¹³⁸. Par ailleurs, la méthylation de la *DNA ligase 1* (LIG1) par G9a permet le recrutement du partenaire spécifique de l'ADN méthylé UHRF1 au niveau des foci de réplication et participe à la méthylation de maintenance ¹²⁵. G9a méthyle également DNMT3a

et influence la méthylation *de novo* par ce biais, ainsi que par son auto-méthylation. En effet, DNMT3a méthylée et G9a autométhylée sont toutes deux reconnues par la protéine MPP8 pour induire la compaction de la chromatine. Enfin, G9a est responsable de la méthylation de H3K9, une marque répressive comme précédemment décrit. Elle réprime donc l'expression génique à la fois par la méthylation de protéines non-histones, impactant la méthylation de l'ADN, et via la méthylation de H3K9 ¹²⁴.

- La réplication et la réparation de l'ADN

SETD8 est une KMT qui joue un rôle critique dans la régulation de la réplication et la réparation de l'ADN. Elle a été initialement décrite comme méthylant H4K20, marque dont le rôle dans le maintien de l'intégrité du génome a été récemment démontré. En méthylant H4K20, SETD8 favorise la compaction de la chromatine et inhibe ainsi la réplication de l'ADN ¹³⁹. À l'inverse, elle méthyle la protéine PCNA (*proliferating cell nuclear antigen*), qui appartient à la famille des clamps glissants ou *sliding clamps*, dont le rôle est de potentialiser la processivité des ADN polymérase lors de la réplication chez les eucaryotes. Sa méthylation sur la lysine K248 par SETD8 augmente sa stabilité en inhibant son ubiquitinylation ¹⁴⁰. Aussi, la méthylation du PCNA favorise son interaction avec l'endonucléase FEN1 (*Flap structure-specific endonuclease 1*), qui est un autre acteur important de la réplication de l'ADN. La dérégulation de cette signalisation entraîne donc une altération de la réplication de l'ADN et induit des dommages à l'ADN, avec un impact possible dans la carcinogenèse. Ainsi, la surexpression de SETD8 est associée à de nombreux cancers ¹³⁹. Son rôle dans la carcinogenèse peut être également lié à sa capacité à méthyler les deux suppresseurs de tumeur p53 et Numb. Un autre acteur de la régulation de la réplication est G9a. Comme mentionné dans la section « La méthylation de l'ADN », la méthylation de LIG1 par G9a promeut son interaction avec UHRF1 et donc la réplication de l'ADN ¹²⁵. PARP1 intervient aussi dans la réplication et notamment dans la réparation de l'ADN, la transcription et le remodelage de la chromatine. Les fonctions de PARP1 sont largement régulées par les PTM, et notamment la méthylation par SMYD2 ¹⁴¹ (cf. section

« Implications de SMYD2 dans le cancer »). Enfin, le suppresseur de tumeur PTEN (*Phosphatase and TENsin homolog*), dont le rôle est de réguler négativement la voie de signalisation de PI3K(phosphatidylinositol 3-kinase)/AKT/mTOR, est méthylé par NSD2, ce qui induit une augmentation de la réparation efficace des dommages de l'ADN. En effet, en cas de cassure de l'ADN double brin, NSD2 diméthyle la lysine K349 de PTEN, qui est spécifiquement reconnue par la protéine 53BP1 (*tumor suppressor p53-binding protein 1*) et permet le recrutement de PTEN sur le site des dommages afin d'assurer une réparation efficace de l'ADN ¹⁴².

La figure 14 offre une vue générale et synthétique de la complexité et de l'intégration des différents processus où la méthylation des lysines joue un rôle essentiel. Un domaine où la recherche est particulièrement prolifique est la cancérologie. La caractérisation de ces différentes signalisations représente donc un enjeu majeur tant pour la compréhension de la physiopathologie que d'un point de vue pharmacologique avec l'identification potentielle de nouvelles cibles thérapeutiques pour améliorer les traitements contre le cancer.

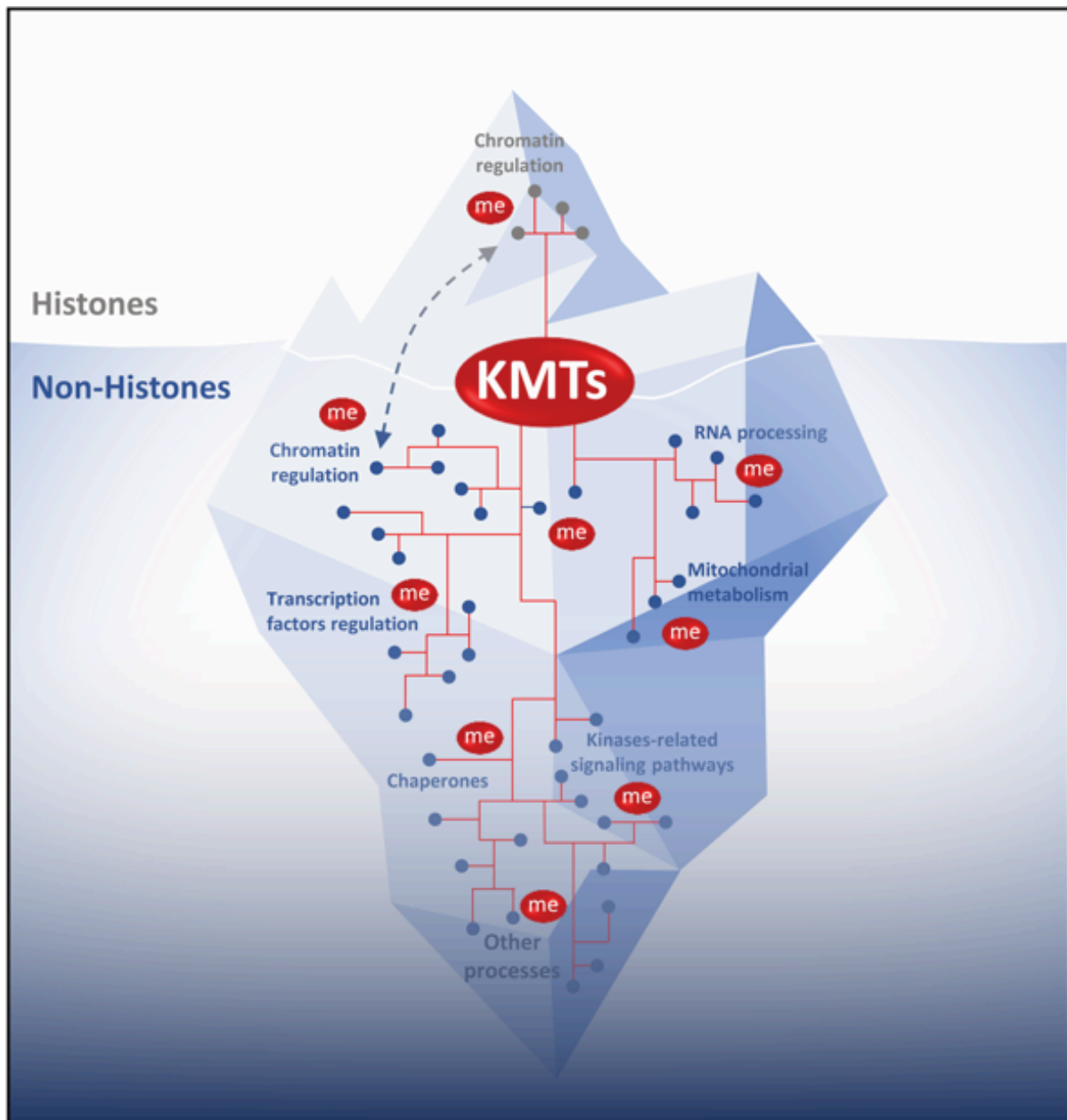


Figure 14 : Résumé des signalisations cellulaires impliquant la méthylation des histones et des protéines non-histone (Ludovic *et al.* Current Protein & Peptide Science 2020).

3. Méthylation des lysines et cancérogenèse

De nombreux évènements de méthylation sont impliqués dans la carcinogénèse. La plupart des voies de signalisation décrites en situation physiologique dans le chapitre précédent sont impliquées dans des processus de carcinogénèse en cas de dérégulation. Ainsi, un certain nombre de KMTs, telles que EZH2 ou SMYD2/SMYD3, ou de protéines substrats régulées

par méthylation, telles que p53 ou Rb, sont particulièrement associés à l'émergence ou à la progression de tumeurs. Cette section propose une synthèse des principales signalisations de méthylation des lysines impliquées dans la cancérogénèse en dehors des protéines SMYD2 et SMYD3 qui feront l'objet d'un chapitre dédié.

a) STAT3

Comme mentionné précédemment, le facteur de croissance STAT3 est un substrat de la KMT EZH2. Lors de l'activation de la voie PI3K/AKT/mTOR, AKT active EZH2 en le phosphorylant, ce qui augmente son activité de méthylation de STAT3 sur sa lysine K180. On observe alors une sur-activation de STAT3, ce qui participe au signal oncogénique en stimulant la prolifération cellulaire et l'expression de différents gènes¹⁴³. La phosphorylation activatrice de EZH2 semble être dépendante de la kinase CDK2 (*cyclin E/cyclin-dependent kinase 2*) comme montré dans un modèle préclinique de cancer du sein triple négatif dans lequel EZH2 inhibe l'expression du récepteur de l'œstrogène ER α ¹⁴⁴. Par ailleurs, EZH2 favorise la carcinogénèse par de multiples, autres mécanismes. En effet, elle méthyle de nombreux autres facteurs de croissance et joue un rôle important dans la régulation de la chromatine. Son potentiel de signalisation non-histone et son implication oncogénique restent cependant à déchiffrer. EZH2 est associée à la survie et la prolifération cellulaires, la transition épithélio-mésenchymateuse, l'invasion, la chimiorésistance et elle impacte le microenvironnement tumoral en régulant différentes cellules immunitaires telles que les lymphocytes T, les *natural killers*, les cellules dendritiques et les macrophages¹³⁰. EZH2 est surexprimée dans 70% des cancers solides et dans de nombreuses hémopathies malignes. Néanmoins, probablement du fait de son rôle physiologique dans l'homéostasie cellulaire, elle semble également avoir, dans certaines conditions, un rôle suppresseur de tumeur, puisque sa perte d'expression est associée à une prolifération tumorale augmentée dans des pathologies telles que les leucémies lymphoblastiques aiguës, les adénocarcinomes pancréatiques dépendant de la voie RAS, ou les carcinomes rénaux à cellules claires¹⁴⁵.

b) p53

p53 est l'un des premiers substrats non-histones observés. Sa régulation complexe par de multiples méthylations de lysine et par d'autres PTM rappelle celle des histones. La méthylation de p53 résulte de l'activité de plusieurs KMT : SMYD2, SETD8, G9a, SETDB1 et SETD7. Les lysines sont nombreuses au niveau du domaine C-ter de p53 et sont essentielles pour contrôler son activité, du fait de la compétition entre différentes marques d'acétylation, d'ubiquitylation et de méthylation. Les premières marques de méthylation identifiées sur p53 sont celles qui affectent la lysine K370 par SMYD2 et K382 par SET8. Ces marques restent les plus étudiées à ce jour. Elles induisent une répression transcriptionnelle de plusieurs gène-cible de p53, dont P21 / WAF1, responsable de l'arrêt du cycle cellulaire en G0/G1 ¹⁴⁶. Plusieurs mécanismes semblent concourir à cette répression. Par exemple, la méthylation K382me1, dépendante de SETD8, permet le recrutement de la protéine L3MBTL1 qui réprime les gènes cibles de p53 en se liant à la chromatine et en induisant sa compaction au niveau des régions de leurs promoteurs ¹⁴⁷. SETD8 méthyle également la protéine Numb qui forme un complexe avec p53, conduisant à la dissociation du complexe et à la dégradation de p53 ¹⁴⁸. De même, la répression de p53 suite à la méthylation K370me1 par SMYD2 diminue sa liaison aux promoteurs des P21 / WAF et MDM2 (*Mouse double minute 2 homolog*). La déplétion en SMYD2 restaure l'apoptose induite par p53 en réponse aux dommages à l'ADN ¹⁴⁹. À l'inverse, les marques de méthylation K370me2 et K82me2 catalysées par SETDB1 et SETD8 sont liées à l'activation de p53. Contrairement à la monométhylation K370me1, la diméthylation K370me2 permet le recrutement du co-activateur p53BP1, qui facilite l'expression des gènes cibles de p53. Cette marque est par ailleurs réversible par l'action de la déméthylase LSD1, qui peut donc être inhibitrice en permettant le passage de K370me2 à K370me1, ou activatrice en permettant le passage de la forme monométhylée à la forme non-méthylée p53K370 ¹⁵⁰. La marque K382me2 catalysée par SETD8 favorise également la liaison au coactivateur p53BP1 ¹⁵¹. Une troisième marque activatrice a été décrite : il s'agit de p53K372me1 par SETD7, qui entraîne une fixation accrue de p53 sur le promoteur de ses gènes-cibles, par le biais du

recrutement de l'acétyltransférase Tip60 qui active p53 en l'acétylant sur ses lysines ¹⁵². De plus, cette marque de méthylation inhibe l'interaction de SMYD2 avec p53, limitant son rôle répresseur. Enfin, G9a méthyle p53 sur sa lysine K373, mais son rôle, activateur ou répresseur, reste discuté ^{153,154}.

c) pRB (Retinoblastome)

La protéine pRB induit l'arrêt du cycle cellulaire en phase G1 et constitue un mécanisme central de la suppression de la prolifération. Cet arrêt du cycle cellulaire passe notamment par la séquestration du facteur de croissance E2F. pRB est finement régulée par différentes PTMs, en particulier l'acétylation, la phosphorylation et la méthylation. La méthylation de pRB par SMYD2 semble inhiber l'activité de pRB par différents mécanismes (détaillés dans la section « Implications de SMYD2 dans le cancer ») ¹⁵⁵. pRB est également méthylée par SETD7 sur la lysine K810. Cet événement favorise le recrutement de la protéine hétérochromatine HP1, qui est nécessaire pour l'arrêt du cycle cellulaire et la répression de l'activité transcriptionnelle dépendante de pRB ¹⁵⁶. Enfin, SETD7 régule aussi indirectement pRB, car l'enzyme peut méthyler la lysine K442 de la protéine MYPT1 (*Myosin phosphatase target subunit 1*), bloquant son interaction avec et aboutissant à une diminution de l'activité de cette dernière ¹⁵⁶. Cette marque est réversible par l'action de LSD1, qui est capable de déméthyler MYPT1K442me ¹⁵⁷.

d) HIF-1 α (Hypoxia-induced factor 1 α)

HIF-1 α participe à la régulation de la transcription et à l'adaptation des cellules en situation d'hypoxie. En hypoxie, SETD7 méthyle HIF-1 α lorsqu'il est localisé dans le noyau, ce qui favorise sa dégradation par augmentation de son ubiquitylation ¹⁵⁸. En situation pathologique, certaines mutations de HIF-1 α , telles que S28Y et R30Q, semblent inhiber la méthylation de la lysine K32 par SETD7, ce qui augmente la concentration intracellulaire de HIF-1 α et entraîne une augmentation de la croissance tumorale. Ceci suggère donc que la méthylation de HIF-1 α joue un rôle important dans la régulation de l'angiogenèse et de la

vascularisation. HIF-1 α est également méthylé par G9a / GLP avec pour conséquence une inhibition de son activité transcriptionnelle sans impact sur sa dégradation. Dans le glioblastome, l'hypoxie induit une diminution de l'expression de G9a, ce qui peut aboutir à une augmentation de l'activité de HIF-1 α ¹⁵⁹.

e) WNT / β -caténine

La voie de signalisation WNT / β -caténine joue un rôle essentiel pour la détermination de la polarité cellulaire au cours de l'embryogenèse et est impliquée dans de nombreux aspects de la cancérogenèse, en particulier la transition épithélio-mésenchymateuse (EMT, Epithelial to Mesenchyme Transition). Dans cette voie, SETD6 méthyle la kinase PAK4 (p21-Activated kinase 4) lorsqu'elle est activée au préalable par p21, ce qui renforce son interaction avec la β -caténine et favorise la progression tumorale. SETD6 est importante pour l'activité de PAK4 car sa déplétion aboutit à des phénotypes similaires à la déplétion en PAK4. L'absence de PAK4 est en effet létale, car essentielle dans le développement embryonnaire ¹⁶⁰. SMYD2 participe aussi à la régulation de la voie de signalisation WNT / β -caténine en méthylant la β -caténine sur sa lysine K133, ce qui régule sa translocation nucléaire ainsi que la transduction du signal et l'expression génique promues par la β -caténine ¹⁶¹.

4. La signalisation par méthylation des lysines : nouvelles cibles thérapeutiques

La plupart des fonctions de KMTs mentionnées ci-dessus peuvent être directement ou indirectement liées au cancer, et le caractère dynamique, réversible et spécifique de leur activité en fait des cibles thérapeutiques potentielles pour tenter de contrôler les différents mécanismes de cancérogenèse. Cette piste thérapeutique n'est en réalité pas récente puisqu'en 1980 l'hypothèse de cibler les méthyltransférases utilisant le cofacteur SAM était déjà évoquée ¹⁶². Ainsi de nouvelles thérapies ciblées axées sur ces PTMs, considérées comme thérapies épigénétiques (« épidrogues ») lorsqu'elles ciblent les histones et donc la

régulation de l'expression génique, sont en cours de développement ¹⁶³. Cette recherche reste néanmoins à un stade très précoce, sans preuve clinique suffisante de leur efficacité à ce jour. Par ailleurs, la sélection des types de tumeurs qui répondraient à de telles thérapies nécessite l'identification de biomarqueurs pronostiques et prédictifs de réponse. Ces marqueurs sont complexes à développer du fait de la spécificité des signalisations de méthylation en fonction des sites et du niveau de méthylation, ainsi que du contexte biologique à l'échelle cellulaire, microenvironnemental et tissulaire. Cette complexité est très bien illustrée par le rôle ambivalent d'EZH2, qui comme précédemment décrit, agit à la fois en tant qu'oncogène dans certains cancers, et semble avoir un rôle suppresseur de tumeur dans d'autres pathologies ¹⁴⁵. La marque H3K27me3 catalysée par EZH2 est promotrice de la prolifération tumorale et est retrouvée dans un certain nombre de lymphomes et de leucémies ainsi que dans l'adénocarcinome prostatique ¹⁶⁴. L'inhibition pharmacologique de EZH2 dans un modèle préclinique de cancer du sein triple négatif permet de restaurer l'expression du récepteur de l'œstrogène ER α et induit une diminution de la progression tumorale en association à l'hormonothérapie ¹⁴⁴. A l'heure actuelle, un certain nombre d'études de phases 1 et 2 évaluant des inhibiteurs de EZH2 sont en cours et un essai de phase 1 récemment publié montre des résultats prometteurs pour le tazemetostat (un inhibiteur spécifique de EZH2) dans le lymphome B non-hodgkinien ¹⁶⁵. Des inhibiteurs de la KMT DOT1L (*disruptor of telomeric silencing 1-like*), qui méthyle H3K79 et dont l'activité est augmentée dans les cancers de la prostate, des ovaires, du sein et dans les leucémies, sont également à l'étude, avec des résultats intéressants dans certaines leucémies ¹⁶⁶. Dans le cancer colorectal, un certain nombre d'épidrogues ciblant la méthylation sont également à l'étude, tels que des inhibiteurs de EZH2 ou DOTL1, mais aussi de SUV39H1, qui méthyle H3K9 avec un effet répresseur sur la transcription. Ainsi, la chaetocine, un métabolite fongique, semble réduire la migration des cellules de cancer colorectal en inhibant l'activité de SUV39H1 dans le cancer colorectal, tout comme dans le cancer du poumon ¹⁶⁷. Une autre KMT essentielle qui fait l'objet d'essais cliniques est SMYD2 (chapitre « III.B.Implications de SMYD2 dans le cancer »). Les KDMs sont aussi ciblées, telle que

KDM5 qui promeut la croissance tumorale via la déméthylation de H3K4me2/3 comme précédemment décrit et dont l'inhibition permet une diminution de la croissance tumorale ¹⁰⁸, et la déméthylase LSD1 ^{113,168,169}. Ces résultats concernent principalement les cancers les plus fréquents et donc les plus étudiés. Il est cependant probable qu'une partie importante de ces mécanismes participe à la progression tumorale de nombreux autres cancers même si non décrits à l'heure actuelle, du fait du rôle ubiquitaire de certains substrats méthylés, tels que les histones. Ainsi, de multiples thérapies, combinaisons de thérapies et nouvelles indications devraient voir le jour dans les années à venir.

III- LYSINE-METHYLTRANSFERASES SET AND MYND DANS LE CANCER

Deux enzymes appartenant à la famille des KMTs à domaines SET and MYND ont récemment été impliquées dans la régulation de voies de signalisations majeures de la carcinogenèse, telles que la prolifération tumorale dépendante de la voie RAS, l'angiogenèse, l'inhibition de p53 ou de PTEN : il s'agit de SMYD2 et SMYD3. Dans cette section, je propose une revue de la littérature des fonctions et caractéristiques de ces enzymes et notamment de leur implication dans le cancer.

A. LYSINE-MÉTHYLTRANSFÉRASES SET AND MYND : GÉNÉRALITÉS

La famille SMYD est composée de 5 membres nommés SMYD 1 à 5. Comme précisé précédemment, ces protéines possèdent un domaine catalytique SET. Le domaine SET est associé à un domaine pré-SET et à un domaine SET-I (SET insertion), comme dans la plupart des KMT à domaine SET. Cependant, la spécificité des protéines SMYDs réside dans la particularité de leur domaine SET d'être séparé en 2 par un domaine MYND de liaison au substrat. La première partie de SET, située à l'extrémité N-ter de la protéine, semble avoir pour rôle de favoriser la liaison au cofacteur SAM et le reste du domaine SET, situé après le domaine MYND, est la portion catalytique à proprement parler. Le domaine MYND possède un motif de type *zinc finger* qui permet la liaison aux substrats de l'enzyme, et a également été décrit comme capable de lier l'ADN^{99,170}. Enfin, à l'exception de SMYD5, les SMYD possèdent un domaine C-ter (CTD), principalement impliqué dans la reconnaissance et la liaison au substrat⁹⁹ (Figure 15).

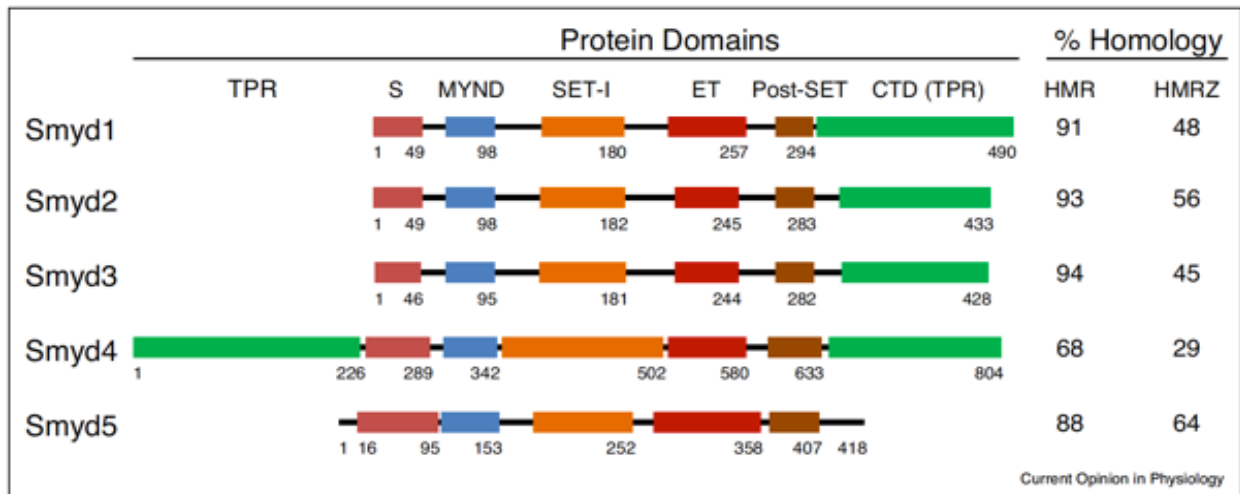


Figure 15 : Représentation linéaire des différents domaines structuraux des protéines de la famille SMYD (Tracy *et al.* Current opinion in Physiology 2018). Les domaines SET (S et ET, en rouge), MYND (bleu), SET-I (orange), Post-SET (marron), et CTD ou TPR (vert) sont représentés. Les nombres correspondent au dernier acide aminé de chaque domaine. Les taux d'homologie sont indiqués à droite, avec HMR correspondant à la comparaison des séquences humaine, de souris et de rat, et HMRZ correspondant aux séquences humaines, de souris, de rat et de poisson zèbre.

Les membres de la famille SMYD semblent avoir différentes capacités de méthylation. Par exemple, SMYD2 peut uniquement monométyler une lysine, alors que SMYD3 est capable de trimétyler une lysine (Tableau 3). Ces KMTs sont à la fois cytoplasmiques et nucléaires. Leur présence au niveau cytoplasmique pourrait être liée à leur liaison à HSP90. En effet, SMYD1, SMYD2 et SMYD3 possèdent un domaine C-terminal contenant 3 motifs TPR. Ces derniers sont chacun constitués de 34 acides aminés formant une hélice- α anti parallèle, et partagent des similarités avec les motifs TPR de nombreuses co-chaperonnes de HSP90, telles que les protéines HOP, CyP40, FKBP51/52 et p23⁹⁹. Ainsi, les protéines SMYD influencent l'activité de HSP90 en tant que co-chaperonnes, voire en la méthylation, comme cela a été montré pour SMYD2 (cf. « III.B. Implications de SMYD2 dans le cancer »). A l'inverse, leur propre activité semble également impactée par leur liaison à HSP90, comme le suggère l'ajout d'HSP90 *in vitro*, aboutissant à l'augmentation des marques de méthylation des histones catalysée par SMYD1, SMYD2 et SMYD3.

Les protéines de la famille SMYD ont initialement été décrites dans la régulation du remodelage de la chromatine et de l'activité transcriptionnelle via la méthylation des histones

	Niveau de méthylation	Localisation cellulaire	Substrats Histones	Substrats non-histone
SMYD1	Mono-, di-, tri-méthylation	Cytoplasme et noyau	H3K4	
SMYD2	Mono-méthylation	Cytoplasme et noyau	H3K4*, H3K36*, H4K20*	p53, HSP90, RB1, ER- α , PTEN, PARP1, EML4-ALK, β -caténine, BMPR2, MAPKAPK3
SMYD3	Mono-, di-, tri-méthylation	Cytoplasme et noyau	H3K4*, H4K5, H4K20*	VEGFR1, MAP3K2, AKT1, HER2
SMYD4	Absence de données	Absence de données	Absence de données	Absence de données
SMYD5	Tri-méthylation	Absence de données	H4K20*	Absence de données

Tableau 3 : Substrats histones et non-histone des KMTs de la famille SMYD
* : Niveau de validation insuffisant de l'événement de méthylation endogène

et via leurs interactions avec des facteurs de transcription. Progressivement, un nombre croissant de cibles non-histone a également été décrit et il est maintenant bien établi que des membres tels que SMYD2 et SMYD3 jouent un rôle important dans la cancérogenèse avec la régulation de nombreuses voies de signalisations oncogéniques. Néanmoins, ces protéines interviennent également dans des processus physiologiques notamment dans le développement du muscle cardiaque et des muscles squelettiques.

- SMYD1 régule la différenciation du muscle cardiaque permettant l'organisation et la contraction des fibres musculaires, notamment via la méthylation de l'histone H3 au niveau de sa lysine K4. Elle semble indispensable dans le développement cardiaque et sa délétion est à l'origine un défaut de maturation du muscle cardiaque chez la souris. Cette traduction phénotypique semble associée à une baisse d'expression de multiples facteurs de

transcription chez les souris *Smyd1^{-/-}*, dont *Hand2* essentiel à la formation du ventricule droit ¹⁷¹.

- SMYD2 est ubiquitaire, et son expression est notamment importante au niveau cardiaque, avec le niveau le plus élevé en période néonatale. Elle a été initialement décrite comme une histone méthyltransférase responsable de la méthylation des histones H3K36 et H3K4, deux marques qui se sont avérées être des artéfacts ^{172,173}. Son caractère indispensable dans le développement cardiaque est controversé avec des résultats discordants. On ne note en effet pas de conséquence phénotypique sévère chez la souris ¹⁷⁴, mais sa délétion chez le poisson-zèbre ou *zebrafish* aboutit à un œdème péricardique, des malformations cardiaques sévères et une diminution de la fraction d'éjection systolique ¹⁷⁵.

- SMYD3 est ubiquitaire et particulièrement abondante dans les testicules et au niveau du muscle squelettique. Elle a été initialement décrite comme une histone-méthyltransférase méthylant H3K4 et H4K20 ^{176,177}, mais elle méthyle en réalité H4K5, sans traduction claire *in vivo* ¹¹⁵. Elle semble impliquée dans le développement du myocarde et des muscles squelettiques, car sa répression (*Smyd3^{-/-}*) dans les embryons *zebrafish* induit un œdème précardique et des anomalies morphologiques cardiaques majeures ¹⁷⁸. Au cours du développement embryonnaire précoce, il a également été démontré que les niveaux d'ARNm de *Smyd3* augmentent après le stade à deux cellules, atteignent un pic au stade à quatre cellules puis diminuent progressivement par la suite. Au stade embryonnaire précoce des souris *Smyd3^{-/-}*, on observe des perturbations du développement de la masse cellulaire interne et du trophoctoderme avec une réduction de la viabilité des souriceaux. Aussi, l'expression des facteurs de transcription Oct4 (octamer-binding transcription factor 4), impliqué dans la pluripotence des cellules souches, et *Cdx2* (caudal type homeobox 2) impliqué dans la formation du placenta, est significativement diminuée dans les embryons *Smyd3^{-/-}*. Un effet similaire sur l'expression d'autres facteurs de transcription tels que *Nanog*

(Nanog Homeobox), Sox2 (sex determining region Y-box 2) et Gata6 (GATA Binding Protein 6) a également été observé confirmant que SMYD3 joue un rôle important dans la détermination précoce des lignées cellulaires et le développement péri-implantatoire¹⁷⁹. Par ailleurs, SMYD3 semble être importante dans le renouvellement des cellules de la muqueuse intestinale car sa répression est associée à une diminution de la croissance cellulaire au niveau des cryptes de Lieberkühn avec une baisse du niveau d'expression des transcrits CcnA3 (Cyclin A3), CcnD1 et CcnE1¹⁸⁰. A l'heure actuelle, malgré des résultats parfois discordants, il semble néanmoins qu'en cas de naissance, les souris *Smyd3*^{-/-} soient viables et fertiles sans conséquence phénotypique majeure^{181,182}. La caractérisation du rôle physiologique de SMYD3 dans le développement embryonnaire reste donc incomplète. Les implications de SMYD2 et SMYD3 dans la cancérogenèse sont nombreuses et seront détaillées dans les chapitres III.B. et III.C.

- SMYD4 et SMYD5 semblent avoir un rôle suppresseur de tumeur. SMYD4 inhibe l'expression de PDGFRa (*Platelet-derived growth factor receptor A*), et sa déplétion aboutit à une augmentation de la prolifération des cellules tumorales, comme cela a été décrit dans le cancer du sein¹⁸³. SMYD5 triméthyle l'histone H4 au niveau de sa lysine K20, ce qui favorise le maintien de la chromatine à l'état d'hétérochromatine et participe à la conservation du génome au cours des différentes étapes de différenciation et de renouvellement cellulaire. À l'inverse, sa déplétion *in vitro* dans les cellules embryonnaires aboutit à l'apparition de cellules transformées et favorise la survenue d'aberrations chromosomiques. De même, lorsque cette déplétion est retrouvée dans les cellules tumorales, elle est associée à une augmentation de la croissance tumorale via l'expression de gènes impliqués dans la prolifération tumorale¹⁸⁴.

B. IMPLICATIONS DE SMYD2 DANS LE CANCER

1. Expression dans les cancers

Chez l'homme, la surexpression de SMYD2 est retrouvée dans de nombreux cancers (Figure 16). Elle a été identifiée comme un facteur de mauvais pronostic dans les carcinomes malpighiens œsophagiens ¹⁸⁵, dans les cancers ORL secondaires au papillomavirus ¹⁸⁶, les cancers du sein ¹⁸⁷, les cancers de l'estomac ¹⁸⁸, les cancers du pancréas, les carcinomes hépatocellulaires ¹⁸⁹, les leucémies aiguës lymphoblastiques pédiatriques ¹⁹⁰ et les cancers papillaires de la thyroïde ¹⁹¹. La surexpression de SMYD2 est également associée à une augmentation de la prolifération tumorale dans un certain nombre de modèles précliniques, tels que les cancers du pancréas induit par l'inflammation chez la souris ¹⁹². Dans le cancer du pancréas, SMYD2 participe à l'initiation et la progression du cancer du pancréas comme montré dans un modèle murin d'adénocarcinome avec des souris présentant une mutation activatrice de l'oncogène Ras ($Ptf1a^{+/Cre};Kras^{+/LSL-G12D}$) et des lésions de pancréatite induites par injections répétées de caeruleïne. Ce modèle permet l'apparition de lésions de métaplasie acino-canalaire, stade le plus précoce de la séquence métaplasie-dysplasie-adénocarcinome, puis de lésions pré cancéreuses de type PanIN (néoplasies intra-épithéliales non invasives). La délétion de *Smyd2* ($Smyd2^{-/-}$) est associée à une diminution de ces lésions, suggérant que SMYD2 intervient dans les stades initiaux de la tumorigénèse pancréatique. De même, en de cas modèle plus agressif avec mutation *Kras* et *p53* combinées ($Ptf1a^{+/Cre};Kras^{+/LSL-G12D};Trp53^{loxP/loxP}$) qui permet l'apparition de lésions d'adénocarcinome, les souris $Smyd2^{-/-}$ présentent une prolifération tumorale ralentie. Enfin, dans un modèle de xénogreffe de cancer du sein triple négatif, la délétion de SMYD2 freine également la prolifération tumorale ¹⁸⁷.

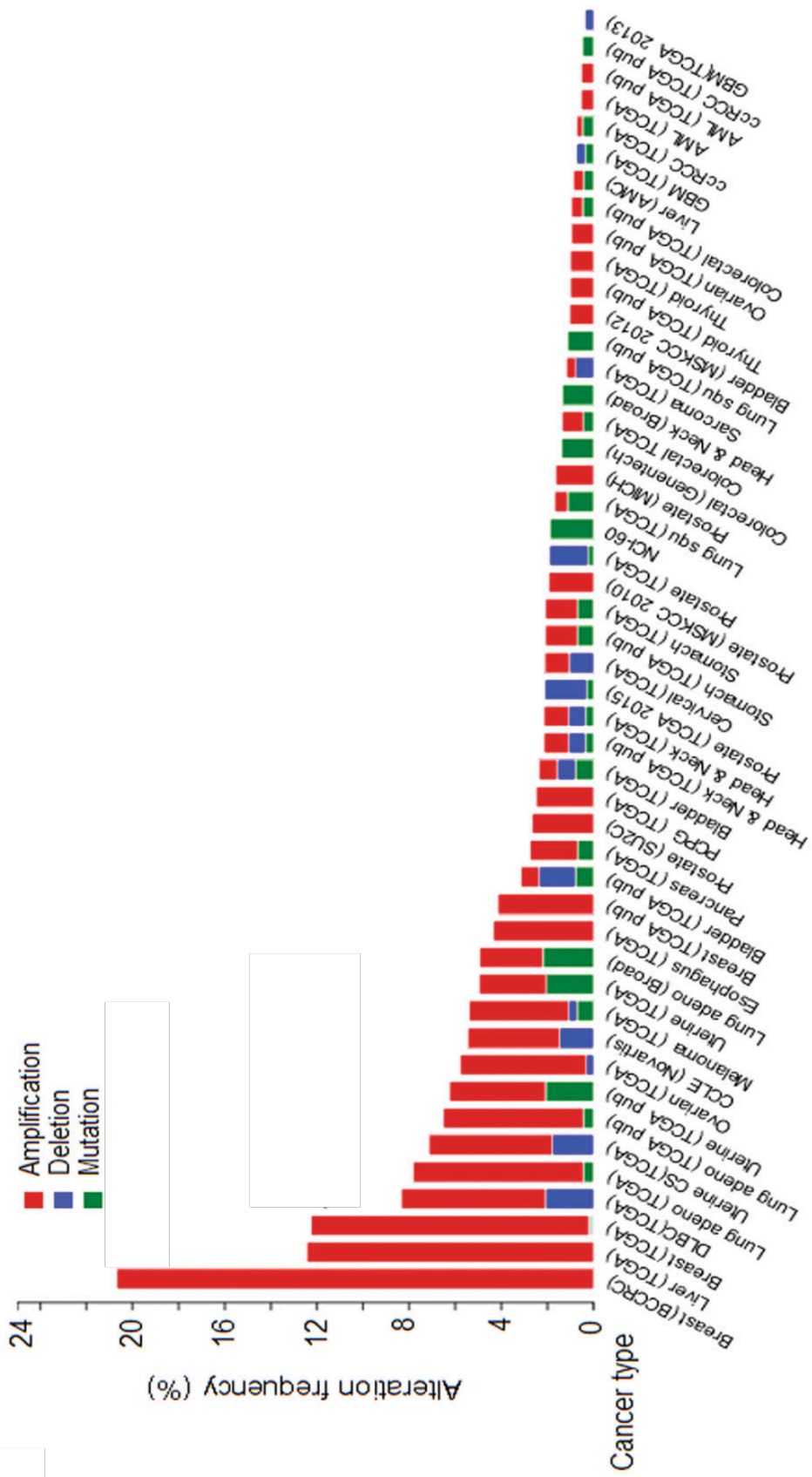


Figure 16 : Profil d'expression de SMYD2 dans les différents cancers (Figure adaptée de Li et al. Cell Death and Disease 2018). Données issues de 91 études via cBioPortal for Cancer Genomics.

2. Les différentes signalisations de SMYD2 dans la carcinogénèse

De nombreux acteurs dont la dérégulation a déjà été impliquée dans la carcinogénèse ont été récemment décrits comme étant méthylés par SMYD2. Les différents substrats de SMYD2 concernés sont détaillés ci-dessous et résumés par la Figure 17.

a) p53

SMYD2 est clairement identifiée comme méthylant le suppresseur de tumeur p53 sur sa lysine K370, avec une action répressive. En effet, la surexpression de SMYD2 est associée à une diminution de l'activité transcriptionnelle de p53 et notamment de sa liaison au promoteur du gène P21 / WAF1. À l'inverse, la répression de SMYD2 par shARN (*short hairpin* ARN) est associée à une augmentation de l'activité des promoteurs dépendants de p53. Sur le plan phénotypique, cette inhibition de p53 par SMYD2 induit une augmentation de la prolifération cellulaire et une diminution de l'apoptose en cas de dommages à l'ADN ¹⁴⁹. Par ailleurs, l'interaction entre SMYD2 et p53 peut être inhibée par l'action d'une autre KMT, SET9, lorsqu'elle méthyle p53 au niveau de sa lysine K372.

b) pRB

SMYD2 inhibe le suppresseur de tumeur pRB en le méthylant au niveau de 2 sites distincts, les lysines K810 et K860. (i) La méthylation de K810 promeut la phosphorylation de pRB sur sa sérine Ser 807/811, médiée par le complexe CDK4/Cyclin D1, interférant avec son interaction avec le facteur de transcription E2F1 qui est responsable de la transition entre les phases G1 et S du cycle cellulaire¹⁹³. (ii) La méthylation de K860 favorise sa liaison au répresseur transcriptionnel L3MBTL1. Ceci pourrait impacter la liaison de pRB à ses interacteurs et serait à l'origine de signaux de prolifération cellulaire, mais ce mécanisme est moins clairement décrit ¹⁵⁵.

c) HSP90

SMYD2 méthyle la protéine HSP90 dont l'activité chaperonne est impliquée dans de nombreux cancers. Cette dernière favorise le maintien de l'activité des oncoprotéines en les protégeant des différents mécanismes cellulaires de régulation et de dégradation. Ainsi, la méthylation de HSP90 par SMYD2 sur les lysines K531 et K574 favorise cette activité chaperonne et, promeut ainsi la prolifération tumorale. A l'inverse, la répression de SMYD2 par shARN induit une diminution de la concentration intracellulaire des protéines oncogéniques substrats de HSP90 telles que Her-2 (*human epidermal growth factor receptor 2*) et CDK4 sans modification de leur ARNm, ce qui suggère que cette baisse est due à une augmentation de leur dégradation ¹⁹³. Par ailleurs, la signalisation SMYD2-HSP90 semble impliquée dans des mécanismes non oncologiques, car elle a été décrite comme régulant l'organisation des myofilaments des muscles squelettiques et cardiaque ¹⁹⁴.

d) ER α

SMYD2 méthyle ER α au niveau de K266 avec un effet répresseur sur son activité transcriptionnelle. En situation de stimulation œstrogénique, cette marque de méthylation est diminuée, ce qui favorise l'activité de transactivation de ER α . A l'inverse, en l'absence de stimulation œstrogénique, cette méthylation s'intensifie et impacte le recrutement de ER α par la chromatine dans des lignées de carcinome mammaire telles que MCF7. Cette perte de fonction est due au fait que la méthylation de K266 gêne l'interaction de ER α avec le complexe p300/cAMP, ce qui a pour conséquence de diminuer le niveau d'acétylation de ER α -K266/268 ¹⁹⁵.

e) PTEN

L'activité du suppresseur de tumeur PTEN (*Phosphatase and tensin homologue*) est régulée négativement par la méthylation de sa lysine K313 par SMYD2. Cette lipide phosphatase inhibe la voie de signalisation PI3K-AKT via la conversion du phosphatidylinositol-3,4,5-triphosphate en phosphatidylinositol-4,5-biphosphate, ce qui antagonise la voie de

signalisation oncogénique AKT. La méthylation de PTEN par SMYD2 semble responsable d'une diminution de sa phosphorylation au niveau de la sérine 380, ce qui inhibe son activité et favorise donc la prolifération tumorale ¹⁹⁶.

f) PARP1

SMYD2 méthyle également la protéine PARP1 (Poly-ADB Ribose Polymérase 1), qui joue un rôle essentiel dans la réparation des cassures simple brin de l'ADN. En l'absence de PARP1 le blocage de la fourche de réplication entraîne l'accumulation de cassures double-brins qui sont réparées par recombinaison homologue. PARP1 est donc impliquée dans la carcinogénèse et la chimiorésistance. Sa méthylation par SMYD2 sur sa lysine K528 amplifie son activité ¹⁴¹.

g) STAT3

Bien qu'aucune caractérisation mécanistique n'ait été réalisée, la régulation à la baisse de SMYD2 dans les cellules de cancer du sein entraîne une diminution globale du niveau de méthylation de STAT3. On ignore si SMYD2 méthyle directement STAT3 dans le contexte d'adénocarcinome mammaire triple négatif, ou si le signal de méthylation résulte d'autres KMT dont l'activité est régulée par SMYD2 ¹⁸⁷. De même, la surexpression de SMYD3 semble associée à l'augmentation de l'activation de STAT3 dans le cancer gastrique, ce qui constitue un facteur de mauvais pronostic ¹⁹⁷.

h) MAPKAPK3

Reynoird et coll. (2016) ont montré que la méthylation de MAPKAPK3 (MAP kinase-activated protein kinase 3) sur sa lysine K355 par SMYD2 était impliquée dans la prolifération tumorale des cancers du pancréas dans un contexte inflammatoire ¹⁹². En effet, MAPKAPK3 est impliquée dans l'inflammation et la réponse au stress, notamment par l'intermédiaire d'IL6 et de TNF α ¹⁹⁸, et sa délétion dans les lignées cellulaires et les modèles de xénogreffe de cellules d'adénocarcinome est associée à une baisse de la prolifération tumorale. De

même, son inhibition pharmacologique chez les souris $Ptf1a^{+/Cre};Kras^{+/LSL-G12D};Trp53^{loxP/loxP}$ était associée à une diminution de la progression tumorale. Par ailleurs, cette MAP kinase peut être phosphorylée par un certain nombre d'enzymes essentielles dans la carcinogenèse telles que ERK1/2, p38 et SAPK (Stress-activated protein kinases)/JNK (Jun amino-terminal kinases) ^{199,200}, ce qui renforce son rôle oncogénique. Néanmoins, la caractérisation des véritables conséquences phénotypiques de sa méthylation par SMYD2 reste incomplète.

D'autres substrats ont été identifiés sans caractérisation précise des conséquences de leur méthylation par SMYD2 tels que p65 Rel ²⁰¹, BMPR2 (Bone Morphogenetic Protein Receptor type 2) ²⁰² and ALK (Anaplastic Lymphoma Kinase) ²⁰³. Enfin, des nouveaux substrats potentiels ont été révélés par des analyses récentes de spectrométrie de masse à grande échelle après un marquage isotopique stable avec des acides aminés en culture cellulaire (SILAC), couplé à un enrichissement par immunoaffinité de peptides porteurs d'une lysine monométhylée, tels que BTF3 (Basic Transcription Factor 3), PDAP1 (PDGFA Associated Protein 1), AHNAK (Neuroblast differentiation-associated protein), et AHNAK2 dans des cellules de carcinome épidermoïde oesophagien ²⁰⁴. Ces données restent très préliminaires et nécessitent de multiples étapes de validation et de caractérisation des conséquences phénotypiques de ces événements de méthylation.

En résumé, SMYD2 semble jouer un rôle pivot dans la carcinogenèse, avec un nombre croissant de substrats décrits dont la méthylation aboutit par divers mécanismes à une augmentation des signaux de prolifération tumorale. L'inhibition pharmacologique de cette dernière représente donc une piste prometteuse dans le traitement du cancer.

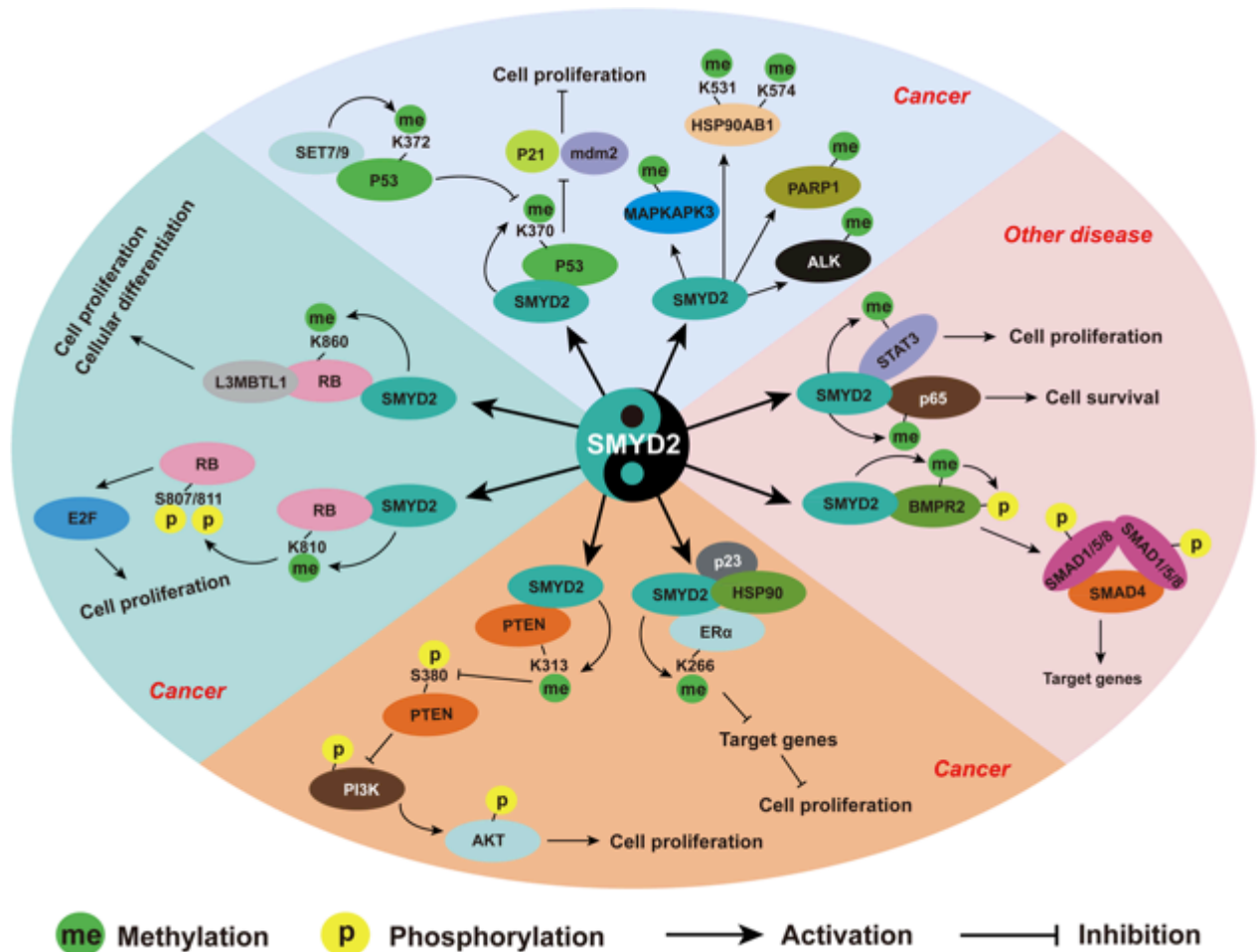


Figure 17 : Résumé des différentes signalisations impliquant SMYD2 (Xin et al. Clinical Epigenetics 2019). Plusieurs protéines non-histone sont méthylées par SMYD2 pendant la carcinogenèse, telles que p53, RB, PTEN, ER α , PARP1, ALK, MAPKAPK3 et HSP90AB1. Ces méthylations influencent souvent d'autres PTMs et sont elles-mêmes impactées par d'autres PTMs, en particulier la phosphorylation.

3. Inhibition pharmacologique de SMYD2

Dans des modèles précliniques de souris *Smyd2*^{-/-}, aucune modification phénotypique majeure et aucun problème de fertilité à l'âge adulte ne sont observés¹⁹². Cette absence d'effets biologiques majeurs suggère que l'inhibition de SMYD2 dans les cancers pourrait entraîner peu d'effets secondaires, ce qui renforce son intérêt pour le développement de nouvelles thérapies ciblées anti-tumorales.

AZ505 est le premier inhibiteur de SMYD2 identifié. C'est un puissant inhibiteur, avec une concentration inhibitrice submicromolaire (IC₅₀ = 0,12 μ M). Il se lie au niveau du site de

liaison peptidique de SMYD2 et inhibe de manière compétitive la liaison aux substrats tel que p53. AZ505 présente une grande sélectivité envers SMYD2, notamment en comparaison aux autres KMTs telles que DOT1L ou EZH2²⁰⁵. Dans des cellules de cancer du sein triple-négatif, l'inhibiteur de SMYD2 AZ505 a montré une efficacité anti-tumorale contre un contrôle par DMSO, chez des souris *nude* avec xénotransplante de cellules MDA-MB231¹⁸⁷. Plus récemment, l'inhibiteur actif de SMYD2 LLY-507 a été testé et présente une sélectivité accrue pour SMYD2 environ 100 fois supérieure par rapport à 24 autres méthyltransférases et une concentration inhibitrice nanomolaire (IC50 <15 nM). L'étude structurale en cristallographie a montré que LLY-507 se lie à SMYD2 au niveau de sa poche de liaison au substrat. Il inhibe la capacité de SMYD2 à méthyler p53 avec une IC50 <15 nM. LLY-507 permet une inhibition dose-dépendante de la prolifération dans plusieurs lignées cellulaires tumorales²⁰⁶.

BAY-598 est un inhibiteur compétitif de SMYD2 avec une sélectivité 100 fois supérieure dans un panel de 32 méthyltransférases. Il semble potentialiser l'efficacité de la doxorubicine dans un modèle de xénotransplante de cellules de cancer de l'œsophage KYSE-150²⁰⁷. De même, lors de l'étude validant MAPKAPK3 comme substrat de SMYD2, BAY-598, en bithérapie avec la gemcitabine, a permis une action synergique, avec un effet antitumoral supérieur à la gemcitabine seule dans l'adénocarcinome pancréatique chez la souris¹⁹².

C. IMPLICATIONS DE SMYD3 DANS LE CANCER

1. Expression dans les cancers

En clinique, la surexpression de SMYD3 semble constituer un facteur de mauvais pronostic indépendant dans de nombreux cancers. Sa surexpression semble constituer un facteur de mauvais pronostic dans le carcinome épidermoïde œsophagien (Figure 18)²⁰⁸, le cancer de l'estomac^{197,209}, les cancers du poumon²¹⁰, le cancer colorectal²¹¹, le carcinome hépatocellulaire²¹², les leucémies²¹³, le cancer du sein²¹⁴, le cancer de la prostate²¹⁵ et les cholangiocarcinomes^{216,217}. Le niveau de démonstration de son intérêt pronostique reste

néanmoins faible à l'heure actuelle car il est souvent basé sur des études rétrospectives avec des effectifs restreints de patients.

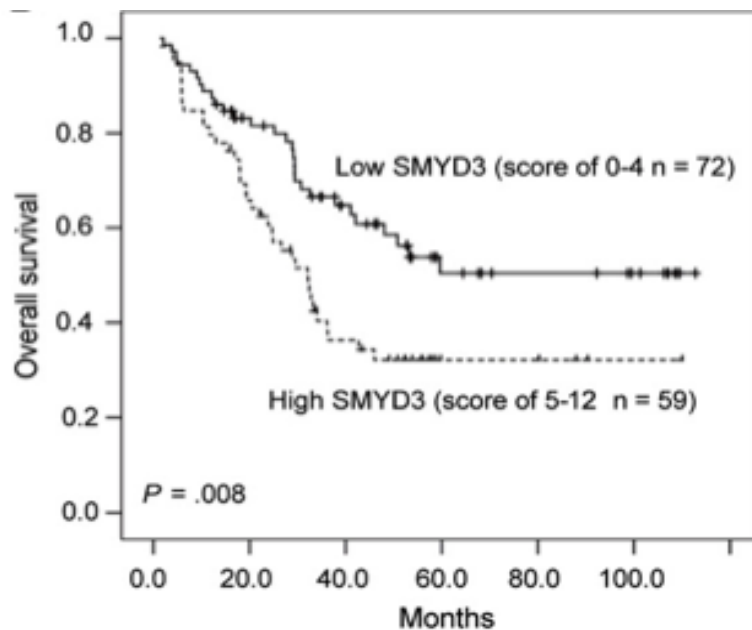


Figure 18 : Impact pronostic de l'expression de SMYD3 sur la survie globale dans le carcinome malpighien de l'œsophage (Zu et al. Human Pathology 2016). Courbe de survie globale selon Kaplan Meier des patients porteurs d'un carcinome malpighien de l'œsophage, surexprimant (High SMYD3) ou non (Low SMYD3) SMYD3 (n = 131 ; p < 0.05, log-rank test).

Le rôle oncogénique de SMYD3 est complexe et varie d'un cancer à l'autre. Elle est impliquée dans la prolifération tumorale des adénocarcinomes pancréatiques dans lesquels elle est surexprimée dans 9% des cas. Dans cette pathologie, elle est impliquée dans la néo-angiogenèse, l'invasion tumorale, le potentiel métastatique, et la trans-différenciation acino-canaulaire (première étape de la séquence *Pancreatic Intraepithelial neoplasia (PanIN)*-1, -2, -3, adénocarcinome)²¹⁰. Le rôle de SMYD3 a également été décrit dans la carcinogenèse des adénocarcinomes colorectaux et des carcinomes hépatocellulaires¹⁸². En effet, chez les souris traitées par diéthylnitrosamine (DEN), un carcinogène qui permet l'obtention de tumeurs hépatiques et coliques, la délétion de SMYD3 est associée à une baisse importante

la prolifération tumorale. Néanmoins, cette dernière n'étant pas complètement abolie, ce qui suggère que SMYD3 est importante pour la prolifération tumorale mais non indispensable. De plus, SMYD3 ne semble pas suffisante car sa surexpression au niveau des hépatocytes sans exposition au DEN, ne permet pas l'obtention de tumeurs hépatiques (Figure 19). Dans ce modèle, SMYD3 est notamment retrouvée au niveau nucléaire et elle semble participer l'expression de nombreux gènes impliqués dans la tumorigenèse. SMYD3 est recrutée au niveau des promoteurs de gènes impliqués dans le cycle cellulaire et la transition épithélio-mésenchymateuse colique et hépatique, étape essentielle dans l'initiation tumorale et la diffusion métastatique. Elle stimule l'expression de multiples gènes impliqués dans ce phénomène dont Snai1 (Snail Family Transcriptional Repressor 1), Fn1 (fibronectin 1) ou Mmp9 (Matrix Metallopeptidase 9), sans modification de la E-cadhérine, protéine clé dans ce processus. Sa délétion entraîne également une perte d'expression d'oncogènes canoniques tels que Myc et Ctnb1 (Catenin Beta 1). Enfin, dans ce modèle, SMYD3 est associée à une augmentation de l'expression de la forme phosphorylée de STAT3 (pSTAT3), synonyme d'activation de la voie de signalisation JAK/STAT3. Cet effet semble indépendant d'IL6, le principal activateur de JAK/STAT3, et passe par l'augmentation de la transcription de Jak1 et Jak2 ¹⁸². Par ailleurs, dans les carcinomes épidermoïdes œsophagiens, SMYD3 amplifie les signaux de prolifération, de migration et d'invasion via la stimulation de la transcription de EZR (Ezrin) et de LOXL2 (Lysyl Oxidase Like 2) ²⁰⁸. Elle est également décrite dans les cancers mammaires dans lesquels, elle stimule la prolifération des cellules tumorales *in vitro* via la transcription de WNT10B (Wnt Family Member 10B) ²¹⁸.

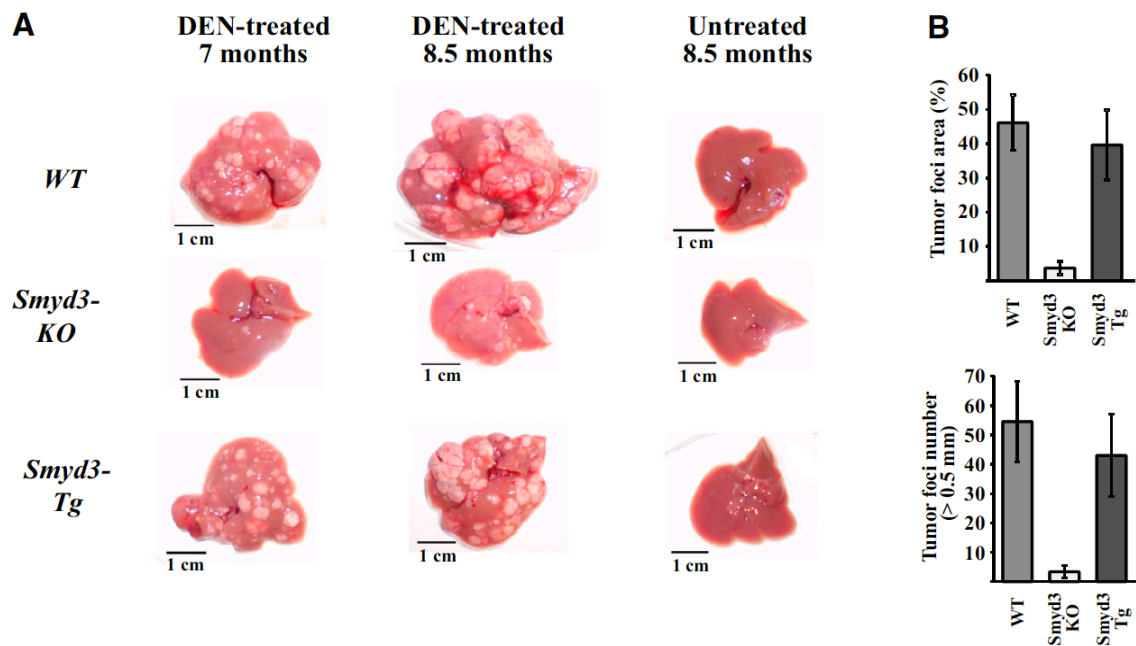


Figure 19 : Impact de SMYD3 sur la prolifération tumorale dans modèle de carcinome hépatocellulaire induit par DEN (Sarris *et al.* Cancer Cell 2016). (A) Illustration de l'aspect macroscopique de foies de souris traitées par DEN à pendant 7 et 8,5 mois, et de souris non traitées à 8,5 mois. Pour chaque condition, un groupe de souris dites sauvages (WT), de souris *Smyd3*^{-/-} (*Smyd3*-KO) et de souris avec surexpression de *Smyd3* au niveau hépatocytaire (*Smyd3*-Tg) ont été étudiées. (B) Quantification du nombre de foyers tumoraux dans les souris traitées par DEN. (moyenne \pm erreur standard de la moyenne (\pm SEM) ; n =

La localisation cellulaire de SMYD3 varie d'un cancer à l'autre, avec une expression plutôt nucléaire dans des modèles de cancer colorectal et de carcinome hépatocellulaire, et cytoplasmique dans des modèles de cancers du pancréas et du poumon dépendants de la voie RAS. À ce jour, le rôle nucléaire de SMYD3 semble principalement lié à la régulation de l'expression de nombreux gènes impliqués dans la tumorigenèse. Il est probablement médié par la méthylation des histones et autres protéines participant au contrôle de l'activité transcriptionnelle mais sa caractérisation reste très lacunaire. Son rôle cytoplasmique est également très partiellement décrit, mais ces dernières années, la caractérisation de la méthylation par SMYD3 d'acteurs majeurs de l'oncogenèse, renforce l'importance de cette KMT dans la régulation de cibles non-histones impliquées dans le

cancer. Les paragraphes suivants détaillent les différentes marques de méthylation catalysées par SMYD3 identifiées à ce jour dans la cancérogenèse.

2. SMYD3 : marques de méthylation impliquées dans la carcinogenèse

a) SMYD3-Histone H4 module la chromatine et promeut la prolifération tumorale

SMYD3 a été initialement décrite comme méthylant H3K4 et H4K20^{176,177}. La première cible s'est avérée être un artéfact *in vitro*, non retrouvé en cellules et sans conséquence phénotypique¹¹⁵. En effet, Van Aller et coll. (2012) ont démontré que la mutation de la lysine K4 en arginine, afin d'empêcher sa méthylation par SMYD3, ne modifiait pas son activité. Grâce à la réalisation de mutations ponctuelles des différentes lysines candidates, le site de méthylation a été localisé sur la lysine K5 de l'histone H4. De plus, en spectrométrie de masse sur extraits cellulaires, l'état de méthylation de H3K4 et de H4K20 ne semblait pas modifié, en situation de délétion ou de surexpression de SMYD3, à l'inverse de H4K5. Ce travail a aussi montré que la délétion de SMYD3 dans la lignée de carcinome hépatocellulaire Hep3B cause une diminution de la prolifération cellulaire et la perte de la capacité à former des colonies. Le même effet s'observe aussi dans la lignée de carcinome mammaire MCF7. La restauration de l'expression de SMYD3 contrecarre ces effets cellulaires, contrairement à la surexpression du mutant SMYD3 F183A dont le site catalytique de méthylation n'est pas fonctionnel¹¹⁵.

b) SMYD3-MAP3K2 favorise l'initiation et la progression tumorales

Une avancée majeure dans la caractérisation de l'implication de SMYD3 dans la carcinogenèse et la progression tumorale a été la démonstration de son rôle dans la méthylation de MAP3K2 (Mitogen-Activated Protein Kinase Kinase Kinase 2), qui appartient à la voie de signalisation Ras/Raf/MEK/ERK. Cette fonction a été décrite dans les adénocarcinomes bronchiques RAS-

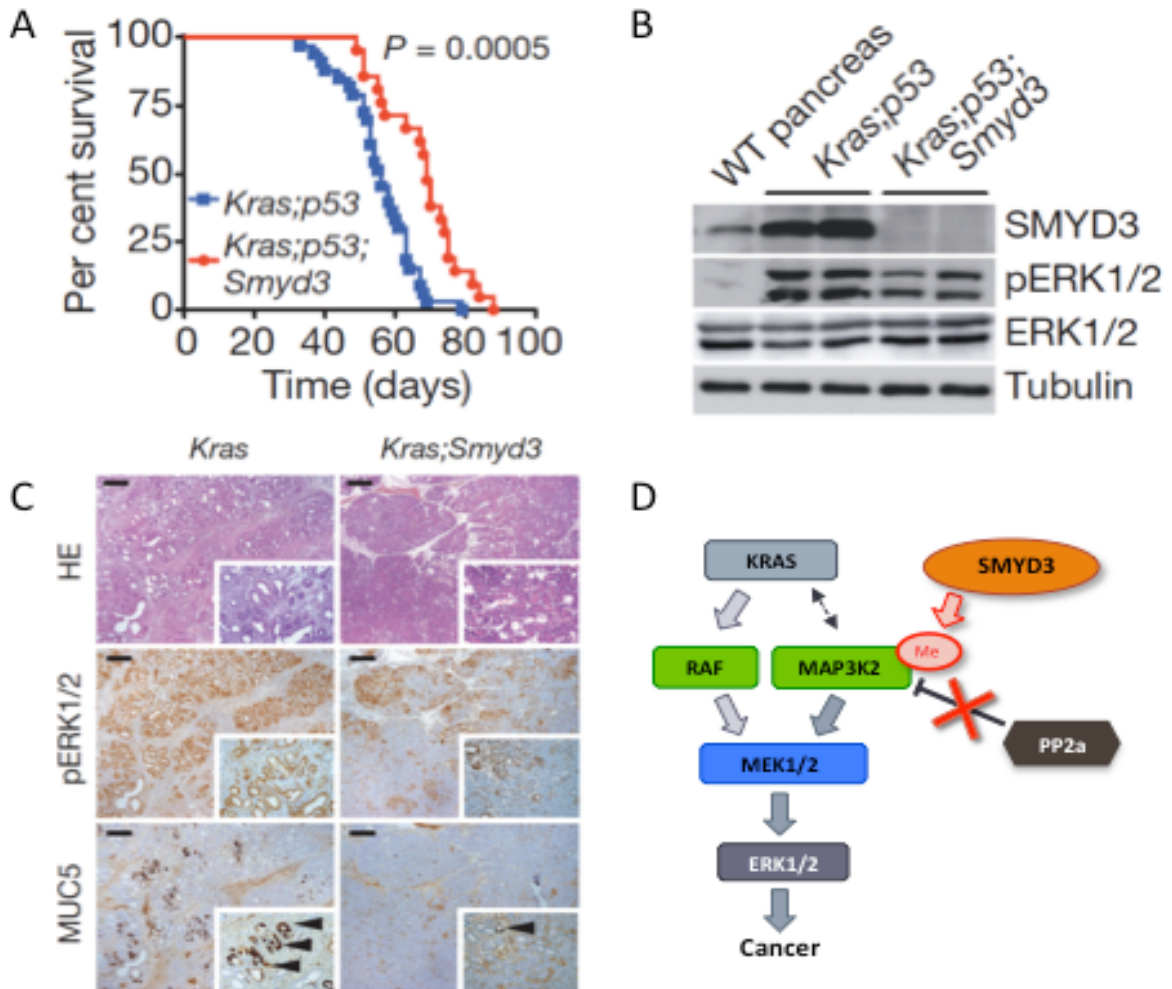


Figure 20 : Signalisation SMYD3-MAP3K2 dans l'adénocarcinome pancréatique (Figure adaptée à partir de Mazur, Reynold et al. Nature 2014). (A) Courbes de survie selon Kaplan-Meier de souris $p48^{+/Cre};Kras^{+/LSL-G12D};Trp53^{loxP/loxP}$ (*Kras;p53*) et de souris $p48^{+/Cre};Kras^{+/LSL-G12D};Trp53^{loxP/loxP};Smyd3^{-/-}$ (*Kras;p53;Smyd3*) avec survie médiane respective de 521 jours et 568,5 jours ($p=0.0005$; test du log-rank). (B) Western Blot avec immunomarquage SMYD3, pERK1/2, ERK1/2 et contrôle tubuline sur des lysats de tumeurs issues de souris *Kras;p53* et *Kras;p53;Smyd3*. (C) Coupes histologiques d'adénocarcinomes pancréatiques avec coloration par hématoxyline et éosine (HE), et immunomarquage pERK1/2 (marqueur de la voie RAS) et MUC5 (marqueur des lésions préneoplasiques de type PanIN (flèches noires)), issues de souris *Kras;p53* et *Kras;p53;Smyd3* (D) Résumé des conséquences de la méthylation de MAP3K2 par SMYD3 sur la voie de signalisation RAS.

dépendants et dans les adénocarcinomes pancréatiques. Dans ces pathologies, SMYD3 semble être principalement cytoplasmique et triméthyle MAP3K2, ce qui bloque son

interaction avec sa phosphatase régulatrice PP2. Ainsi, l'action de SMYD3 a pour conséquence d'activer MAP3K2 de manière irréversible, ce qui amplifie le signal RAS et donc la phosphorylation de ERK1/2 (Figures 20B et C), avec pour conséquence la prolifération cellulaire et l'inhibition de l'apoptose. Ces données ont été démontrées par test de méthylation *in vitro* en lignées cellulaires, puis validées en modèle murin d'adénocarcinome pancréatique avec souris transgéniques $p48^{+/Cre};Kras^{+/LSL-G12D};Trp53^{loxP/loxP}$ (*Kras;p53*) et d'adénocarcinome pulmonaire avec souris $Kras^{+/LSL-G12D};Trp53^{loxP/loxP}$ traitées par instillation intra-pulmonaire d'adénovirus exprimant la Cre recombinase. Chez les souris *Kras;p53*, la croissance tumorale est ralentie (Figure 20A) tandis que la restitution de l'expression de SMYD3 augmente la croissance tumorale. La méthylation de MAP3K2 a été démontrée comme amplifiant la signalisation RAS, mais ce mécanisme reste accessoire et son inhibition ne bloque pas à elle seule la prolifération dépendante de la voie RAS ¹⁸¹ (Figure 20D).

c) SMYD3-VEGFR1 stimule la néoangiogenèse

SMYD3 triméthyle VEGFR1 (Vascular endothelial growth factor receptor 1) sur sa lysine K831, un récepteur surexprimé dans de nombreux cancers tels que le carcinome hépatocellulaire, le cancer du sein ou le cancer colorectal. Cette méthylation se situe à seulement 3 acides aminés en N-ter du motif GXGXXG hautement conservé au sein du site catalytique de la kinase et semble augmenter l'activité kinase du récepteur VEGFR1 ²¹⁹.

d) SMYD3-AKT1 potentialise l'activité de AKT1

La protéine AKT1 (RAC-alpha serine/threonine-protein kinase) est une serine/thréonine kinase de la voie de la signalisation PI3K/AKT/mTOR, surexprimée dans de nombreux cancers et dont le rôle a été bien établi dans la carcinogenèse ²²⁰. De nombreuses PTMs ont été décrites, telles que la phosphorylation ou l'ubiquitinylation et récemment la triméthylation de sa lysine 14 par SMYD3. Cette méthylation survient dans le domaine PH (*pleckstrin homology*) essentiel à son activation et est associée à une augmentation de la

phosphorylation sur la thréonine 308, promouvant ainsi la progression tumorale. A l'inverse, la répression de SMYD3, de même que la mutation de AKT1 K14, est associée à une diminution de son niveau de phosphorylation dans les cellules cancéreuses ²²¹.

e) SMYD3-HER2 augmente l'activité oncogénique de HER2

HER2 (Human Epidermal Growth Factor Receptor-2) est un récepteur à activité tyrosine kinase amplifié dans de nombreux cancers dont notamment les cancers du sein et de l'estomac. SMYD3 triméthyle HER2 sur sa lysine K175, induisant ainsi son homodimérisation avec une augmentation de son autophosphorylation et donc de son activité oncogénique ²²².

En résumé, SMYD3 a une action oncogénique via la régulation de la transcription par les modifications conformationnelles de la chromatine (méthylation H4K5) et par l'activation de voies de signalisation majeures telles que MAP3K2 dans la voie RAS, la néoangiogenèse tumorale via VEGFR1, ou encore AKT1 et HER2 (Figure 21).

3. Inhibition pharmacologique de SMYD3

La répression de SMYD3 impacte la prolifération tumorale dans de nombreux modèles cellulaires *in vitro*, tels que l'adénocarcinome du pancréas et du poumon dépendant de la voie Ras, le cancer colorectal, ou encore le carcinome hépatocellulaire ^{181,182}. L'inhibiteur BCI-121 réduit l'activité de SMYD3 *in vitro* dans des cellules de cancer colorectal, avec une diminution du niveau de méthylation de H4K5me3, de même que dans d'autres modèles de cancer du poumon, de la prostate, des ovaires ou du pancréas ²²³. Deux autres molécules, le sulfonamide

EPZ031686 et le sulfamide EPZ-030456 ont montré des résultats prometteurs *in vitro* et des données satisfaisantes de biodisponibilité ont été obtenues *in vivo* pour EPZ031686, sans donnée d'efficacité ²²⁴. La novobiocine, un inhibiteur de HSP90, semble bloquer l'interaction de HSP90 avec SMYD3 et diminuer l'activité de cette dernière, entraînant une baisse de la

prolifération et de la migration des cellules tumorales de manière dose-dépendante dans des cellules de cancer du sein telles que les MDA-MB-231 ²²⁵. Plusieurs petites molécules inhibant SMYD3 sont en cours de développement mais aucune n'offre à l'heure actuelle des résultats satisfaisants en matière d'efficacité anti-tumorale en phase préclinique. Le développement de combinaisons de thérapies inhibant à plusieurs niveaux les signalisations impliquant SMYD3 est donc une piste plus prometteuse, tel qu'un double blocage de la voie RAS en associant un inhibiteur de SMYD3 avec un inhibiteur de MEK-1 en cas de tumeur KRAS mutée par exemple, d'autant plus qu'une surexpression de SMYD3 est observée dans cette situation ²²⁶.

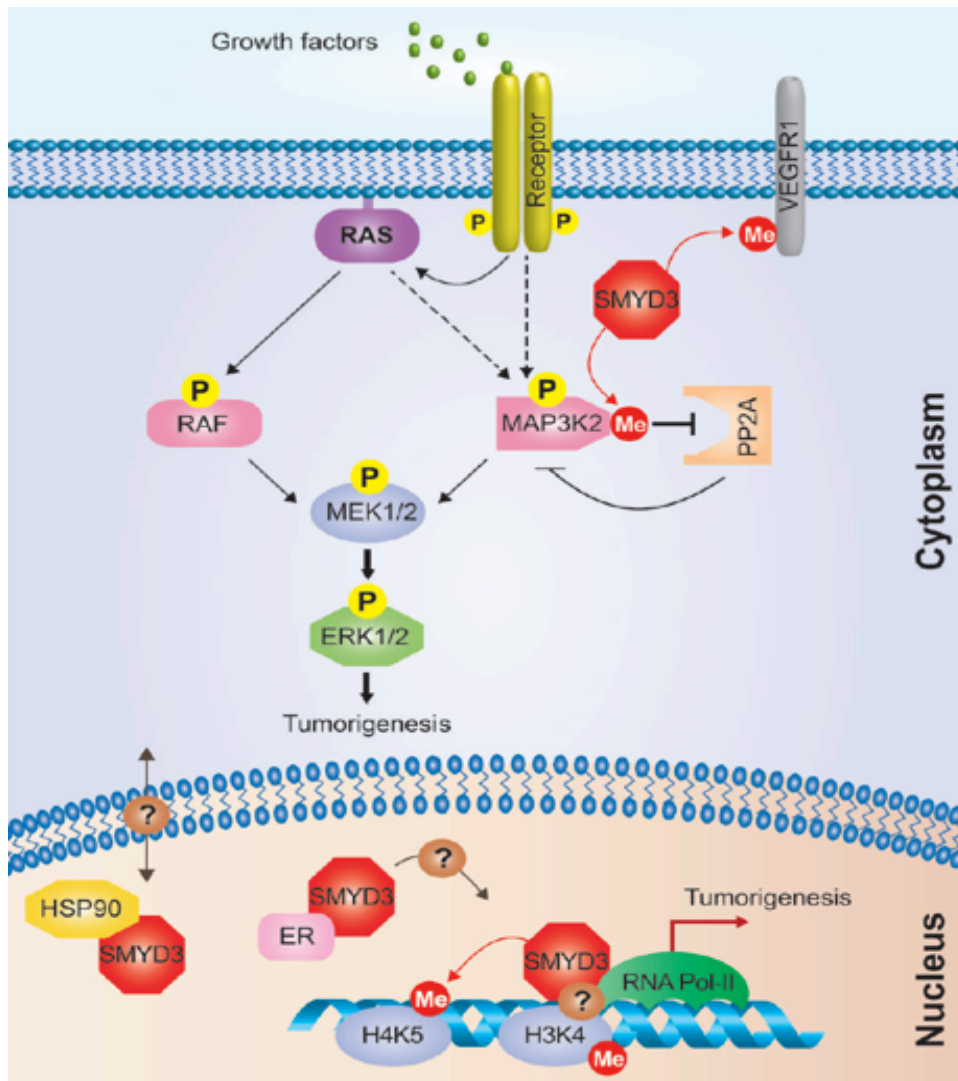


Figure 21 : Résumé des fonctions nucléaires et cytoplasmiques de SMYD3 (Mazur et al. Transl Cancer Res 2016). Au niveau cytoplasmique, SMYD3 méthyle MAP3K2 ce qui bloque son interaction avec sa phosphatase PP2A, aboutissant à une amplification du signal RAS. La méthylation de VEGFR1 constitue un autre mécanisme d'amplification de la prolifération tumorale par SMYD3. Au niveau nucléaire, SMYD3 semble faciliter la transcription en se fixant à la fois à l'ADN, à l'ARN-polymérase II, ainsi qu'en méthylant H4K5 avec un impact sur le remodelage de la chromatine. L'interaction de SMYD3 avec HSP90 joue peut-être un rôle dans la présence de SMYD3 dans le noyau. SMYD3 se lie au récepteur des œstrogènes ER, influençant son activité de facteur de transcription.

CONCLUSION

Le protéome est extrêmement riche et d'une très grande diversité, d'un type cellulaire à l'autre, à différents temps de la vie d'un organisme, et d'une situation physiologique ou pathologique à l'autre. Cette grande diversité est le fruit de multiples étapes pré-transcriptionnelle, post-transcriptionnelles, traductionnelles et post-traductionnelles, qui induisent une formidable variabilité dans l'expression des gènes, puis des protéines et de leur fonctionnalité. En effet, au niveau pré- et per-transcriptionnel, les différents éléments tels que la régulation de la chromatine, la méthylation de l'ADN et les différentes PTMs survenant au niveau des histones, vont favoriser la production de certains ARNm, en répondant à des mécanismes de régulation complexes qui mettent en jeu les événements sus-cités, de manière parfois inter-dépendante, synergique, voire compétitive. Ensuite, l'épissage, suivi des différentes modifications post-traductionnelles, va parachever la grande diversité du protéome en générant de multiples transcrits traduits en de multiples protéines, avec des fonctions diverses selon leur séquence et les différentes PTMs qui les modifient. La méthylation est une PTM particulièrement importante dans les signalisations cellulaires au niveau des histones et de nombreuses protéines non-histone. La méthylation des lysines, est un mécanisme de régulation fin et complexe, qui va impacter les interactions protéine-protéine et des fonctions protéiques précises, influençant ainsi le fonctionnement des cellules dans de nombreuses situations physiologiques et pathologiques. La méthylation des lysines a été largement étudiée dans le cancer, dans lequel ses implications sont multiples. Deux KMTs de la même famille, SMYD2 et SMYD3 sont surexprimées dans de nombreux cancers et sont souvent associées à un mauvais pronostic. Elles interviennent dans des signalisations canoniques de la carcinogenèse. SMYD2 par exemple va réguler négativement l'activité de suppresseurs de tumeurs, tels que p53, PTEN, Rb, ou augmenter l'activité de systèmes de réparation de l'ADN, tels que PARP1, favorisant ainsi la chimiorésistance. SMYD3 amplifie le signal oncogénique RAS dans certains cancers en

méthylant la kinase MAP3K2. De même, elle active différentes voies oncogéniques via la méthylation de VEGFR1, HER2 ou AKT1.

Le caractère ubiquitaire de ces protéines, leur surexpression dans un grand nombre de cancers, et les récentes données immatures issues d'études de spectrométrie de masse ou d'analyses bioinformatiques, suggèrent que SMYD2 et SMYD3 sont associées à d'autres signalisations impliquées dans la carcinogenèse, dont la caractérisation conditionne la compréhension du rôle oncogénique complexe de ces KMTs et le développement de nouvelles approches thérapeutiques.

L'objectif de la première partie de ce travail de thèse consiste en l'identification et la confirmation *in vitro* de nouveaux substrats méthylés par SMYD2 et SMYD3, afin d'initier des projets de recherche sur des signalisations d'intérêt influencées par ces événements de méthylation.

La seconde partie de cette thèse porte sur la caractérisation moléculaire précise des conséquences phénotypiques de la méthylation de RNF113A par SMYD3 dans un modèle de cancer du poumon à petites cellules.

REFERENCES

- 1 Fincham, J. R. S. Epigenetic Mechanisms of Gene Regulation. Edited by V. E. A. Russo, R. A. Martienssen and A. D. Riggs. Cold Spring Harbor Laboratory Press, 1996. 693+xii pages. Price \$125. ISBN 0 87969 490 4. *Genetical Research* 69, 159-162, doi:10.1017/S0016672397229320 (1997).
- 2 Luger, K. & Hansen, J. C. Nucleosome and chromatin fiber dynamics. *Curr Opin Struct Biol* 15, 188-196, doi:10.1016/j.sbi.2005.03.006 (2005).
- 3 Talbert, P. B. & Henikoff, S. Histone variants on the move: substrates for chromatin dynamics. *Nat Rev Mol Cell Biol* 18, 115-126, doi:10.1038/nrm.2016.148 (2017).
- 4 Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell research* 21, 381-395, doi:10.1038/cr.2011.22 (2011).
- 5 Jenuwein, T. & Allis, C. D. Translating the Histone Code. *Science* 293, 1074-1080, doi:10.1126/science.1063127 (2001).
- 6 Wolffe, A. P. Developmental regulation of chromatin structure and function. *Trends in Cell Biology* 1, 61-66, doi:https://doi.org/10.1016/0962-8924(91)90091-M (1991).
- 7 Falkenberg, K. J. & Johnstone, R. W. Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders. *Nat Rev Drug Discov* 13, 673-691, doi:10.1038/nrd4360 (2014).
- 8 Allfrey, V. G., Faulkner, R. & Mirsky, A. E. ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS. *Proceedings of the National Academy of Sciences* 51, 786-794, doi:10.1073/pnas.51.5.786 (1964).
- 9 Wang, Z. *et al.* Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet* 40, 897-903, doi:10.1038/ng.154 (2008).
- 10 Rossetto, D., Avvakumov, N. & Cote, J. Histone phosphorylation: a chromatin modification involved in diverse nuclear events. *Epigenetics* 7, 1098-1108, doi:10.4161/epi.21975 (2012).
- 11 Seeler, J. S. & Dejean, A. Nuclear and unclear functions of SUMO. *Nat Rev Mol Cell Biol* 4, 690-699, doi:10.1038/nrm1200 (2003).
- 12 Hershko, A. & Ciechanover, A. THE UBIQUITIN SYSTEM. *Annual Review of Biochemistry* 67, 425-479, doi:10.1146/annurev.biochem.67.1.425 (1998).
- 13 Zhang, T., Cooper, S. & Brockdorff, N. The interplay of histone modifications – writers that read. *EMBO reports* 16, 1467-1481, doi:10.15252/embr.201540945 (2015).
- 14 Rea, S. *et al.* Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406, 593-599, doi:10.1038/35020506 (2000).
- 15 Lee, J. S. *et al.* Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. *Cell* 131, 1084-1096, doi:10.1016/j.cell.2007.09.046 (2007).
- 16 Bartke, T. *et al.* Nucleosome-interacting proteins regulated by DNA and histone methylation. *Cell* 143, 470-484, doi:10.1016/j.cell.2010.10.012 (2010).
- 17 Venner, H. & Reinert, H. Possible role of methylated DNA bases for the transcription of the genetic information. *Zeitschrift für allgemeine Mikrobiologie* 13, 613-624, doi:10.1002/jobm.19730130711 (1973).

- 18 Watt, F. & Molloy, P. L. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes & Development* 2, 1136-1143, doi:10.1101/gad.2.9.1136 (1988).
- 19 Waterland, R. A. & Jirtle, R. L. Transposable Elements: Targets for Early Nutritional Effects on Epigenetic Gene Regulation. *Molecular and Cellular Biology* 23, 5293-5300, doi:10.1128/mcb.23.15.5293-5300.2003 (2003).
- 20 Schubeler, D. Function and information content of DNA methylation. *Nature* 517, 321-326, doi:10.1038/nature14192 (2015).
- 21 Jones, P. A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* 13, 484-492, doi:10.1038/nrg3230 (2012).
- 22 Tahiliani, M. *et al.* Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science* 324, 930-935, doi:10.1126/science.1170116 (2009).
- 23 Dor, Y. & Cedar, H. Principles of DNA methylation and their implications for biology and medicine. *The Lancet* 392, 777-786, doi:10.1016/s0140-6736(18)31268-6 (2018).
- 24 Nuber, U. A. *et al.* Up-regulation of glucocorticoid-regulated genes in a mouse model of Rett syndrome. *Human Molecular Genetics* 14, 2247-2256, doi:10.1093/hmg/ddi229 (2005).
- 25 Tudor, M., Akbarian, S., Chen, R. Z. & Jaenisch, R. Transcriptional profiling of a mouse model for Rett syndrome reveals subtle transcriptional changes in the brain. *Proceedings of the National Academy of Sciences* 99, 15536-15541, doi:10.1073/pnas.242566899 (2002).
- 26 Sansom, O. J. *et al.* Deficiency of Mbd2 suppresses intestinal tumorigenesis. *Nature Genetics* 34, 145-147, doi:10.1038/ng1155 (2003).
- 27 Cook, P. C. *et al.* A dominant role for the methyl-CpG-binding protein Mbd2 in controlling Th2 induction by dendritic cells. *Nat Commun* 6, 6920, doi:10.1038/ncomms7920 (2015).
- 28 Brown, S. E., Suderman, M. J., Hallett, M. & Szyf, M. DNA demethylation induced by the methyl-CpG-binding domain protein MBD3. *Gene* 420, 99-106, doi:10.1016/j.gene.2008.05.009 (2008).
- 29 Hendrich, B., Hardeland, U., Ng, H.-H., Jiricny, J. & Bird, A. The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature* 401, 301-304, doi:10.1038/45843 (1999).
- 30 Laget, S. *et al.* The human proteins MBD5 and MBD6 associate with heterochromatin but they do not bind methylated DNA. *PLoS One* 5, e11982, doi:10.1371/journal.pone.0011982 (2010).
- 31 Meilinger, D. *et al.* Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells. *EMBO Rep* 10, 1259-1264, doi:10.1038/embor.2009.201 (2009).
- 32 Filion, G. J. *et al.* A family of human zinc finger proteins that bind methylated DNA and repress transcription. *Mol Cell Biol* 26, 169-181, doi:10.1128/MCB.26.1.169-181.2006 (2006).
- 33 Baylin, S. B. & Jones, P. A. A decade of exploring the cancer epigenome - biological and translational implications. *Nature reviews. Cancer* 11, 726-734, doi:10.1038/nrc3130 (2011).
- 34 Ley, T. J. *et al.* DNMT3A Mutations in Acute Myeloid Leukemia. *New England Journal of Medicine* 363, 2424-2433, doi:10.1056/NEJMoa1005143 (2010).

- 35 Ko, M. *et al.* Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* 468, 839-843, doi:10.1038/nature09586 (2010).
- 36 Bader, S. *et al.* Somatic frameshift mutations in the MBD4 gene of sporadic colon cancers with mismatch repair deficiency. *Oncogene* 18, 8044-8047, doi:10.1038/sj.onc.1203229 (1999).
- 37 Mongan, N. P., Emes, R. D. & Archer, N. Detection and analysis of RNA methylation. *F1000Res* 8, F1000 Faculty Rev-1559, doi:10.12688/f1000research.17956.1 (2019).
- 38 Huang, H. *et al.* Histone H3 trimethylation at lysine 36 guides m(6)A RNA modification co-transcriptionally. *Nature* 567, 414-419, doi:10.1038/s41586-019-1016-7 (2019).
- 39 Daffis, S. *et al.* 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature* 468, 452-456, doi:10.1038/nature09489 (2010).
- 40 Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470-476, doi:10.1038/nature07509 (2008).
- 41 Baralle, F. E. & Giudice, J. Alternative splicing as a regulator of development and tissue identity. *Nat Rev Mol Cell Biol* 18, 437-451, doi:10.1038/nrm.2017.27 (2017).
- 42 Lee, S. C. & Abdel-Wahab, O. Therapeutic targeting of splicing in cancer. *Nat Med* 22, 976-986, doi:10.1038/nm.4165 (2016).
- 43 Chabot, B. & Shkreta, L. Defective control of pre-messenger RNA splicing in human disease. *J Cell Biol* 212, 13-27, doi:10.1083/jcb.201510032 (2016).
- 44 Climente-Gonzalez, H., Porta-Pardo, E., Godzik, A. & Eyras, E. The Functional Impact of Alternative Splicing in Cancer. *Cell Rep* 20, 2215-2226, doi:10.1016/j.celrep.2017.08.012 (2017).
- 45 Simon, J. M. *et al.* Variation in chromatin accessibility in human kidney cancer links H3K36 methyltransferase loss with widespread RNA processing defects. *Genome Research* 24, 241-250, doi:10.1101/gr.158253.113 (2014).
- 46 Shahrouki, P. & Larsson, E. The non-coding oncogene: a case of missing DNA evidence? *Front Genet* 3, 170, doi:10.3389/fgene.2012.00170 (2012).
- 47 Carthew, R. W. & Sontheimer, E. J. Origins and Mechanisms of miRNAs and siRNAs. *Cell* 136, 642-655, doi:10.1016/j.cell.2009.01.035 (2009).
- 48 Mette, M. F., Aufsatz, W., Winden, J. v. d., Matzke, M. A. & Matzke, A. J. M. Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *The EMBO Journal* 19, 5194-5201, doi:10.1093/emboj/19.19.5194 (2000).
- 49 Sleutels, F., Zwart, R. & Barlow, D. P. The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* 415, 810-813, doi:10.1038/415810a (2002).
- 50 Brennecke, J. *et al.* Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* 128, 1089-1103, doi:10.1016/j.cell.2007.01.043 (2007).
- 51 Houwing, S. *et al.* A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. *Cell* 129, 69-82, doi:10.1016/j.cell.2007.03.026 (2007).
- 52 Pedrazzini, T. Le cœur des ARN non codants. *Med Sci (Paris)* 31, 261-267 (2015).
- 53 Khoury, G. A., Baliban, R. C. & Floudas, C. A. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci Rep* 1, doi:10.1038/srep00090 (2011).

- 54 Choudhary, C. *et al.* Lysine Acetylation Targets Protein Complexes and Co-Regulates Major Cellular Functions. *Science* 325, 834, doi:10.1126/science.1175371 (2009).
- 55 Tan, M. *et al.* Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 146, 1016-1028, doi:10.1016/j.cell.2011.08.008 (2011).
- 56 Chen, Y. *et al.* Lysine Propionylation and Butyrylation Are Novel Post-translational Modifications in Histones. *Molecular & Cellular Proteomics* 6, 812-819, doi:10.1074/mcp.M700021-MCP200 (2007).
- 57 Peng, C. *et al.* The First Identification of Lysine Malonylation Substrates and Its Regulatory Enzyme. *Molecular & Cellular Proteomics* 10, M111.012658, doi:10.1074/mcp.M111.012658 (2011).
- 58 Weinert, B. T. *et al.* Lysine succinylation is a frequently occurring modification in prokaryotes and eukaryotes and extensively overlaps with acetylation. *Cell Rep* 4, 842-851, doi:10.1016/j.celrep.2013.07.024 (2013).
- 59 Wisniewski, J. R., Zougman, A. & Mann, M. Nepsilon-formylation of lysine is a widespread post-translational modification of nuclear proteins occurring at residues involved in regulation of chromatin function. *Nucleic Acids Res* 36, 570-577, doi:10.1093/nar/gkm1057 (2008).
- 60 Hochstrasser, M. Origin and function of ubiquitin-like proteins. *Nature* 458, 422-429, doi:10.1038/nature07958 (2009).
- 61 van Slyke, D. D. & Sinex, F. M. THE COURSE OF HYDROXYLATION OF LYSINE TO FORM HYDROXYLYSINE IN COLLAGEN. *Journal of Biological Chemistry* 232, 797-806 (1958).
- 62 Johansen, M. B., Kiemer, L. & Brunak, S. Analysis and prediction of mammalian protein glycation. *Glycobiology* 16, 844-853, doi:10.1093/glycob/cwl009 (2006).
- 63 Lanouette, S., Mongeon, V., Figeys, D. & Couture, J. F. The functional diversity of protein lysine methylation. *Molecular Systems Biology* 10, 724-724, doi:10.1002/msb.134974 (2014).
- 64 Zhang, X., Huang, Y. & Shi, X. Emerging roles of lysine methylation on non-histone proteins. *Cell Mol Life Sci* 72, 4257-4272, doi:10.1007/s00018-015-2001-4 (2015).
- 65 Sims, R. J. & Reinberg, D. Is there a code embedded in proteins that is based on post-translational modifications? *Nat Rev Mol Cell Biol* 9, 815-820 (2008).
- 66 Hicke, L. Protein regulation by monoubiquitin. *Nature Reviews Molecular Cell Biology* 2, 195-201, doi:10.1038/35056583 (2001).
- 67 Walsh, C. T., Garneau-Tsodikova, S. & Gatto, G. J., Jr. Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew Chem Int Ed Engl* 44, 7342-7372, doi:10.1002/anie.200501023 (2005).
- 68 GAO, J. & XU, D. in *Biocomputing 2012* 94-103.
- 69 Radivojac, P. *et al.* Intrinsic disorder and functional proteomics. *Biophys J* 92, 1439-1456, doi:10.1529/biophysj.106.094045 (2007).
- 70 Wright, P. E. & Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol* 16, 18-29, doi:10.1038/nrm3920 (2015).
- 71 Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6, 197-208, doi:10.1038/nrm1589 (2005).

- 72 Theillet, F. X. *et al.* Cell signaling, post-translational protein modifications and NMR spectroscopy. *J Biomol NMR* 54, 217-236, doi:10.1007/s10858-012-9674-x (2012).
- 73 Lu, S. C. S-Adenosylmethionine. *The International Journal of Biochemistry & Cell Biology* 32, 391-395, doi:[https://doi.org/10.1016/S1357-2725\(99\)00139-9](https://doi.org/10.1016/S1357-2725(99)00139-9) (2000).
- 74 Hyun, K., Jeon, J., Park, K. & Kim, J. Writing, erasing and reading histone lysine methylations. *Exp Mol Med* 49, e324, doi:10.1038/emm.2017.11 (2017).
- 75 Murn, J. & Shi, Y. The winding path of protein methylation research: milestones and new frontiers. *Nature Reviews Molecular Cell Biology* 18, 517-527, doi:10.1038/nrm.2017.35 (2017).
- 76 Chen, Y. & Zhu, W. G. Biological function and regulation of histone and non-histone lysine methylation in response to DNA damage. *Acta Biochim Biophys Sin (Shanghai)* 48, 603-616, doi:10.1093/abbs/gmw050 (2016).
- 77 Petrossian, T. C. & Clarke, S. G. Uncovering the Human Methyltransferasome. *Molecular & Cellular Proteomics* 10, M110.000976, doi:10.1074/mcp.M110.000976 (2011).
- 78 Hamamoto, R., Saloura, V. & Nakamura, Y. Critical roles of non-histone protein lysine methylation in human tumorigenesis. *Nature reviews. Cancer* 15, 110-124, doi:10.1038/nrc3884 (2015).
- 79 Blanc, R. S. & Richard, S. Arginine Methylation: The Coming of Age. *Mol Cell* 65, 8-24, doi:10.1016/j.molcel.2016.11.003 (2017).
- 80 Bedford, M. T. & Clarke, S. G. Protein arginine methylation in mammals: who, what, and why. *Mol Cell* 33, 1-13, doi:10.1016/j.molcel.2008.12.013 (2009).
- 81 Caceres, T. B. *et al.* Phe71 in Type III Trypanosomal Protein Arginine Methyltransferase 7 (TbPRMT7) Restricts the Enzyme to Monomethylation. *Biochemistry* 57, 1349-1359, doi:10.1021/acs.biochem.7b01265 (2018).
- 82 Shi, H., Liu, L. & Wang, Z. Improving the efficacy and safety of engineered T cell therapy for cancer. *Cancer Lett* 328, 191-197, doi:10.1016/j.canlet.2012.09.015 (2013).
- 83 Gayatri, S. & Bedford, M. T. Readers of histone methylarginine marks. *Biochim Biophys Acta* 1839, 702-710, doi:10.1016/j.bbagr.2014.02.015 (2014).
- 84 Thandapani, P., O'Connor, T. R., Bailey, T. L. & Richard, S. Defining the RGG/RG motif. *Mol Cell* 50, 613-623, doi:10.1016/j.molcel.2013.05.021 (2013).
- 85 Yang, Y. & Bedford, M. T. Protein arginine methyltransferases and cancer. *Nature reviews. Cancer* 13, 37-50, doi:10.1038/nrc3409 (2013).
- 86 Feng, Y. *et al.* Mammalian protein arginine methyltransferase 7 (PRMT7) specifically targets RXR sites in lysine- and arginine-rich regions. *J Biol Chem* 288, 37010-37025, doi:10.1074/jbc.M113.525345 (2013).
- 87 Fuhrmann, J., Clancy, K. W. & Thompson, P. R. Chemical biology of protein arginine modifications in epigenetic regulation. *Chem Rev* 115, 5413-5461, doi:10.1021/acs.chemrev.5b00003 (2015).
- 88 Reintjes, A. *et al.* Asymmetric arginine dimethylation of RelA provides a repressive mark to modulate TNF α /NF- κ B response. *Proceedings of the National Academy of Sciences* 113, 4326-4331, doi:10.1073/pnas.1522372113 (2016).

- 89 Bao, J. *et al.* The arginine methyltransferase CARM1 represses p300•ACT•CREM τ activity and is required for spermiogenesis. *Nucleic Acids Research* 46, 4327-4343, doi:10.1093/nar/gky240 (2018).
- 90 Biggar, K. K. & Li, S. S. Non-histone protein methylation as a regulator of cellular signalling and function. *Nat Rev Mol Cell Biol* 16, 5-17, doi:10.1038/nrm3915 (2015).
- 91 Tsukada, Y. *et al.* Histone demethylation by a family of JmjC domain-containing proteins. *Nature* 439, 811-816, doi:10.1038/nature04433 (2006).
- 92 Smyth, M. J., Ngiow, S. F., Ribas, A. & Teng, M. W. Combination cancer immunotherapies tailored to the tumour microenvironment. *Nat Rev Clin Oncol* 13, 143-158, doi:10.1038/nrclinonc.2015.209 (2016).
- 93 Walport, L. J. *et al.* Arginine demethylation is catalysed by a subset of JmjC histone lysine demethylases. *Nat Commun* 7, 11974, doi:10.1038/ncomms11974 (2016).
- 94 Comb, D. G., Sarkar, N. & Pinzino, C. J. The Methylation of Lysine Residues in Protein. *Journal of Biological Chemistry* 241, 1857-1862 (1966).
- 95 Chuikov, S. *et al.* Regulation of p53 activity through lysine methylation. *Nature* 432, 353-360, doi:http://www.nature.com/nature/journal/v432/n7015/suppinfo/nature03117_S1.html (2004).
- 96 Chi, P., Allis, C. D. & Wang, G. G. Covalent histone modifications--miswritten, misinterpreted and mis-erased in human cancers. *Nature reviews. Cancer* 10, 457-469, doi:10.1038/nrc2876 (2010).
- 97 Tschiersch, B. *et al.* The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. *The EMBO journal* 13, 3822-3831 (1994).
- 98 Dillon, S. C., Zhang, X., Trievel, R. C. & Cheng, X. The SET-domain protein superfamily: protein lysine methyltransferases. *Genome biology* 6, 227-227, doi:10.1186/gb-2005-6-8-227 (2005).
- 99 Spellmon, N., Holcomb, J., Trescott, L., Sirinupong, N. & Yang, Z. Structure and function of SET and MYND domain-containing proteins. *Int J Mol Sci* 16, 1406-1428, doi:10.3390/ijms16011406 (2015).
- 100 Min, J., Zhang, X., Cheng, X., Grewal, S. I. & Xu, R. M. Structure of the SET domain histone lysine methyltransferase Clr4. *Nat Struct Biol* 9, 828-832, doi:10.1038/nsb860 (2002).
- 101 Carlson, S. M. & Gozani, O. Nonhistone Lysine Methylation in the Regulation of Cancer Pathways. *Cold Spring Harb Perspect Med* 6, doi:10.1101/cshperspect.a026435 (2016).
- 102 Herz, H. M., Garruss, A. & Shilatifard, A. SET for life: biochemical activities and biological functions of SET domain-containing proteins. *Trends Biochem Sci* 38, 621-639, doi:10.1016/j.tibs.2013.09.004 (2013).
- 103 Falnes, P. O., Jakobsson, M. E., Davydova, E., Ho, A. & Malecki, J. Protein lysine methylation by seven-beta-strand methyltransferases. *Biochem J* 473, 1995-2009, doi:10.1042/BCJ20160117 (2016).
- 104 Martin, J. L. & McMillan, F. M. SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. *Current Opinion in Structural Biology* 12, 783-793, doi:[https://doi.org/10.1016/S0959-440X\(02\)00391-3](https://doi.org/10.1016/S0959-440X(02)00391-3) (2002).

- 105 Schubert, H. L., Blumenthal, R. M. & Cheng, X. Many paths to methyltransfer: a chronicle of convergence. *Trends in Biochemical Sciences* 28, 329-335, doi:10.1016/s0968-0004(03)00090-2 (2003).
- 106 Dimitrova, E., Turberfield, A. H. & Klose, R. J. Histone demethylases in chromatin biology and beyond. *EMBO Rep* 16, 1620-1639, doi:10.15252/embr.201541113 (2015).
- 107 Majello, B., Gorini, F., Sacca, C. D. & Amente, S. Expanding the Role of the Histone Lysine-Specific Demethylase LSD1 in Cancer. *Cancers (Basel)* 11, doi:10.3390/cancers11030324 (2019).
- 108 Zheng, Y.-C. *et al.* Lysine demethylase 5B (KDM5B): A potential anti-cancer drug target. *European Journal of Medicinal Chemistry* 161, 131-140, doi:10.1016/j.ejmech.2018.10.040 (2019).
- 109 Roth, G. S., Casanova, A. G., Lemonnier, N. & Reynoird, N. Lysine methylation signaling in pancreatic cancer. *Curr Opin Oncol* 30, 30-37, doi:10.1097/CCO.0000000000000421 (2018).
- 110 Murray, K. The Occurrence of ϵ -N-Methyl Lysine in Histones. *Biochemistry* 3, 10-15, doi:10.1021/bi00889a003 (1964).
- 111 Allis, C. D. & Jenuwein, T. The molecular hallmarks of epigenetic control. *Nat Rev Genet* 17, 487-500, doi:10.1038/nrg.2016.59 (2016).
- 112 Plass, C. *et al.* Mutations in regulators of the epigenome and their connections to global chromatin patterns in cancer. *Nat Rev Genet* 14, 765-780, doi:10.1038/nrg3554 (2013).
- 113 Morera, L., Lubbert, M. & Jung, M. Targeting histone methyltransferases and demethylases in clinical trials for cancer therapy. *Clin Epigenetics* 8, 57, doi:10.1186/s13148-016-0223-4 (2016).
- 114 Zentner, G. E., Tesar, P. J. & Scacheri, P. C. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res* 21, 1273-1283, doi:10.1101/gr.122382.111 (2011).
- 115 Van Aller, G. S. *et al.* Smyd3 regulates cancer cell phenotypes and catalyzes histone H4 lysine 5 methylation. *Epigenetics* 7, 340-343, doi:10.4161/epi.19506 (2012).
- 116 He, R. & Kidder, B. L. H3K4 demethylase KDM5B regulates global dynamics of transcription elongation and alternative splicing in embryonic stem cells. *Nucleic Acids Res* 45, 6427-6441, doi:10.1093/nar/gkx251 (2017).
- 117 Lesch, B. J. & Page, D. C. Poised chromatin in the mammalian germ line. *Development* 141, 3619-3626, doi:10.1242/dev.113027 (2014).
- 118 Han, D. *et al.* Lysine methylation of transcription factors in cancer. *Cell Death Dis* 10, 290, doi:10.1038/s41419-019-1524-2 (2019).
- 119 Sze, C. C. *et al.* Histone H3K4 methylation-dependent and -independent functions of Set1A/COMPASS in embryonic stem cell self-renewal and differentiation. *Genes Dev* 31, 1732-1737, doi:10.1101/gad.303768.117 (2017).
- 120 Fang, L. *et al.* SET1A-Mediated Mono-Methylation at K342 Regulates YAP Activation by Blocking Its Nuclear Export and Promotes Tumorigenesis. *Cancer Cell* 34, 103-118 e109, doi:10.1016/j.ccell.2018.06.002 (2018).
- 121 Scheer, S. & Zaph, C. The Lysine Methyltransferase G9a in Immune Cell Differentiation and Function. *Front Immunol* 8, 429-429, doi:10.3389/fimmu.2017.00429 (2017).

- 122 Shinkai, Y. & Tachibana, M. H3K9 methyltransferase G9a and the related molecule GLP. *Genes Dev* 25, 781-788, doi:10.1101/gad.2027411 (2011).
- 123 Yoshida, K. *et al.* The transcription factor ATF7 mediates lipopolysaccharide-induced epigenetic changes in macrophages involved in innate immunological memory. *Nat Immunol* 16, 1034-1043, doi:10.1038/ni.3257 (2015).
- 124 Chang, Y. *et al.* MPP8 mediates the interactions between DNA methyltransferase Dnmt3a and H3K9 methyltransferase GLP/G9a. *Nat Commun* 2, 533, doi:10.1038/ncomms1549 (2011).
- 125 Ferry, L. *et al.* Methylation of DNA Ligase 1 by G9a/GLP Recruits UHRF1 to Replicating DNA and Regulates DNA Methylation. *Mol Cell* 67, 550-565 e555, doi:10.1016/j.molcel.2017.07.012 (2017).
- 126 Milite, C. *et al.* The emerging role of lysine methyltransferase SETD8 in human diseases. *Clin Epigenetics* 8, 102, doi:10.1186/s13148-016-0268-4 (2016).
- 127 Kwiatkowski, S. *et al.* SETD3 protein is the actin-specific histidine N-methyltransferase. *Elife* 7, doi:10.7554/eLife.37921 (2018).
- 128 Wilkinson, A. W. *et al.* SETD3 is an actin histidine methyltransferase that prevents primary dystocia. *Nature* 565, 372-376, doi:10.1038/s41586-018-0821-8 (2019).
- 129 He, A. *et al.* PRC2 directly methylates GATA4 and represses its transcriptional activity. *Genes & Development* 26, 37-42, doi:10.1101/gad.173930.111 (2012).
- 130 Dasgupta, M., Dermawan, J. K. T., Willard, B. & Stark, G. R. STAT3-driven transcription depends upon the dimethylation of K49 by EZH2. *Proceedings of the National Academy of Sciences* 112, 3985-3990, doi:10.1073/pnas.1503152112 (2015).
- 131 Yang, J. *et al.* Reversible methylation of promoter-bound STAT3 by histone-modifying enzymes. *Proceedings of the National Academy of Sciences* 107, 21499-21504, doi:10.1073/pnas.1016147107 (2010).
- 132 Yang, X.-D. *et al.* Negative regulation of NF- κ B action by Set9-mediated lysine methylation of the RelA subunit. *The EMBO Journal* 28, 1055-1066, doi:10.1038/emboj.2009.55 (2009).
- 133 Levy, D. *et al.* Lysine methylation of the NF-kappaB subunit RelA by SETD6 couples activity of the histone methyltransferase GLP at chromatin to tonic repression of NF-kappaB signaling. *Nat Immunol* 12, 29-36, doi:10.1038/ni.1968 (2011).
- 134 Vougiouklakis, T., Hamamoto, R., Nakamura, Y. & Saloura, V. The NSD family of protein methyltransferases in human cancer. *Epigenomics* 7, 863-874, doi:10.2217/epi.15.32 (2015).
- 135 Lu, T. *et al.* Regulation of NF-kappaB by NSD1/FBXL11-dependent reversible lysine methylation of p65. *Proc Natl Acad Sci U S A* 107, 46-51, doi:10.1073/pnas.0912493107 (2010).
- 136 Cho, H. S. *et al.* Enhanced HSP70 lysine methylation promotes proliferation of cancer cells through activation of Aurora kinase B. *Nat Commun* 3, 1072, doi:10.1038/ncomms2074 (2012).
- 137 Jakobsson, M. E. *et al.* Identification and characterization of a novel human methyltransferase modulating Hsp70 protein function through lysine methylation. *J Biol Chem* 288, 27752-27763, doi:10.1074/jbc.M113.483248 (2013).
- 138 Estève, P.-O. *et al.* Regulation of DNMT1 stability through SET7-mediated lysine methylation in mammalian cells. *Proceedings of the National Academy of Sciences* 106, 5076-5081, doi:10.1073/pnas.0810362106 (2009).

- 139 Shoab, M. *et al.* Histone H4K20 methylation mediated chromatin compaction threshold ensures genome integrity by limiting DNA replication licensing. *Nat Commun* 9, 3704, doi:10.1038/s41467-018-06066-8 (2018).
- 140 Takawa, M. *et al.* Histone lysine methyltransferase SETD8 promotes carcinogenesis by deregulating PCNA expression. *Cancer Res* 72, 3217-3227, doi:10.1158/0008-5472.CAN-11-3701 (2012).
- 141 Piao, L. *et al.* The Histone Methyltransferase SMYD2 Methylates PARP1 and Promotes Poly(ADP-ribosylation) Activity in Cancer Cells. *Neoplasia* 16, 257-264.e252, doi:10.1016/j.neo.2014.03.002 (2014).
- 142 Zhang, J. *et al.* PTEN Methylation by NSD2 Controls Cellular Sensitivity to DNA Damage. *Cancer Discov* 9, 1306-1323, doi:10.1158/2159-8290.CD-18-0083 (2019).
- 143 Kim, E. *et al.* Phosphorylation of EZH2 activates STAT3 signaling via STAT3 methylation and promotes tumorigenicity of glioblastoma stem-like cells. *Cancer Cell* 23, 839-852, doi:10.1016/j.ccr.2013.04.008 (2013).
- 144 Nie, L. *et al.* CDK2-mediated site-specific phosphorylation of EZH2 drives and maintains triple-negative breast cancer. *Nat Commun* 10, 5114, doi:10.1038/s41467-019-13105-5 (2019).
- 145 Gan, L. *et al.* Epigenetic regulation of cancer progression by EZH2: from biological insights to therapeutic potential. *Biomark Res* 6, 10, doi:10.1186/s40364-018-0122-2 (2018).
- 146 West, L. E. & Gozani, O. Regulation of p53 function by lysine methylation. *Epigenomics* 3, 361-369, doi:10.2217/epi.11.21 (2011).
- 147 West, L. E. *et al.* The MBT repeats of L3MBTL1 link SET8-mediated p53 methylation at lysine 382 to target gene repression. *J Biol Chem* 285, 37725-37732, doi:10.1074/jbc.M110.139527 (2010).
- 148 Dhami, G. K. *et al.* Dynamic methylation of Numb by Set8 regulates its binding to p53 and apoptosis. *Mol Cell* 50, 565-576, doi:10.1016/j.molcel.2013.04.028 (2013).
- 149 Huang, J. *et al.* Repression of p53 activity by Smyd2-mediated methylation. *Nature* 444, 629-632, doi:10.1038/nature05287 (2006).
- 150 Huang, J. *et al.* p53 is regulated by the lysine demethylase LSD1. *Nature* 449, 105-108, doi:10.1038/nature06092 (2007).
- 151 Kachirskaia, I. *et al.* Role for 53BP1 Tudor domain recognition of p53 dimethylated at lysine 382 in DNA damage signaling. *J Biol Chem* 283, 34660-34666, doi:10.1074/jbc.M806020200 (2008).
- 152 Kurash, J. K. *et al.* Methylation of p53 by Set7/9 mediates p53 acetylation and activity in vivo. *Mol Cell* 29, 392-400, doi:10.1016/j.molcel.2007.12.025 (2008).
- 153 Rada, M. *et al.* Human EHMT2/G9a activates p53 through methylation-independent mechanism. *Oncogene* 36, 922-932, doi:10.1038/onc.2016.258 (2017).
- 154 Huang, J. *et al.* G9a and Glp methylate lysine 373 in the tumor suppressor p53. *J Biol Chem* 285, 9636-9641, doi:10.1074/jbc.M109.062588 (2010).
- 155 Saddic, L. A. *et al.* Methylation of the retinoblastoma tumor suppressor by SMYD2. *J Biol Chem* 285, 37733-37740, doi:10.1074/jbc.M110.137612 (2010).
- 156 Munro, S., Khaire, N., Inche, A., Carr, S. & La Thangue, N. B. Lysine methylation regulates the pRb tumour suppressor protein. *Oncogene* 29, 2357-2367, doi:10.1038/onc.2009.511 (2010).

- 157 Cho, H. S. *et al.* Demethylation of RB regulator MYPT1 by histone demethylase LSD1 promotes cell cycle progression in cancer cells. *Cancer Res* 71, 655-660, doi:10.1158/0008-5472.CAN-10-2446 (2011).
- 158 Kim, Y. *et al.* Methylation-dependent regulation of HIF-1 α stability restricts retinal and tumour angiogenesis. *Nat Commun* 7, 10347, doi:10.1038/ncomms10347 (2016).
- 159 Bao, L. *et al.* Methylation of hypoxia-inducible factor (HIF)-1 α by G9a/GLP inhibits HIF-1 transcriptional activity and cell migration. *Nucleic Acids Res* 46, 6576-6591, doi:10.1093/nar/gky449 (2018).
- 160 Vershinin, Z., Feldman, M., Chen, A. & Levy, D. PAK4 Methylation by SETD6 Promotes the Activation of the Wnt/ β -Catenin Pathway. *J Biol Chem* 291, 6786-6795, doi:10.1074/jbc.M115.697292 (2016).
- 161 Deng, X. *et al.* Critical roles of SMYD2-mediated β -catenin methylation for nuclear translocation and activation of Wnt signaling. *Oncotarget* 8 (2017).
- 162 Borchardt, R. T. S-Adenosyl-L-methionine-dependent macromolecule methyltransferases: potential targets for the design of chemotherapeutic agents. *Journal of Medicinal Chemistry* 23, 347-357, doi:10.1021/jm00178a001 (1980).
- 163 Jones, P. A., Issa, J. P. & Baylin, S. Targeting the cancer epigenome for therapy. *Nat Rev Genet* 17, 630-641, doi:10.1038/nrg.2016.93 (2016).
- 164 Kim, K. H. & Roberts, C. W. Targeting EZH2 in cancer. *Nat Med* 22, 128-134, doi:10.1038/nm.4036 (2016).
- 165 Italiano, A. *et al.* Tazemetostat, an EZH2 inhibitor, in relapsed or refractory B-cell non-Hodgkin lymphoma and advanced solid tumours: a first-in-human, open-label, phase 1 study. *The Lancet Oncology* 19, 649-659, doi:10.1016/s1470-2045(18)30145-1 (2018).
- 166 Sarno, F., Nebbioso, A. & Altucci, L. DOT1L: a key target in normal chromatin remodelling and in mixed-lineage leukaemia treatment. *Epigenetics*, 1-15, doi:10.1080/15592294.2019.1699991 (2019).
- 167 Patnaik, S. & Anupriya. Drugs Targeting Epigenetic Modifications and Plausible Therapeutic Strategies Against Colorectal Cancer. *Front Pharmacol* 10, 588, doi:10.3389/fphar.2019.00588 (2019).
- 168 Copeland, R. A., Solomon, M. E. & Richon, V. M. Protein methyltransferases as a target class for drug discovery. *Nat Rev Drug Discov* 8, 724-732, doi:10.1038/nrd2974 (2009).
- 169 Hamamoto, R. & Nakamura, Y. Dysregulation of protein methyltransferases in human cancer: An emerging target class for anticancer therapy. *Cancer Sci* 107, 377-384, doi:10.1111/cas.12884 (2016).
- 170 Tracy, C. *et al.* The Smyd Family of Methyltransferases: Role in Cardiac and Skeletal Muscle Physiology and Pathology. *Curr Opin Physiol* 1, 140-152, doi:10.1016/j.cophys.2017.10.001 (2018).
- 171 Tan, X., Rotllant, J., Li, H., DeDeyne, P. & Du, S. J. SmyD1, a histone methyltransferase, is required for myofibril organization and muscle contraction in zebrafish embryos. *Proceedings of the National Academy of Sciences of the United States of America* 103, 2713, doi:10.1073/pnas.0509503103 (2006).

- 172 Leinhart, K. & Brown, M. SET/MYND Lysine Methyltransferases Regulate Gene Transcription and Protein Activity. *Genes (Basel)* 2, 210-218, doi:10.3390/genes2010210 (2011).
- 173 Brown, M. A., Sims, R. J., 3rd, Gottlieb, P. D. & Tucker, P. W. Identification and characterization of Smyd2: a split SET/MYND domain-containing histone H3 lysine 36-specific methyltransferase that interacts with the Sin3 histone deacetylase complex. *Mol Cancer* 5, 26, doi:10.1186/1476-4598-5-26 (2006).
- 174 Yi, X., Jiang, X. J. & Fang, Z. M. Histone methyltransferase SMYD2: ubiquitous regulator of disease. *Clin Epigenetics* 11, 112, doi:10.1186/s13148-019-0711-4 (2019).
- 175 Sese, B., Barrero, M. J., Fabregat, M. C., Sander, V. & Izpisua Belmonte, J. C. SMYD2 is induced during cell differentiation and participates in early development. *Int J Dev Biol* 57, 357-364, doi:10.1387/ijdb.130051ji (2013).
- 176 Foreman, K. W. *et al.* Structural and functional profiling of the human histone methyltransferase SMYD3. *PLoS One* 6, e22290, doi:10.1371/journal.pone.0022290 (2011).
- 177 Hamamoto, R. *et al.* SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nat Cell Biol* 6, 731-740, doi:10.1038/ncb1151 (2004).
- 178 Imhof, A. *et al.* Smyd3 Is Required for the Development of Cardiac and Skeletal Muscle in Zebrafish. *PLoS ONE* 6, e23491, doi:10.1371/journal.pone.0023491 (2011).
- 179 Suzuki, S. *et al.* Histone methyltransferase Smyd3 regulates early embryonic lineage commitment in mice. *Reproduction* 150, 21-30, doi:10.1530/REP-15-0019 (2015).
- 180 Bottino, C., Peserico, A., Simone, C. & Caretti, G. SMYD3: An Oncogenic Driver Targeting Epigenetic Regulation and Signaling Pathways. *Cancers* 12, 142, doi:10.3390/cancers12010142 (2020).
- 181 Mazur, P. K. *et al.* SMYD3 links lysine methylation of MAP3K2 to Ras-driven cancer. *Nature* 510, 283-287, doi:10.1038/nature13320 (2014).
- 182 Sarris, M. E., Moulos, P., Haroniti, A., Giakountis, A. & Talianidis, I. Smyd3 Is a Transcriptional Potentiator of Multiple Cancer-Promoting Genes and Required for Liver and Colon Cancer Development. *Cancer Cell* 29, 354-366, doi:10.1016/j.ccell.2016.01.013 (2016).
- 183 Hu, L., Zhu, Y. T., Qi, C. & Zhu, Y. J. Identification of Smyd4 as a potential tumor suppressor gene involved in breast cancer development. *Cancer Res* 69, 4067-4072, doi:10.1158/0008-5472.CAN-08-4097 (2009).
- 184 Kidder, B. L. *et al.* SMYD5 Controls Heterochromatin and Chromosome Integrity during Embryonic Stem Cell Differentiation. *Cancer Res* 77, 6729-6745, doi:10.1158/0008-5472.CAN-17-0828 (2017).
- 185 Komatsu, S. *et al.* Overexpression of SMYD2 relates to tumor cell proliferation and malignant outcome of esophageal squamous cell carcinoma. *Carcinogenesis* 30, 1139-1146, doi:10.1093/carcin/bgp116 (2009).
- 186 Ohtomo-Oda, R. *et al.* SMYD2 overexpression is associated with tumor cell proliferation and a worse outcome in human papillomavirus-unrelated nonmultiple head and neck carcinomas. *Hum Pathol* 49, 145-155, doi:10.1016/j.humpath.2015.08.025 (2016).

- 187 Li, L. X. *et al.* Lysine methyltransferase SMYD2 promotes triple negative breast cancer progression. *Cell Death Dis* 9, 326, doi:10.1038/s41419-018-0347-x (2018).
- 188 Komatsu, S. *et al.* Overexpression of SMYD2 contributes to malignant outcome in gastric cancer. *Br J Cancer* 112, 357-364, doi:10.1038/bjc.2014.543 (2015).
- 189 Zuo, S. R. *et al.* Positive Expression of SMYD2 is Associated with Poor Prognosis in Patients with Primary Hepatocellular Carcinoma. *J Cancer* 9, 321-330, doi:10.7150/jca.22218 (2018).
- 190 Sakamoto, L. H., Andrade, R. V., Felipe, M. S., Motoyama, A. B. & Pittella Silva, F. SMYD2 is highly expressed in pediatric acute lymphoblastic leukemia and constitutes a bad prognostic factor. *Leuk Res* 38, 496-502, doi:10.1016/j.leukres.2014.01.013 (2014).
- 191 Xu, W., Chen, F., Fei, X., Yang, X. & Lu, X. Overexpression of SET and MYND Domain-Containing Protein 2 (SMYD2) Is Associated with Tumor Progression and Poor Prognosis in Patients with Papillary Thyroid Carcinoma. *Med Sci Monit* 24, 7357-7365, doi:10.12659/MSM.910168 (2018).
- 192 Reynoird, N. *et al.* Coordination of stress signals by the lysine methyltransferase SMYD2 promotes pancreatic cancer. *Genes Dev* 30, 772-785, doi:10.1101/gad.275529.115 (2016).
- 193 Cho, H. S. *et al.* RB1 methylation by SMYD2 enhances cell cycle progression through an increase of RB1 phosphorylation. *Neoplasia* 14, 476-486, doi:10.1593/neo.12656 (2012).
- 194 Du, S. J., Tan, X. & Zhang, J. SMYD proteins: key regulators in skeletal and cardiac muscle development and function. *Anat Rec (Hoboken)* 297, 1650-1662, doi:10.1002/ar.22972 (2014).
- 195 Zhang, X. *et al.* Regulation of estrogen receptor alpha by histone methyltransferase SMYD2-mediated protein methylation. *Proc Natl Acad Sci U S A* 110, 17284-17289, doi:10.1073/pnas.1307959110 (2013).
- 196 Nakakido, M. *et al.* Dysregulation of AKT Pathway by SMYD2-Mediated Lysine Methylation on PTEN. *Neoplasia* 17, 367-373, doi:10.1016/j.neo.2015.03.002 (2015).
- 197 Liu, Y. *et al.* Overexpression of SMYD3 was associated with increased STAT3 activation in gastric cancer. *Med Oncol* 32, 404, doi:10.1007/s12032-014-0404-y (2015).
- 198 Wysk, M., Yang, D. D., Lu, H.-T., Flavell, R. A. & Davis, R. J. Requirement of mitogen-activated protein kinase kinase 3 (MKK3) for tumor necrosis factor-induced cytokine expression. *Proceedings of the National Academy of Sciences* 96, 3763-3768, doi:10.1073/pnas.96.7.3763 (1999).
- 199 Ludwig, S. *et al.* 3pK, a novel mitogen-activated protein (MAP) kinase-activated protein kinase, is targeted by three MAP kinase pathways. *Molecular and Cellular Biology* 16, 6687-6697, doi:10.1128/mcb.16.12.6687 (1996).
- 200 McLaughlin, M. M. *et al.* Identification of Mitogen-activated Protein (MAP) Kinase-activated Protein Kinase-3, a Novel Substrate of CSBP p38 MAP Kinase. *Journal of Biological Chemistry* 271, 8488-8492, doi:10.1074/jbc.271.14.8488 (1996).
- 201 Li, L. X. *et al.* Lysine methyltransferase SMYD2 promotes cyst growth in autosomal dominant polycystic kidney disease. *J Clin Invest* 127, 2751-2764, doi:10.1172/JCI90921 (2017).
- 202 Gao, S. *et al.* The lysine methyltransferase SMYD2 methylates the kinase domain of type II receptor BMPR2 and stimulates bone morphogenetic protein signaling. *Journal of Biological Chemistry* 292, 12702-12712, doi:10.1074/jbc.M117.776278 (2017).

- 203 Wang, R. *et al.* Effects of SMYD2-mediated EML4-ALK methylation on the signaling pathway and growth in non-small-cell lung cancer cells. *Cancer Sci* 108, 1203-1209, doi:10.1111/cas.13245 (2017).
- 204 Olsen, J. B. *et al.* Quantitative Profiling of the Activity of Protein Lysine Methyltransferase SMYD2 Using SILAC-Based Proteomics. *Molecular & Cellular Proteomics* 15, 892-905, doi:10.1074/mcp.M115.053280 (2016).
- 205 Ferguson, Andrew D. *et al.* Structural Basis of Substrate Methylation and Inhibition of SMYD2. *Structure* 19, 1262-1273, doi:10.1016/j.str.2011.06.011 (2011).
- 206 Nguyen, H. *et al.* LLY-507, a Cell-active, Potent, and Selective Inhibitor of Protein-lysine Methyltransferase SMYD2. *J Biol Chem* 290, 13641-13653, doi:10.1074/jbc.M114.626861 (2015).
- 207 Eggert, E. *et al.* Discovery and Characterization of a Highly Potent and Selective Aminopyrazoline-Based in Vivo Probe (BAY-598) for the Protein Lysine Methyltransferase SMYD2. *Journal of Medicinal Chemistry* 59, 4578-4600, doi:10.1021/acs.jmedchem.5b01890 (2016).
- 208 Zhu, Y. *et al.* SMYD3 stimulates EZR and LOXL2 transcription to enhance proliferation, migration, and invasion in esophageal squamous cell carcinoma. *Human Pathology* 52, 153-163, doi:10.1016/j.humpath.2016.01.012 (2016).
- 209 Liu, Y. *et al.* Overexpression of SMYD3 and matrix metalloproteinase-9 are associated with poor prognosis of patients with gastric cancer. *Tumour Biol* 36, 4377-4386, doi:10.1007/s13277-015-3077-z (2015).
- 210 Giakountis, A., Moulos, P., Sarris, M. E., Hatzis, P. & Talianidis, I. Smyd3-associated regulatory pathways in cancer. *Semin Cancer Biol* 42, 70-80, doi:10.1016/j.semcancer.2016.08.008 (2017).
- 211 Li, B. *et al.* SMYD3 promoter hypomethylation is associated with the risk of colorectal cancer. *Future Oncology* 14, 1825-1834, doi:10.2217/fon-2017-0682 (2018).
- 212 Binh, M. T. *et al.* Upregulation of SMYD3 and SMYD3 VNTR 3/3 polymorphism increase the risk of hepatocellular carcinoma. *Sci Rep* 10, 2797, doi:10.1038/s41598-020-59667-z (2020).
- 213 Oliveira-Santos, W. *et al.* Residual expression of SMYD2 and SMYD3 is associated with the acquisition of complex karyotype in chronic lymphocytic leukemia. *Tumour Biol* 37, 9473-9481, doi:10.1007/s13277-016-4846-z (2016).
- 214 Song, J. *et al.* Expression patterns and the prognostic value of the SMYD family members in human breast carcinoma using integrative bioinformatics analysis. *Oncology Letters*, doi:10.3892/ol.2019.10054 (2019).
- 215 Vieira, F. Q. *et al.* SMYD3 contributes to a more aggressive phenotype of prostate cancer and targets Cyclin D2 through H4K20me3. *Oncotarget* 6 (2015).
- 216 Guo, N. *et al.* Hepatitis C virus core upregulates the methylation status of the RASSF1A promoter through regulation of SMYD3 in hilar cholangiocarcinoma cells. *Acta Biochim Biophys Sin (Shanghai)* 43, 354-361, doi:10.1093/abbs/gmr021 (2011).
- 217 Zeng, B. *et al.* Epigenetic regulation of miR-124 by hepatitis C virus core protein promotes migration and invasion of intrahepatic cholangiocarcinoma cells by targeting SMYD3. *FEBS Lett* 586, 3271-3278, doi:10.1016/j.febslet.2012.06.049 (2012).

- 218 Hamamoto, R. *et al.* Enhanced SMYD3 expression is essential for the growth of breast cancer cells. *Cancer Sci* 97, 113-118, doi:10.1111/j.1349-7006.2006.00146.x (2006).
- 219 Kunizaki, M. *et al.* The lysine 831 of vascular endothelial growth factor receptor 1 is a novel target of methylation by SMYD3. *Cancer Res* 67, 10759-10765, doi:10.1158/0008-5472.CAN-07-1132 (2007).
- 220 Laplante, M. & Sabatini, D. M. mTOR signaling in growth control and disease. *Cell* 149, 274-293, doi:10.1016/j.cell.2012.03.017 (2012).
- 221 Yoshioka, Y. *et al.* SMYD3-mediated lysine methylation in the PH domain is critical for activation of AKT1. *Oncotarget* 7 (2016).
- 222 Yoshioka, Y. *et al.* Protein lysine methyltransferase SMYD3 is involved in tumorigenesis through regulation of HER2 homodimerization. *Cancer Med* 6, 1665-1672, doi:10.1002/cam4.1099 (2017).
- 223 Peserico, A. *et al.* A SMYD3 Small-Molecule Inhibitor Impairing Cancer Cell Growth. *Journal of Cellular Physiology* 230, 2447-2460, doi:10.1002/jcp.24975 (2015).
- 224 Mitchell, L. H. *et al.* Novel Oxindole Sulfonamides and Sulfamides: EPZ031686, the First Orally Bioavailable Small Molecule SMYD3 Inhibitor. *ACS Med Chem Lett* 7, 134-138, doi:10.1021/acsmedchemlett.5b00272 (2016).
- 225 Luo, X. G., Zou, J. N., Wang, S. Z., Zhang, T. C. & Xi, T. Novobiocin decreases SMYD3 expression and inhibits the migration of MDA-MB-231 human breast cancer cells. *IUBMB Life* 62, 194-199, doi:10.1002/iub.288 (2010).
- 226 Gaedcke, J. *et al.* Mutated KRAS results in overexpression of DUSP4, a MAP-kinase phosphatase, and SMYD3, a histone methyltransferase, in rectal carcinomas. *Genes Chromosomes Cancer* 49, 1024-1034, doi:10.1002/gcc.20811 (2010).

**PARTIE I : IDENTIFICATION DE
NOUVELLES SIGNALISATIONS DE SMYD2
ET SMYD3 DANS LE CANCER**

I. RESUME DE LA PARTIE I

Introduction : La méthylation des protéines est un mécanisme essentiel de régulation des fonctions cellulaires, impliqué dans de nombreux processus physiologiques et pathologiques. Les lysines méthyltransférases SMYD2 et SMYD3 sont surexprimées dans de nombreux cancers dans lesquels elles sont classiquement associées à une augmentation de la prolifération tumorale et à un mauvais pronostic.

Méthode : Ce chapitre porte sur une étude visant à identifier des nouveaux substrats prometteurs de SMYD2 et SMYD3 grâce à deux techniques complémentaires de protéomique à haut débit : le ProtoArray® et le SILAC-3xMBT *pulldown*. Les candidats sélectionnés ont ensuite été produits dans des bactéries et validés comme substrats en utilisant des réactions de méthylation *in vitro*. L'identification du site a ensuite été réalisée par des analyses de spectrométrie de masse et des tests de méthylation *in vitro* avec des mutants ponctuels K>R et K>A des différents sites suspectés. Enfin, grâce à la production d'anticorps méthyl-spécifiques des lysines méthylées de chaque substrat d'intérêt, les événements de méthylation ont été confirmés en modèle d'expression ectopique dans des cellules HEK 293T puis en lignées cellulaires de cancer du pancréas MIA-PaCa2 sauvage, *Smyd2*^{-/-} et *Smyd3*^{-/-} pour validation de la méthylation endogène des différents substrats.

Résultats : Nous avons démontré la méthylation *in vitro* de 4 nouveaux substrats de SMYD2 (RTF1, BCAR3, RRP44 et SARG), ainsi que 3 nouveaux substrats de SMYD3 (RNF113A, HSP90 et RIOK1). Nous avons caractérisé le site et le niveau de méthylation des 3 substrats suivants : RTF1 K587me1, BCAR3 K334me1 et RNF113A K20me3. Enfin, nous avons confirmé en cellule ces nouvelles signalisations de méthylation, en confirmant les événements de méthylation RTF1, BCAR3 et RNF113A dans les cellules 293T et les

méthylations endogènes BCAR3 et RNF113A dans la lignée cellulaire MIA PaCa-2 de l'adénocarcinome canalaire pancréatique.

Conclusion : Nous avons démontré ici la validité de nos méthodes pour l'identification de nouveaux substrats et confirmé in vitro et dans les cellules que SMYD2 méthyle RTF1 K587me1 et BCAR3 K334me1, et SMYD3 méthyle RNF113A K20me3, initiant de nouvelles études pour mieux caractériser le rôle de SMYD2 et SMYD3 dans le cancer.

ARTICLE

Identification of new SMYD2 and SMYD3 signalings in cancer

Roth S. Gaël^{1,2,3}, Casanova Alexandre¹, Couté Yohan⁴, Tardif Marianne⁴, Gozani Or⁷,
Hainaut Pierre¹, Reynoird Nicolas¹

1. Institute for Advanced Biosciences - INSERM U1209/CNRS UMR 5309/Université Grenoble-Alpes, France.
2. Clinique universitaire d'Hépatogastroentérologie et Oncologie digestive, CHU Grenoble-Alpes
3. Faculté de médecine, Université Grenoble-Alpes
4. Biosciences and Biotechnology Institute of Grenoble – INSERM U1038/CEA , France.
5. Department of Biology, Stanford University, USA.

Key words : SMYD2, SMYD3, lysine methylation signaling, cancer

ABSTRACT

Introduction: Methylation of proteins is essential in the regulation of cell functions in physiological and pathological settings. Lysine methyl-transferases SMYD2 and SMYD3 are overexpressed in many cancers and are associated with tumor proliferation and bad prognosis.

Method: In this study we identified new promising substrates of SMYD2 and SMYD3 thanks to two complementary techniques of high-throughput proteomics: ProtoArray® and SILAC-3xMBT pulldown.

Selected candidates were then produced in bacteria and validated as substrates using *in vitro* methylation reactions. Site identification was then realized by mass spectrometry analyses and *in vitro* methylation assays with point mutants K>R. Finally, newly identified methylation events were confirmed by ectopic expression in 293-T cells, and further confirmed as physiological by detection of endogenous methylations in MIA-PaCa2 pancreatic adenocarcinoma cells.

Results: We demonstrated the *in vitro* methylation of 4 new substrates of SMYD2 (RTF1, BCAR3, RRP44 and SARG), as well as 3 new substrates for SMYD3 (RNF113A, HSP90 and RIOK1). Moreover, we characterized the sites and the levels of methylation for 3 substrates as follow: RTF1 K587me1, BCAR3 K334me1 and RNF113A K20me3. Finally, we justified the relevance of these potential new lysine methylation signalings, by confirming RTF1, BCAR3 and RNF113A methylation events in the Human embryonic kidney 293T cells and BCAR3 and RNF113A endogenous methylations in the pancreatic ductal adenocarcinoma MIA PaCa-2 cell line either wild type, Smyd2^{-/-} or Smyd3^{-/-}.

Conclusion: We demonstrated here the validity of our methods for the identification of new substrates and confirmed *in vitro* and in cells that SMYD2 methylates RTF1 K587me1 and BCAR3 K334me1, and SMYD3 methylates RNF113A K20me3, initiating new studies to better characterized SMYD2 and SMYD3 role in cancer.

INTRODUCTION

Covalent post-translational modifications (PTMs) of proteins are essential for cell signaling and fine modulation of cell's adaptation to its environment. Recent high-throughput proteomic analyses of PTMs revealed a new layer of complexity in the regulation of cells' functions in physiological and pathological settings. Indeed, aberrant modulations of these PTMs constitute a strong source of pathophysiological mechanisms leading to a broad spectrum of diseases, in particular in carcinogenesis. While a great number of different PTMs exists among all amino acids on proteins, lysine is one of the most differentially modified residues. To this day, a large number of lysine PTMs is described and can be divided in 3 categories: acetylation ¹ and related PTMs such as crotonylation ², propionylation, butyrylation ³, succinylated ⁴, malonylation ⁵, formylation ⁶; ubiquitinylation or ubiquitinyl-like PTMs such as SUMOylation, ISGylation and NEDDylation ⁷; and other lysine modifications such as hydroxylation ⁸, glycation ⁹, and methylation ^{10,11}.

Methylation is a particularly essential PTM in cell physiology and S-adenosylmethionine (SAM), the methyl donor in methylation reactions, is one of the most abundant cofactors in cells. Protein methylation can occur either on arginine ¹² or lysine ¹⁰ residues, and this modification is also described on DNA ¹³ and RNA ¹⁴. Lysine methylation is a complex and finely tuned signaling, implicated in a large range of diseases, but it is still poorly characterized to this day. This PTM is a dynamic and reversible process involving two highly regulated sets of enzymes, called lysine methyltransferases (KMTs) and lysine demethylases (KDMs), which respectively catalyze the addition and removal of lysine methylation. A third set of proteins can transduce a specific response for each methylation event, by direct recognition of the modified lysine thanks to specific methyl-binding domains (MBDs) in a manner specified by the extent and context of the methylation event ¹⁵⁻¹⁷. Lysine residues can either be monomethylated, dimethylated or trimethylated, and each methylation state of a single lysine can potentially result in distinct biological outcomes ¹⁸.

Lysine methylation is known since early 1960's ¹⁹ and the first methylated substrates described were histones ²⁰. Rapidly, evidence of non-histone proteins grew with the

description of methylation of cytochromes in wheat germs and neurospora in the late 1960's²¹. Protein methylation is a very prolific field of research with an increasing number of KMTs and substrates described. The link between lysine methylation and cancer was then made with multiple examples of cancer-related signaling influenced by lysine methylation, such as EZH2, a KMT which plays a central role in cancer and whose pharmacological inhibition is actively studied^{22,23}, or the fact that the major tumor suppressor p53 is strongly modified on its C-ter extremity with diverse methylation marks regulating its activity²⁴. Other KMTs are studied with growing literature on their implication in cancer. The two KMTs SMYD2 and SMYD3 belong to the SET and MYND domain-containing proteins (SMYD) family, a class of KMTs containing an atypical catalytic SET domain which is split by a MYND binding domain²⁵. Both SMYD2 and SMYD3 were recently described as methylating essential signaling involved in carcinogenesis and seems to be overexpressed and associated with poor prognosis in many cancers.

SMYD2 seems to be ubiquitously expressed in tissues and is known to monomethylate its substrates. It was first classified as a histone methyl-transferase methylating H3K36 and H3K4, two marks that probably corresponds to *in vitro* artefacts observations^{26,27}. It was then described as monomethylating the tumor suppressor p53 on its lysine (K) 370, inhibiting its binding to the promoter of p21 target genes and leading to transcriptional repression with pro oncogenic consequences²⁸. To date, SMYD2 has been involved in esophageal squamous cell carcinoma (ESCC)²⁹, papilloma-induced oropharyngeal carcinoma³⁰, breast cancer³¹, gastric cancer³², lung and pancreatic ductal adenocarcinoma³³, hepatocellular carcinoma (HCC)³⁴, acute lymphoblastic leukemia³⁵ and papillary thyroid cancer³⁶. Other inhibiting marks on tumor suppressors catalyzed by SMYD2 have been observed, such as RB (retinoblastoma protein) K810me1 and K860me1 which promotes its binding to repressive transcriptional factor L3MBTL1 (Lethal(3)malignant brain tumor-like protein 1) leading to cell proliferation^{37,38}, PTEN (phosphatase and tensin homologue) K313me1 which leads to PI3K/AKT/mTOR pathway activation³⁹, and ER α (estrogen receptor alpha) K266me1 which downregulates ER α 's downstream signalings⁴⁰. Inversely, SMYD2 methylates oncogenic

proteins such as HSP90 (heat shock protein 90) K531me1 and K574me1, PARP1 (Poly(ADP-ribose) polymerase 1) K528me1 which increases resistance to chemotherapeutic agents ⁴¹, and MAPKAPK3 (MAPK Activated Protein Kinase 3) K355me1 ³³ which is involved in cell proliferation and tumor growth in inflammation-induced pancreatic ductal adenocarcinoma ³³.

On its side, SMYD3 is able to trimethylate its substrates in a processive manner, capable of using either non-, mono-, or di-methylated lysine as template. Fewer substrates have been described so far but its role in carcinogenesis seems essential too, as it appears to be overexpressed in many different cancers such as HCC ^{42,43}, cholangiocarcinoma ^{44,45}, colorectal cancer (CRC) ⁴⁶, ESCC ⁴⁷, gastric cancer ^{48,49}, lung and pancreatic ductal adenocarcinoma ⁵⁰, breast cancer ⁵¹, prostate cancer ⁵², and leukemias ⁵³. SMYD3 is one of the first KMTs shown as an epigenetic regulator promoting carcinogenesis, through histone H3 K4me3 ⁵⁴ and/or histone H4 K20me3 ⁵⁵. These signaling were recently invalidated, with clear demonstration of the non-canonical histone methylation site H4K5 as the single SMYD3 physiologic activity on histone ⁵¹. Few years ago, Mazur, Reynoird and colleagues demonstrated the first mechanistic evidence of oncogenic signaling for SMYD3, with MAP3K2 K260me3 promoting RAS signaling and tumor proliferation in RAS-dependent lung and pancreatic adenocarcinomas ⁵⁰. This mechanism was quite exciting at the time, as this mark does not induce the binding to a specific interactor, but rather the loss of the possible interaction of MAP3K2 with PP2a (protein phosphatase 2a), indirectly enhancing its activity. SMYD3 is also known to methylate VEGFR on its lysine K381, located on its catalytic site ⁵⁶, as well as HER2 (human epidermal growth factor receptor 2) on K175 ⁵⁷ and AKT1 (serine/threonine-protein kinase 1) on K14 ⁵⁸. However, downstream biological links with tumor progression of these signaling are still poorly characterized.

SMYD2 and SMYD3 are overexpressed in many various types of cancers, and only few substrates are clearly characterized with a good level of evidence, especially for SMYD3. Due to the complexity of lysine methylation detection, several substrates were described by *in vitro* analyses only, but could not be confirmed in cells. Also, in several cases the activity

of supposed substrates or signaling seemed to be associated with an increased expression of SMYD2 or SMYD3, but without clear understanding of the exact mechanism that linked the KMTs to these supposed substrates. As illustrated by the example of EZH2⁵⁹, KMTs are involved in many different pathways dependent of biological contexts, and we postulate that many major signaling and substrates of SMYD2 and SMYD3 are still unidentified. Finding new oncogenic pathways involving SMYD2 and SMYD3 seems essential to better characterize their role in cancer and to improve cancer treatment by developing new therapies targeting these signaling.

Therefore, the aim of this study was to discover and validate new methylation signaling of SMYD2 and SMYD3 in cancer cells, to allow a better understanding of their implication in cancer and to offer new perspectives in the development of anticancer targeted therapies.

RESULTS

Identification of SMYD2 and SMYD3 new substrates

We used the ProtoArray® platform from Invitrogen containing over 9,500 human proteins, as template to identify new potential substrates for SMYD2 and SMYD3 *in vitro*. Protein microchips were exposed either to the active or inactive recombinant KMT with radioactive H³-SAM as previously described, to obtain a radiolabelling of methylated proteins⁶⁰ (Figure 1A and 1B). We observed 251 hits for SMYD2 (Supplementary Table 1) and 22 hits for SMYD3 (Supplementary Table 2), with several common false-positives being auto-methylated methyltransferases. Among these potential substrates, several were already validated substrates of SMYD2 (HSP90, MAPKAPK3, p53) and SMYD3 (MAP3K2), demonstrating the efficiency of this technique. The large discrepancy between SMYD2 and SMYD3 size of candidate's pool could be explained in different ways. First, SMYD3 may be an extremely specific KMT while SMYD2 could be more related to poorly specific KMTs such as G9a or SETD7. Second, SMYD2's specificity could be dictated by cellular contexts, post-translational modifications or cofactors missing in this *in vitro* assay. This is suggested by the capacity of SMYD2 to methylate several residues on free histone H3 *in vitro*, while the enzyme is not able to methylate H3 within a nucleosome (unpublished data).

Therefore, we decided to use a second proteomic approach, the SILAC-3xMBT pulldown, to identify in a more physiological context the activity of SMYD2 and SMYD3. This second method consists in SILAC labeling of cell extracts in presence or absence of the recombinant KMT of interest (Stable Isotope Labeling of Amino Acids in Cell Culture), followed by an enrichment technique using the 3xMBT domain of L3MBTL1 protein. This domain has a pan-specific affinity for mono- and di-methylated lysines, with minimal sequence specificity and better capacity than commercially available pan-methyl antibodies (data unpublished). This allows for an efficient enrichment of all lysine-methylated proteins from a specific cell extract, and the distinction between the differently labelled population of proteins by MS/MS^{61,62}. Proteins potentially methylated by the KMTs were identified by liquid chromatography and

mass spectrometry analysis (LC-MS/MS) (Figure 1D). We identified with good confidence 55 hits for SMYD2 (Supplementary Table 3) and 31 hits for SMYD3 (Supplementary Table 4). Among them, two SMYD2 substrates were previously identified in literature (HSP90 and Rb) and one for SMYD3 (MAP3K2), validating the effectiveness of this technique. However, it is interesting to see that very few hits were common between the protoarray and the SILAC MBT techniques, strongly arguing for the importance of biological context in proper KMT's substrates identification. Thus, it is more likely that candidate-substrates from the SILAC MBT technique are better physiological substrates, especially for SMYD2, as the enzyme seems highly aspecific *in vitro*. Moreover, the reliability of SMYD3's hits by SILAC-3xMBT can be questioned as this enzyme seems to trimethylate in a processive way its substrates (i.e. doing mono- to di- to tri-methylation) while the 3xMBT domain used to enrich the pool of methylated proteins only binds to mono- and di-methylated lysines.

Therefore, each of the two proteomic techniques we used provided us with a global list of interesting potential new substrates for SMYD2 and SMYD3. By combining these approaches and by performing discriminatory analyses to help identifying specific and physiologic substrates, we established a final listing of 6 potential SMYD2's substrates, and 15 SMYD3's substrates we aimed to confirm and study further (Table 1). Among different criteria, clear function in cancer or identification in other proteomics studies were important parameters for this selection. For example, several SMYD2's substrates we discovered were also identified in a similar proteomic approach used by Olsen and colleagues. They performed a SILAC-based labeling, and compared methylated proteins enriched by immunoaffinity with mono-methyl-lysine (Kme1) antibody from cells expressing or depleted for SMYD2 (by either shRNA or inhibitory compound) ⁶³.

Confirmation of methylation events *in vitro*

On the basis of these lists of potential SMYD2 and SMYD3 substrates, we successfully cloned 14 candidates, and produced 11 recombinant proteins of good quality for further analyses as detailed in Table 2. Recombinant SMYD2 and SMYD3 wild type (wt) and

inactive mutants enzymes were also purified, and their enzymatic activity was validated on substrates already characterized such as MAPKAPK3, p53 and HSP90 for SMYD2^{28,33,64} and MAP3K2 for SMYD3⁵⁰ (data not shown).

We performed *in vitro* methylation tests with a tritium-labeled radioactive SAM cofactor, considered as the most reliable method for detecting methylation in a sensitive and unbiased manner when a methylation event is undescribed and no specific antibody is available. We confirmed the methylation of 4 new SMYD2 substrates (Figure 2A). Among them, RTF1 (Paf1/RNA polymerase II complex component) is a component of the PAF1 complex which plays multiple roles, notably in histone modifications and remodeling of chromatin with Histone 2B monoubiquitination. RTF1 also plays a role in transcription, by regulating the elongation phase and termination phase with 3' processing of polyadenylated and nonpolyadenylated Pol II transcripts⁶⁵. Our second validated substrate is BCAR3 (breast cancer anti-estrogen resistance 3), a chaperone protein which seems to be involved in tumor aggressiveness and resistance to estrogen inhibition in breast cancer patients⁶⁶. SMYD2 also methylates the exosome complex exonuclease RRP44 or DIS3 (Figure 2A), which is a nuclear exoribonucleosidase belonging to the exosome complex involved in RNA degradation⁶⁷. Finally, SMYD2 methylates SARG (specifically androgen-regulated gene protein) with a strong intensity of methylation compared to RTF1's methylation (Figure 2B). Regarding SMYD3, we confirmed 3 new substrates: HSP90, RNF113A (RING finger protein 113A) and RIOK1 (RIO kinase 1) (Figure 2C, 2D and 2E). HSP90 is a chaperone protein, involved in many cancers by stabilizing oncogenic proteins, and is already known to be methylated by SMYD2^{64,68}. RNF113A, is a RING finger E3 ubiquitin-protein ligase which non-sense mutation c.901 C> T, p.Q301 * appears to be responsible for trichotiodystrophy⁶⁹, a rare disease. It is also a E3-ligase, responsible for ubiquitination of proteins such as CXC4⁷⁰ or the ASCC3 complex (Activating signal cointegrator 1 complex subunit 3) in DNA alkylation repair⁷¹. This substrate was also particularly interesting as it was the only one strongly methylated by SMYD3 among various candidates tested (Figure 2C and 2D).

Finally, RIOK1 is an already known substrate of the KMT SETD7 and is implicated in tumor growth ⁷².

Identification of sites and level of methylation

We then decided to focus on the 5 most promising substrates regarding data from 3xMBT pulldown and ProtoArray® experiments, intensity of *in vitro* methylation assay, level of implication in cancer and potential downstream signaling pathways affected by these methylations: RTF1, BCAR3, RRP44, RNF113a and HSP90.

Lysine methylation is usually very specific and KMTs classically methylate their substrates on a unique site. Therefore, we aimed to identify the methylated lysine of each candidate using MS/MS analysis on purified methylated substrates from non-radioactive methylation assays with or without the active KMT (Figure 3A). In their study, Olsen *et al.* already identified several potential sites for SMYD2 substrates such as RTF1 K555me1 and K587me1, BCAR3 K5me1 and K334me1, RRP44 K331me1 and K552me1 ⁶³. Our analyses identified specifically RTF1 K587me1, BCAR3 K334me1, RRP44 K331me1 and K552me1. To verify if all identified lysines were truly methylated and to exclude false positive results, we performed point mutations by replacing these lysines by arginines (K>R). This change is sufficient to lose KMT specific methylation without inducing a significant structural change as both lysine and arginine are positively charged basic amino acids with comparable steric hindrance. For each lysine, we repeated radiolabeled methylation assays with wild type and K>R mutant proteins. Radioactive methylation assays revealed that RTF1 is specifically monomethylated *in vitro* by SMYD2 on its lysine K587, as methylation signal was completely abrogated using RTF1 K587K>R mutant (Figure 3B). Similarly, BCAR3 is only monomethylated on its lysine K334 (Figure 3C). Importantly, RTF1 K555 and BCAR3 K5 potential sites of methylation identified by Olsen *et al.* constitute false positive as their specific mutations did not alter methylation by SMYD2 (Figure 3B, 3C). Similarly, our analyses suggest that RRP44 K552 is not methylated by SMYD2, however mutation of lysine K331 does not totally shut down the methylation signal neither, as well as double mutant K331R and K552R (Figure 3D).

Consequently, we suspect a second unidentified lysine apart K331 to be methylated by SMYD2, within a peptide not covered by the trypsin digestion used for MSMS analysis.

Concerning SMYD3 substrates, HSP90 methylation site was not identified by mass spectrometry and we did not conduct further experiments. However, MS/MS analyses confirmed that RNF113A is trimethylated by SMYD3, but were not able to conclusively discriminate between RNF113A K20 and K21 as specific or dual potential methylation sites. Methylation reactions with the RNF113A K20R mutant showed no radioactive signal, in favor of complete loss of methylation, while the K21R mutant also showed a significant decrease in methylation signal (Figure 3E). We therefore performed K>A mutants (replacement of the target lysine with an alanine) which showed that K20A mutation totally shut down the methylation signal of RNF113A, when K21A mutation did not (Figure 3F). This confirms that K20 is the only lysine methylated by SMYD3, and suggests that K21R mutant might somehow impairs the proper binding of SMYD3 to RNF113A, indirectly inhibiting SMYD3's capacity to methylate RNF113A on K20.

Based on these identified methylated sites, we then produced biotinylated peptides comprising the target lysine in the monomethylated or unmethylated state for SMYD2 substrates such as RTF1 (Figure 4A), and BCAR3 (Figure 4B) as well as non-methylated, mono, di or tri-methylated state for RNF113A (Figure 4C). These peptides are essentially developed to validate the specificity of methyl-antibodies before immunodetection of methylation marks in cells as explained further. Nonetheless, they also allowed us to confirm monomethylation and trimethylation of substrates, respectively catalyzed by SMYD2 and SMYD3. Indeed, we repeated new *in vitro* radioactive methylation assays using these peptides and the recombinant KMTs and confirmed that SMYD2 is able to methylate RTF1 K587me₀ and BCAR3 K334me₀ peptide and cannot methylate RTF1 K587me₁ and BCAR3 K334me₁ anymore (data not shown). Concerning SMYD3, we performed radioactive, methylation assays using RNF113A K20me₀, me₁, me₂, me₃ incubated with recombinant SMYD3, as well as non-radioactive assays with immunodetection on dotblots using RNF113A K20me₃ specific antibody. We confirmed that SMYD3 is able to use unmodified,

mono and di-methylated lysines as templates to catalyze trimethylation in a processive manner (Figure 4C).

Confirmation of methylation events in cells

Confirmation of methylation events in cells, in overexpression systems such as transfection of 293T or more biologically relevant models such as endogenous methylation in cancer cell lines, is essential before running further explorations due to high risks of false positive results by *in vitro* assays. To efficiently detect these events, we first produced methyl specific antibodies by immunizing rabbits with peptides containing the methylated lysine of each substrate. Rabbits sera were then purified using the methylated peptides as bait, followed by a counter-purification with the corresponding unmethylated peptides to ensure for high antibody specificity. We validated the sensitivity and specificity of these antibodies by peptide dotblots for RTF1 K587me1 (Figure 4A), BCAR3 K334me1 (Figure 4B), RNF113A K20me3 (Figure 4C). Methyl specific antibodies were also tested on *in vitro* non-radioactive methylation assays, using full length recombinant RTF1 or BCAR3 either WT or K>R mutant with recombinant SMYD2 (Figure 4D), as well as full length recombinant RNF113A WT or mutant with recombinant SMYD3 (Figure 4F). These experiments confirmed the great sensitivity and specificity of our antibodies under these conditions.

However, such specificity and sensitivity need to be assessed using cell extracts. First, human cell line 293T derived from HEK 293 cells, were transfected with pCDNA plasmids in order to obtain the co-expression of SMYD2 or SMYD3 with its respective substrates either WT or K>R mutants. We detected RTF1 K587me1 in case of RTF1 WT + SMYD2 overexpression, and we no longer detected this mark in the absence of SMYD2 expression, or in case of RTF1 K587R mutant + SMYD2 expression (Figure 5A). We also observed BCAR3 K334me1 only in case of BCAR3 WT with SMYD2 expression, with the absence of this methylation mark in other conditions (Figure 5B). Finally, RNF113A K20me3 was only detected in case of RNF113A WT with SMYD3 expression (Figure 5C). These results suggest that our antibodies are specific for the targeted lysine, as mutant are not recognized

at all, and specific for the methylated status of these lysines as immunodetection only occurs in the condition of wild type substrate + active enzyme expression.

Thus, we identified the site and the level of methylation of RTF1, BCAR3 and RNF113A, and we confirmed that these events can occur in 293T cells. This experiment constitutes an advantageous system to bring out this type of interaction but does not confer a sufficient biological relevance. The next wanted to confirm these new signaling at the endogenous level in cancer cell lines that constitute a more biologically relevant model, to confirm their relevance in cancer and to justify the study of their implications in carcinogenesis.

Endogenous methylation confirmation

After screening 21 different cancer cell lines regarding SMYD2 and SMYD3 expression as well as their different substrates, we selected the MIA PaCa-2, a pancreatic cancer cell line in which the different substrates and both SMYD2 and SMYD3 appeared to be highly expressed compared to other human cell lines tested (Figure 6A). Based on both immunodetection on different cell compartment extracts obtained by cell fractionation, and immunofluorescence analyses, we observed a predominantly nuclear expression of RNF113A (Figure 6B and 6C), while BCAR3 is essentially cytoplasmic (Figure 6D). The results obtained with different RTF1 antibodies commercially available that we tested in our experiments were unfortunately inconclusive. Interestingly, SMYD2 and SMYD3 were present in both cytoplasm and nucleus of MIA PaCa-2 cells (Figure 6B and D).

We engineered cell lines for stable repression of SMYD2 or SMYD3, by cloning and transfecting selective shRNAs in either pSicoR (constitutive repression) or pLKO-Tet-On (inducible repression) plasmids (Figure 6). We observed BCAR3 K334 mono-methylation by SMYD2 in MIA PaCa-2 cells after immunoprecipitation with anti-BCAR3 antibody and immunodetection with BCAR3 K334me1 specific antibody. This mark was not detected in MIA PaCa-2 Smyd2^{-/-} cells (Figure 6E). As well, we confirmed the endogenous trimethylation of RNF113A K20 by SMYD3 in MIA PaCa-2 cells after immunopurification with anti-RNF113A antibody followed by immunodetection using the RNF113A K20me3 specific

antibody. As expected, SMYD3's repression correlates with the loss of RNF113A K20me3 mark in MIA PaCa-2 (Figure 6F). These endogenous methylation marks could not be observed without prior immunoprecipitation, probably because of the low concentration of these endogenous marks and the prevalence of heavily methylated proteins on non-purified cell extracts, bypassing the relatively good specificity of our methyl-antibodies.

Here, we were able to identify a list of potential new substrates of the oncogenic lysine methyltransferases SMYD2 and SMYD3 using two different and complementary unbiased proteomics approaches. We further demonstrated the quality of our identification by confirming *in vitro*, the methylation of 4 new substrates of SMYD2 (RTF1, BCAR3, RRP44 and SARG), as well as 3 new substrates for SMYD3 (RNF113A, HSP90 and RIOK1). Moreover, we characterized the methylation sites and the levels of methylation for RTF1, BCAR3 and RNF113A. Finally, we justified the relevance of these potential new lysine methylation signaling in physio/pathological contexts, by confirming RTF1, BCAR3 and RNF113A methylation events in the Human embryonic kidney 293T cells and BCAR3 and RNF113A endogenous methylations in the pancreatic ductal adenocarcinoma MIA PaCa-2 cell line.

DISCUSSION AND PERSPECTIVES

Methylation is a very frequent and ubiquitous event in cells, regulating numbers of signalings. In this study, we used 2 different high-throughput proteomic approaches without *a priori* to identify new substrates of SMYD2 and SMYD3, 2 KMTs largely involved in cancer.

First, we performed ProtoArray®, a technique that detect enzyme-substrate catalytic reaction independent of cellular specific contexts and variable substrates' availability in cells. This experiment constitutes an interesting first step to find new substrates without requiring a difficult enrichment step of methylated proteins, We obtained 22 hits for SMYD3 with ProtoArray®, 13 of them were selected as promising substrates and among the 11 we were able to test, only 3 of them were truly methylated *in vitro*: RNF113A, HSP90 and RIOK1. Regarding SMYD2, 251 hits were obtained, raising the question of the false to true positives

ratio and the efficiency of such methods for strongly active and aspecific KMT *in vitro*. This illustrated the fact that this method is interesting for SMYD3, but requires multiple verification steps as the risk of false positive is still important, and that it should be used with extreme caution for the identification of new substrates of SMYD2 or any other robust KMT.

We then used SILAC-3xMBT pulldown, comparing protein cell extracts exposed to the active or inactive form of the studied KMTs. This technique was particularly adapted in SMYD2 situation as it allows an enrichment of all mono- and di-methylated lysines with relatively mild influence of the amino acids sequence surrounding the methylated lysines ⁷³. We obtained 55 hits for SMYD2, and after selection of 5 substrates that seemed particularly promising, we were able to prove methylation of 4 of them *in vitro* and 2 in cells. This method is less adapted for SMYD3's trimethylated substrates screening, as among 31 hits, none of the 3 potential substrates we selected was confirmed by direct *in vitro* methylation assay.

SMYD2 is intensively studied with active research of new substrates and growing data in literature mostly based on high-throughput proteomic approaches. As previously mentioned, new potential SMYD2 substrates were recently revealed by large scale MS/MS analyses on ESCC cells extracts after SILAC coupled with immunoaffinity enrichment of monomethyl-lysine peptides by Olsen and colleagues. Interestingly, both their and our analyses lead to the discovery of significantly redundant substrates, strengthening the identification of these shared candidates by two unrelated studies ⁵³.

In this study, we detailed the different steps of our validation strategy for KMTs-substrates' reactions, that could be considered as a bottom-up approach, to confirm new methylation signalings. After the first steps, we selected the most promising candidates, first based on biological relevance if the substrates represent a major interest in cellular physiology or in carcinogenesis, but also for biochemical reasons such as the strength of the methylation signal, or pragmatical considerations such as the rapidity to validate first steps.

We were able to confirm methylation site *in vitro* as well as methylation events in cells for RNF113A, BCAR3 and RTF1. We did not succeed to localize the methylated lysine in RRP44 after mutagenesis of the 2 suspected lysines revealed by MS/MS analyses. This was

probably due to the presence of a third lysine whose site was not compatible with digestion reactions used for MS/MS analyses.

RTF1 is of major interest as it is involved in the regulation of RNA polymerase II as well as the modification of histones such as Histone 2 ubiquitination, and the regulation of chromatin. Its methylation could target one of its functions and provide a new mechanism characterizing the importance of SMYD2 in the regulation of chromatin ^{65,74,75}.

BCAR3 was initially described as a gene whose overexpression was associated with tamoxifen resistance and/or tumor aggressiveness in breast cancer patients ⁶⁶. Inversely, it is surprisingly described as a good prognostic marker in multiple myeloma ⁷⁶. Finally, BCAR3 protein seems involved in tumor invasiveness and cell motility through Rac1 activation ⁷⁷. It was recently confirmed that BCAR3 is promoting adhesion disassembly, tumor cell invasion and migration, through its interaction with p130^{cas} which also induces Rac1 activity ⁷⁸. This mechanism is of major interest in the understanding of carcinogenesis and particularly in breast cancer where SMYD2 is overexpressed and associated with bad prognosis ³¹.

As previously mentioned, RNF113A, is a RING finger E3 ubiquitin-protein ligase highly conserved between species finger with a domain, very conserved, a predominantly nuclear expression, which is overexpressed in many cancers including PDAC and colorectal cancer based on The Human Atlas data. Moreover, non-sense mutation c.901 C> T, p.Q301 of RNF113A has been linked to trichothiodystrophy ⁶⁹. Recently, RNF113A ubiquitin ligase activity was described as ubiquitinating CXCR4 on its lysine K331 in cells, leading to CXCR4 degradation, and inhibition of the CXCR4-CXCL12 signaling, involved in chemotactic recruitment of lymphocytes during inflammation ⁷⁰. Furthermore, RNF113A ubiquitin ligase activity is involved in alkylation damage repair through its interaction with the ASCC (activating signal cointegrator complex) pathway, and interestingly, hypersensitivity to alkylating agents is observed in patients with trichothiodystrophy ⁷¹. This candidate substrate was of particular interest as it was the only one that was confirmed as being methylated *in vitro* and in cells by SMYD3. Thus, it may represent an essential novel signaling explaining

the role of SMYD3 in RAS-independent cancers where SMYD3-MAP3K2 signaling cannot explain the link between overexpression of SMYD3 and increased tumor proliferation.

In this study, we demonstrated the efficiency of exhaustive high-throughput proteomics strategies to discover new substrates for the lysine methyltransferases SMYD2 and SMYD3, through 3xMBT pulldown and ProtoArray®. We further confirmed several selected hits among our list of potential candidates as true substrates of SMYD2 and SMYD3. We observed that SMYD2 monomethylates BCAR3 at K334 and RTF1 at K587, and that SMYD3 trimethylates RNF113 at K20 in MIA PaCa-2 cells. These new relevant signaling will require further explorations to characterize the biological consequences of these new methylation events and their potential involvement in SMYD2- and SMYD3-driven carcinogenesis.

MATERIAL AND METHOD

Protein Array

Assay was conducted as previously described⁶⁰. Briefly, a ProtoArray® (Invitrogen, version 5.0) was incubated overnight with GST alone or recombinant GST-tagged KMT (SMYD2 or SMYD3) and the tritium radiolabeled cofactor ³H-S-adenosyl-methionine in both conditions. Methylation was then revealed by autoradiography.

SILAC-3xMBT pulldown

SILAC consisted in cells cultured with a medium enriched in Arginines and Lysines harboring either light carbon and azote isotopes (C12 and N14; “K0” and “R0”) in one condition or heavy isotopes (C 13 and N15; “K6” and “R10”) in the other condition. This allows for an efficient quantification between the differently labelled population of proteins by MS/MS^{61,62}. SILAC-3xMBT pulldown was performed in replicate, using either light or heavy labelled total protein extracts from HeLa cells, incubated with either active KMT (aKMT) or inactive KMT (iKMT). In both conditions, S-adenosyl-methionine was added. Then, opposite reactions

were pooled together (light + aKMT / Heavy + iKMT or light + iKMT / Heavy + aKMT) and 3xMBT pulldown was performed for each pair of reactions.

Protein pull-downs were performed as previously described⁶¹. Briefly, 1 mg of total protein extract was prepared with heavy and light isotopic labeling and enriched by incubation with recombinant 3xMBT-GST domains coupled on GSH sepharose beads. After glutathione elution, enriched methylated proteins were purified by removing 3xMBT-GST by dialysis in high salt buffer and recapture on GSH-Sepharose beads. Purified methylated proteins were separated by SDS-PAGE and analyzed by in-gel tryptic digest followed by LC-MS/MS. Peptides were identified using MaxQuant version 1.2.2.5 (Cox and Mann, 2008), and candidate methylated peptides were verified by manual inspection.

Plasmid, cloning and shRNA construction

Bacterial expression plasmids were created using pGEX-6P1 vector. Transient mammalian expression plasmids were created using pCDNA3.1 HA and pCDNA3.1 Myc. The different inserts were amplified by PCR using either cDNA or specific clones from the human ORFeome library as template. Single-point mutations of RTF1, BCAR3, RRP44 and RNF113A were generated using the Q5 site directed mutagenesis kit protocol (New England Biolabs Inc), and clones were confirmed by DNA sequencing. SMYD2, SMYD3 inducible shRNA targeting UTR regions were cloned in pLKO-Tet-On vector carrying a puromycin resistance gene, as follow:

- Targeted Human SMYD2 UTR 3' sequence : GTCTGAATCTTGAACCTTA
- Targeted Human SMYD3 UTR 3' sequence : GAACGCAGTCAGAGGGAAA

Expression and purification of recombinant proteins

For expression of GST tagged recombinant proteins, transformed BL21 cells were induced at OD=0.6 with 0.1 mM IPTG overnight at 20 °C, and proteins were purified using Glutathione Sepharose® 4B beads (GE Healthcare) and either eluted in 10 mM reduced L-glutathione (Sigma) or cleaved from the GST tag using purified Precision enzyme.

Methylation assays

In vitro methylation assays were performed using 1 to 2 μg of recombinant proteins or peptides incubated with 1 μg of recombinant methyltransferases and 0.1 mM of S-adenosyl-methionine (SAM, Sigma) or 2 μCi 3H-SAM in buffer containing 50 mM Tris-HCl (pH 8.0), 10% glycerol, 20 mM KCl, 5 mM MgCl_2 , and 1 mM PMSF at 30 °C overnight. The reaction mixture was resolved by SDS-PAGE, followed by autoradiography, Coomassie stain (Pierce) or mass spectrometry analysis.

Cell culture, reagents and stable knockdown

293T, and PDAC cell lines SW1990, CAPAN1, HPAF2, CFPAC, MIA-PaCa2, Paca3, YPAC, T3M4 and PDAC8902, were grown in Dulbecco's modified Eagle's medium (GIBCO) supplemented with 10% fetal calf serum (FCS, GIBCO), 100 units ml^{-1} penicillin/streptomycin and glutamine. All cells were cultured at 37 °C in a humidified incubator with 5% CO_2 . For transient expression, cells were transfected using Lipofectamine transfection reagent and collected 24 to 36 h later. For stable knockdown, cells were transduced with lentiviral shRNA constructs using the packaging vectors pGagpol and p Δ 8.2, followed by 2 $\mu\text{g}/\text{mL}$ puromycin selection for one week. For rescue experiments, cells were transduced with retroviral pBabe and pMSCV constructs using packaging pGag and pVSVg vectors, followed by 50 $\mu\text{g}/\text{mL}$ hygromycin selection for one week.

Cell extracts, immunoprecipitation and immunoblot analysis

For total cell extracts, cells were lysed in RIPA buffer (10 mM Tris-HCl pH 8, 140 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1% Triton, 0.1% SDS, 1 mM PMSF, protease inhibitors (Roche) and a phosphatase inhibitor cocktail (Sigma-Aldrich) for 15 min. Cell fractionation was performed by collecting supernatant (cytoplasmic fraction) after 5 min of 1,300g centrifugation following a 10 min incubation in hypotonic buffer (10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM MgCl_2 , 0.34 M sucrose, 10% glycerol, 1 mM DTT, 0.05% Triton, and

protease inhibitors). The pellet was then incubated 15 min in LSDB250 buffer (glycerol 20%, MgCl₂ 3 mM, HEPES pH 7.9 50 mM, KCl 250 mM, DTT 0.5mM, PMSF 0.5 mM, NP40 0,1%, protease inhibitors), and centrifuge at 15,000g for 10 min. The supernatant was collected as soluble nuclear extract, and the pellet was further extracted in LSDB250 buffer with sonication (chromatin fraction). Protein concentration was determined by BCA assay (Pierce).

For immunoprecipitation, cells were lysed in either LSDB250 buffer for total cell extract, or hypotonic buffer for cytoplasmic extract, and same amount of protein extracts were incubated with specific antibody overnight at 4 °C. Extracts were then incubated with protein A Sepharose beads (GE Healthcare) for 2 h at 4 °C.

Proteins were resolved by SDS–PAGE, transferred to nitrocellulose membrane and analyzed by immunoblot. Antibodies used were as follows: SMYD2 (D14H7 CST), SMYD3 (home-made and D2Q4V CST), RTF1K587me1, BCAR3K331me1 and RNF113K20me3 (generated by Eurogenetec ®); beta-tubulin (05-661 Millipore), RTF1, BCAR3, RNF113a, HA, Flag antibodies.

Immunofluorescence

Cells were cultured on coverslips pre-coated with 10ug / ml of fibronectin and fixed with PFA 4% (Paraformaldéhyde, Sigma) for 10 minutes at 20°C. Permeabilization was done with PBS-Tween 0.1% -Triton 0.1% for 2 minutes at 20 ° C. Non-specific sites were blocked with 2.5% BSA for 1 hour, followed by incubation with the primary antibody of interest for 2 h at 20°C in the blocking solution, washes and final incubation with the secondary antibody (AlexaFluor 488 cross-absorbing, Thermo Fisher) and phalloidin (Texas Red, Thermo Fisher) for 1 h at 20°C. Mounting of coverslips on glass slides was performed using ProLong™ Gold Antifade Mountant (Invitrogen). Images were captured with a X63 magnification using the inverted confocal microscope Andromeda TILL-FEI (EMCCD iXon 897 camera ; aPlan-Apochromat 63x/1.46 Oil Korr M27, WD 0.1 mm objective) and analyzed using ImageJ software.

Peptide production

After identification of the methylation site, 15-aminoacids biotinylated peptides containing the methylated lysine were designed based on the following sequence: Biotin - Ahx - DQVCTFLF Kme3 KPGRKG - CONH₂ (example of RNF113A K20me₃). Peptides of RTF1K587me₀/me₁, BCAR3K331me₀/me₁ and RNF113A K20me₀/me₁/me₂/me₃ peptides were produced by Covalab ®.

Dotblot

Biotinylated peptides were either directly spotted on nitrocellulose membranes at different concentrations, or spotted after methylation assays with recombinant KMT and SAM. Membranes were then dried O/N, blocked in 5% BSA TBS-Tween 0.1% and then incubated with either streptavidin antibody (for loading control) or methyl specific antibodies (α -RTF1K587me₁, α -BCAR3K331me₁ and α -RNF113K20me₃).

Mass spectrometry analyses

Non-radioactive *in vitro* methylation assays extracts were analyzed by in-gel tryptic digestion followed by LC-MS/MS using an OrbiTrap VELOS pro. Results were analyzed by with MaxQuant version 1.2.2.5 software (Cox and Mann, 2008).

REFERENCES

- 1 Choudhary, C. *et al.* Lysine Acetylation Targets Protein Complexes and Co-Regulates Major Cellular Functions. *Science* 325, 834, doi:10.1126/science.1175371 (2009).
- 2 Tan, M. *et al.* Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 146, 1016-1028, doi:10.1016/j.cell.2011.08.008 (2011).
- 3 Chen, Y. *et al.* Lysine Propionylation and Butyrylation Are Novel Post-translational Modifications in Histones. *Molecular & Cellular Proteomics* 6, 812-819, doi:10.1074/mcp.M700021-MCP200 (2007).
- 4 Peng, C. *et al.* The First Identification of Lysine Malonylation Substrates and Its Regulatory Enzyme. *Molecular & Cellular Proteomics* 10, M111.012658, doi:10.1074/mcp.M111.012658 (2011).
- 5 Weinert, B. T. *et al.* Lysine succinylation is a frequently occurring modification in prokaryotes and eukaryotes and extensively overlaps with acetylation. *Cell Rep* 4, 842-851, doi:10.1016/j.celrep.2013.07.024 (2013).
- 6 Wisniewski, J. R., Zougman, A. & Mann, M. Nepsilon-formylation of lysine is a widespread post-translational modification of nuclear proteins occurring at residues involved in regulation of chromatin function. *Nucleic Acids Res* 36, 570-577, doi:10.1093/nar/gkm1057 (2008).
- 7 Hochstrasser, M. Origin and function of ubiquitin-like proteins. *Nature* 458, 422-429, doi:10.1038/nature07958 (2009).
- 8 van Slyke, D. D. & Sinex, F. M. THE COURSE OF HYDROXYLATION OF LYSINE TO FORM HYDROXYLYSINE IN COLLAGEN. *Journal of Biological Chemistry* 232, 797-806 (1958).
- 9 Johansen, M. B., Kiemer, L. & Brunak, S. Analysis and prediction of mammalian protein glycation. *Glycobiology* 16, 844-853, doi:10.1093/glycob/cwl009 (2006).
- 10 Lanouette, S., Mongeon, V., Figeys, D. & Couture, J. F. The functional diversity of protein lysine methylation. *Molecular Systems Biology* 10, 724-724, doi:10.1002/msb.134974 (2014).
- 11 Zhang, X., Huang, Y. & Shi, X. Emerging roles of lysine methylation on non-histone proteins. *Cell Mol Life Sci* 72, 4257-4272, doi:10.1007/s00018-015-2001-4 (2015).
- 12 Blanc, R. S. & Richard, S. Arginine Methylation: The Coming of Age. *Mol Cell* 65, 8-24, doi:10.1016/j.molcel.2016.11.003 (2017).
- 13 Schubeler, D. Function and information content of DNA methylation. *Nature* 517, 321-326, doi:10.1038/nature14192 (2015).
- 14 Mongan, N. P., Emes, R. D. & Archer, N. Detection and analysis of RNA methylation. *F1000Res* 8, F1000 Faculty Rev-1559, doi:10.12688/f1000research.17956.1 (2019).
- 15 Roth, G. S., Casanova, A. G., Lemonnier, N. & Reynoird, N. Lysine methylation signaling in pancreatic cancer. *Curr Opin Oncol* 30, 30-37, doi:10.1097/CCO.0000000000000421 (2018).
- 16 Dillon, S. C., Zhang, X., Trievel, R. C. & Cheng, X. The SET-domain protein superfamily: protein lysine methyltransferases. *Genome biology* 6, 227-227, doi:10.1186/gb-2005-6-8-227 (2005).

- 17 Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D. & Patel, D. J. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* 14, 1025-1040, doi:10.1038/nsmb1338 (2007).
- 18 Hyun, K., Jeon, J., Park, K. & Kim, J. Writing, erasing and reading histone lysine methylations. *Exp Mol Med* 49, e324, doi:10.1038/emm.2017.11 (2017).
- 19 Allfrey, V. G., Faulkner, R. & Mirsky, A. E. ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS. *Proceedings of the National Academy of Sciences* 51, 786-794, doi:10.1073/pnas.51.5.786 (1964).
- 20 Murray, K. The Occurrence of ϵ -N-Methyl Lysine in Histones. *Biochemistry* 3, 10-15, doi:10.1021/bi00889a003 (1964).
- 21 DeLange, R. J., Glazer, A. N. & Smith, E. L. Presence and Location of an Unusual Amino Acid, ϵ -N-Trimethyllysine, in Cytochrome c of Wheat Germ and Neurospora. *Journal of Biological Chemistry* 244, 1385-1388 (1969).
- 22 Gan, L. *et al.* Epigenetic regulation of cancer progression by EZH2: from biological insights to therapeutic potential. *Biomark Res* 6, 10, doi:10.1186/s40364-018-0122-2 (2018).
- 23 Kim, K. H. & Roberts, C. W. Targeting EZH2 in cancer. *Nat Med* 22, 128-134, doi:10.1038/nm.4036 (2016).
- 24 West, L. E. & Gozani, O. Regulation of p53 function by lysine methylation. *Epigenomics* 3, 361-369, doi:10.2217/epi.11.21 (2011).
- 25 Tracy, C. *et al.* The Smyd Family of Methyltransferases: Role in Cardiac and Skeletal Muscle Physiology and Pathology. *Curr Opin Physiol* 1, 140-152, doi:10.1016/j.cophys.2017.10.001 (2018).
- 26 Leinhart, K. & Brown, M. SET/MYND Lysine Methyltransferases Regulate Gene Transcription and Protein Activity. *Genes (Basel)* 2, 210-218, doi:10.3390/genes2010210 (2011).
- 27 Brown, M. A., Sims, R. J., 3rd, Gottlieb, P. D. & Tucker, P. W. Identification and characterization of Smyd2: a split SET/MYND domain-containing histone H3 lysine 36-specific methyltransferase that interacts with the Sin3 histone deacetylase complex. *Mol Cancer* 5, 26, doi:10.1186/1476-4598-5-26 (2006).
- 28 Huang, J. *et al.* Repression of p53 activity by Smyd2-mediated methylation. *Nature* 444, 629-632, doi:10.1038/nature05287 (2006).
- 29 Komatsu, S. *et al.* Overexpression of SMYD2 relates to tumor cell proliferation and malignant outcome of esophageal squamous cell carcinoma. *Carcinogenesis* 30, 1139-1146, doi:10.1093/carcin/bgp116 (2009).
- 30 Ohtomo-Oda, R. *et al.* SMYD2 overexpression is associated with tumor cell proliferation and a worse outcome in human papillomavirus-unrelated nonmultiple head and neck carcinomas. *Hum Pathol* 49, 145-155, doi:10.1016/j.humpath.2015.08.025 (2016).
- 31 Li, L. X. *et al.* Lysine methyltransferase SMYD2 promotes triple negative breast cancer progression. *Cell Death Dis* 9, 326, doi:10.1038/s41419-018-0347-x (2018).

- 32 Komatsu, S. *et al.* Overexpression of SMYD2 contributes to malignant outcome in gastric cancer. *Br J Cancer* 112, 357-364, doi:10.1038/bjc.2014.543 (2015).
- 33 Reynoird, N. *et al.* Coordination of stress signals by the lysine methyltransferase SMYD2 promotes pancreatic cancer. *Genes Dev* 30, 772-785, doi:10.1101/gad.275529.115 (2016).
- 34 Zuo, S. R. *et al.* Positive Expression of SMYD2 is Associated with Poor Prognosis in Patients with Primary Hepatocellular Carcinoma. *J Cancer* 9, 321-330, doi:10.7150/jca.22218 (2018).
- 35 Sakamoto, L. H., Andrade, R. V., Felipe, M. S., Motoyama, A. B. & Pittella Silva, F. SMYD2 is highly expressed in pediatric acute lymphoblastic leukemia and constitutes a bad prognostic factor. *Leuk Res* 38, 496-502, doi:10.1016/j.leukres.2014.01.013 (2014).
- 36 Xu, W., Chen, F., Fei, X., Yang, X. & Lu, X. Overexpression of SET and MYND Domain-Containing Protein 2 (SMYD2) Is Associated with Tumor Progression and Poor Prognosis in Patients with Papillary Thyroid Carcinoma. *Med Sci Monit* 24, 7357-7365, doi:10.12659/MSM.910168 (2018).
- 37 Cho, H. S. *et al.* RB1 methylation by SMYD2 enhances cell cycle progression through an increase of RB1 phosphorylation. *Neoplasia* 14, 476-486, doi:10.1593/neo.12656 (2012).
- 38 Saddic, L. A. *et al.* Methylation of the retinoblastoma tumor suppressor by SMYD2. *J Biol Chem* 285, 37733-37740, doi:10.1074/jbc.M110.137612 (2010).
- 39 Nakakido, M. *et al.* Dysregulation of AKT Pathway by SMYD2-Mediated Lysine Methylation on PTEN. *Neoplasia* 17, 367-373, doi:10.1016/j.neo.2015.03.002 (2015).
- 40 Zhang, X. *et al.* Regulation of estrogen receptor alpha by histone methyltransferase SMYD2-mediated protein methylation. *Proc Natl Acad Sci U S A* 110, 17284-17289, doi:10.1073/pnas.1307959110 (2013).
- 41 Piao, L. *et al.* The Histone Methyltransferase SMYD2 Methylates PARP1 and Promotes Poly(ADP-ribosyl)ation Activity in Cancer Cells. *Neoplasia* 16, 257-264.e252, doi:10.1016/j.neo.2014.03.002 (2014).
- 42 Binh, M. T. *et al.* Upregulation of SMYD3 and SMYD3 VNTR 3/3 polymorphism increase the risk of hepatocellular carcinoma. *Sci Rep* 10, 2797, doi:10.1038/s41598-020-59667-z (2020).
- 43 Sarris, M. E., Moulos, P., Haroniti, A., Giakountis, A. & Talianidis, I. Smyd3 Is a Transcriptional Potentiator of Multiple Cancer-Promoting Genes and Required for Liver and Colon Cancer Development. *Cancer Cell* 29, 354-366, doi:10.1016/j.ccell.2016.01.013 (2016).
- 44 Guo, N. *et al.* Hepatitis C virus core upregulates the methylation status of the RASSF1A promoter through regulation of SMYD3 in hilar cholangiocarcinoma cells. *Acta Biochim Biophys Sin (Shanghai)* 43, 354-361, doi:10.1093/abbs/gmr021 (2011).
- 45 Zeng, B. *et al.* Epigenetic regulation of miR-124 by hepatitis C virus core protein promotes migration and invasion of intrahepatic cholangiocarcinoma cells by targeting SMYD3. *FEBS Lett* 586, 3271-3278, doi:10.1016/j.febslet.2012.06.049 (2012).
- 46 Li, B. *et al.* SMYD3 promoter hypomethylation is associated with the risk of colorectal cancer. *Future Oncology* 14, 1825-1834, doi:10.2217/fo-2017-0682 (2018).

- 47 Zhu, Y. *et al.* SMYD3 stimulates EZR and LOXL2 transcription to enhance proliferation, migration, and invasion in esophageal squamous cell carcinoma. *Human Pathology* 52, 153-163, doi:10.1016/j.humpath.2016.01.012 (2016).
- 48 Liu, Y. *et al.* Overexpression of SMYD3 was associated with increased STAT3 activation in gastric cancer. *Med Oncol* 32, 404, doi:10.1007/s12032-014-0404-y (2015).
- 49 Liu, Y. *et al.* Overexpression of SMYD3 and matrix metalloproteinase-9 are associated with poor prognosis of patients with gastric cancer. *Tumour Biol* 36, 4377-4386, doi:10.1007/s13277-015-3077-z (2015).
- 50 Mazur, P. K. *et al.* SMYD3 links lysine methylation of MAP3K2 to Ras-driven cancer. *Nature* 510, 283-287, doi:10.1038/nature13320 (2014).
- 51 Van Aller, G. S. *et al.* Smyd3 regulates cancer cell phenotypes and catalyzes histone H4 lysine 5 methylation. *Epigenetics* 7, 340-343, doi:10.4161/epi.19506 (2012).
- 52 Vieira, F. Q. *et al.* SMYD3 contributes to a more aggressive phenotype of prostate cancer and targets Cyclin D2 through H4K20me3. *Oncotarget* 6 (2015).
- 53 Oliveira-Santos, W. *et al.* Residual expression of SMYD2 and SMYD3 is associated with the acquisition of complex karyotype in chronic lymphocytic leukemia. *Tumour Biol* 37, 9473-9481, doi:10.1007/s13277-016-4846-z (2016).
- 54 Hamamoto, R. *et al.* SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nat Cell Biol* 6, 731-740, doi:10.1038/ncb1151 (2004).
- 55 Foreman, K. W. *et al.* Structural and functional profiling of the human histone methyltransferase SMYD3. *PLoS One* 6, e22290, doi:10.1371/journal.pone.0022290 (2011).
- 56 Kunizaki, M. *et al.* The lysine 831 of vascular endothelial growth factor receptor 1 is a novel target of methylation by SMYD3. *Cancer Res* 67, 10759-10765, doi:10.1158/0008-5472.CAN-07-1132 (2007).
- 57 Yoshioka, Y. *et al.* Protein lysine methyltransferase SMYD3 is involved in tumorigenesis through regulation of HER2 homodimerization. *Cancer Med* 6, 1665-1672, doi:10.1002/cam4.1099 (2017).
- 58 <SMYD3 trimethylates AKT1-Oncotarget 2016.pdf>.
- 59 Yamaguchi, H. & Hung, M. C. Regulation and Role of EZH2 in Cancer. *Cancer Res Treat* 46, 209-222, doi:10.4143/crt.2014.46.3.209 (2014).
- 60 Levy, D. *et al.* A proteomic approach for the identification of novel lysine methyltransferase substrates. *Epigenetics & Chromatin* 4, 19, doi:10.1186/1756-8935-4-19 (2011).
- 61 Carlson, S. M., Moore, K. E., Green, E. M., Martín, G. M. & Gozani, O. Proteome-wide enrichment of proteins modified by lysine methylation. *Nature protocols* 9, 37-50, doi:10.1038/nprot.2013.164 (2014).
- 62 Carlson, S. M., Moore, K. E., Green, E. M., Martín, G. M. & Gozani, O. Proteome-wide enrichment of proteins modified by lysine methylation. *Nat. Protocols* 9, 37-50, doi:10.1038/nprot.2013.164
- <http://www.nature.com/nprot/journal/v9/n1/abs/nprot.2013.164.html> - supplementary-information (2014).

- 63 Olsen, J. B. *et al.* Quantitative Profiling of the Activity of Protein Lysine Methyltransferase SMYD2 Using SILAC-Based Proteomics. *Molecular & Cellular Proteomics* 15, 892-905, doi:10.1074/mcp.M115.053280 (2016).
- 64 Donlin, L. T. *et al.* Smyd2 controls cytoplasmic lysine methylation of Hsp90 and myofilament organization. *Genes & Development* 26, 114-119, doi:10.1101/gad.177758.111 (2012).
- 65 Cao, Q. F. *et al.* Characterization of the Human Transcription Elongation Factor Rtf1: Evidence for Nonoverlapping Functions of Rtf1 and the Paf1 Complex. *Mol Cell Biol* 35, 3459-3470, doi:10.1128/MCB.00601-15 (2015).
- 66 van Agthoven, T. *et al.* Relevance of breast cancer antiestrogen resistance genes in human breast cancer progression and tamoxifen resistance. *J Clin Oncol* 27, 542-549, doi:10.1200/JCO.2008.17.1462 (2009).
- 67 Zinder, J. C., Wasmuth, E. V. & Lima, C. D. Nuclear RNA Exosome at 3.1 Å Reveals Substrate Specificities, RNA Paths, and Allosteric Inhibition of Rrp44/Dis3. *Mol Cell* 64, 734-745, doi:10.1016/j.molcel.2016.09.038 (2016).
- 68 Hamamoto, R., Toyokawa, G., Nakakido, M., Ueda, K. & Nakamura, Y. SMYD2-dependent HSP90 methylation promotes cancer cell proliferation by regulating the chaperone complex formation. *Cancer Lett* 351, 126-133, doi:10.1016/j.canlet.2014.05.014 (2014).
- 69 Corbett, M. A. *et al.* A novel X-linked trichothiodystrophy associated with a nonsense mutation in RNF113A. *J Med Genet* 52, 269-274, doi:10.1136/jmedgenet-2014-102418 (2015).
- 70 Lear, T. *et al.* RING finger protein 113A regulates C-X-C chemokine receptor type 4 stability and signaling. *Am J Physiol Cell Physiol* 313, C584-C592, doi:10.1152/ajpcell.00193.2017 (2017).
- 71 Brickner, J. R. *et al.* A ubiquitin-dependent signalling axis specific for ALKBH-mediated DNA dealkylation repair. *Nature* 551, 389-393, doi:10.1038/nature24484 (2017).
- 72 Hong, X. *et al.* Targeting posttranslational modifications of RIOK1 inhibits the progression of colorectal and gastric cancers. *Elife* 7, doi:10.7554/eLife.29511 (2018).
- 73 Moore, K. E. *et al.* A general molecular affinity strategy for global detection and proteomic analysis of lysine methylation. *Mol Cell* 50, 444-456, doi:10.1016/j.molcel.2013.03.005 (2013).
- 74 Ng, H. H., Dole, S. & Struhl, K. The Rtf1 component of the Paf1 transcriptional elongation complex is required for ubiquitination of histone H2B. *J Biol Chem* 278, 33625-33628, doi:10.1074/jbc.C300270200 (2003).
- 75 Van Oss, S. B. *et al.* The Histone Modification Domain of Paf1 Complex Subunit Rtf1 Directly Stimulates H2B Ubiquitylation through an Interaction with Rad6. *Mol Cell* 64, 815-825, doi:10.1016/j.molcel.2016.10.008 (2016).
- 76 Zhang, W. *et al.* Prediction and prognostic significance of BCAR3 expression in patients with multiple myeloma. *J Transl Med* 16, 363, doi:10.1186/s12967-018-1728-8 (2018).
- 77 Wilson, A. L., Schrecengost, R. S., Guerrero, M. S., Thomas, K. S. & Bouton, A. H. Breast cancer antiestrogen resistance 3 (BCAR3) promotes cell motility by regulating actin cytoskeletal and adhesion remodeling in invasive breast cancer cells. *PLoS One* 8, e65678, doi:10.1371/journal.pone.0065678 (2013).

- 78 Cross, A. M. *et al.* Breast cancer antiestrogen resistance 3-p130(Cas) interactions promote adhesion disassembly and invasion in breast cancer cells. *Oncogene* 35, 5850-5859, doi:10.1038/onc.2016.123 (2016).

TABLES

Table 1: Selected new candidates for SMYD2 and SMYD3. 6 candidates for SMYD2 and 15 for SMYD3 were selected based on combination of ProtoArray® and 3xMBT pull down data and their biological relevance. SMYD2's substrate redundant with Olsen *et al.* study are marked with an asterisk (*).

KMT	Generic name	Protein Identification	Detection technique
SMYD2	BCAR3	Breast cancer anti-estrogen resistance protein 3	3xMBT *
SMYD2	LSM14B	Protein LSM14 homolog B	3xMBT *
SMYD2	RRP44	Exosome complex exonuclease	3xMBT *
SMYD2	RTF1	RNA polymerase-associated protein RTF1 homolog	3xMBT *
SMYD2	SARG2	Specifically androgen-regulated gene protein	3xMBT *
SMYD2	ACTC1	Actin, alpha cardiac muscle 1	3xMBT
SMYD3	ACTC1	Actin, alpha cardiac muscle 1	3xMBT
SMYD3	ARHGAP17	Rho GTPase activating protein 17	ProtoArray®
SMYD3	ARHGEF5	Rho guanine nucleotide exchange factor 5	3xMBT
SMYD3	ASZ1	Ankyrin repeat, SAM and basic leucine zipper domain-containing protein 1	ProtoArray®
SMYD3	CLUAP1	Clusterin associated protein 1	ProtoArray®
SMYD3	HSP90	Heat shock protein 90 kDa	3xMBT
SMYD3	LMNA	Lamin A/C	ProtoArray®
SMYD3	LSM14B	Protein LSM14 homolog B	3xMBT
SMYD3	NOC2L	Nucleolar complex associated 2 homolog	ProtoArray®
SMYD3	NPM1	Nucleophosmin (<i>S. cerevisiae</i>)	ProtoArray®
SMYD3	PIN4	Peptidyl-prolyl cis-trans isomerase NIMA-interacting 4	ProtoArray®
SMYD3	RIOK1	RIO kinase 1 (yeast)	ProtoArray®
SMYD3	RNF113A	RING finger protein 113A	ProtoArray®
SMYD3	RUVBL2	RuvB-like 2	3xMBT
SMYD3	SP100	SP100 nuclear antigen	ProtoArray®

Table 2: Production of recombinant candidate substrates and *in vitro* confirmation of methylation events with recombinant SMYD2 or SMYD3.

KMT	Substrate	Cloning	Recombinant protein production	Methylation assay result
SMYD2	BCAR3	Yes	Yes	Methylated
SMYD2	LSM14B	Yes	No	N/A
SMYD2	RRP44	Yes	Yes	Methylated
SMYD2	RTF1	Yes	Yes	Methylated
SMYD2	SARG2	Yes	Yes	Methylated
SMYD2	ACTC1	Yes	Yes	Not methylated
SMYD3	ACTC1	Yes	Yes	Not methylated
SMYD3	ARHGAP17	No	N/A	N/A
SMYD3	ARHGEF5	No	N/A	N/A
SMYD3	ASZ1	No	N/A	N/A
SMYD3	CLUAP1	No	N/A	N/A
SMYD3	HSP90	Yes	Yes	Methylated
SMYD3	LMNA	No	N/A	N/A
SMYD3	LSM14B	Yes	No	N/A
SMYD3	NOC2L	Yes	No	N/A
SMYD3	NPM1	Yes	Yes	Not methylated
SMYD3	PIN4	Yes	Yes	Not methylated
SMYD3	RIOK1	Yes	Yes	Methylated
SMYD3	RNF113A	Yes	Yes	Methylated
SMYD3	RUVBL2	Yes	Yes	Not methylated
SMYD3	SP100	Yes	Yes	Not methylated

FIGURES

Figure 1: Identification of new candidate substrates for SMYD2 and SMYD3. (A) Identification of new SMYD3 substrates using ProtoArray[®] microchip incubated with recombinant SMYD3 with reaction buffer and radiolabeled S-adenosyl methionine (SAM) overnight. The right panel is a close-up of the indicated area on the array. (B) Representative flow chart of SILAC 3xMBT pulldown with HeLa cells extracts exposed to active or inactive recombinant SMYD2. (C) Example of mass spectrometry analysis results of SMYD2 SILAC 3xMBT pulldown.

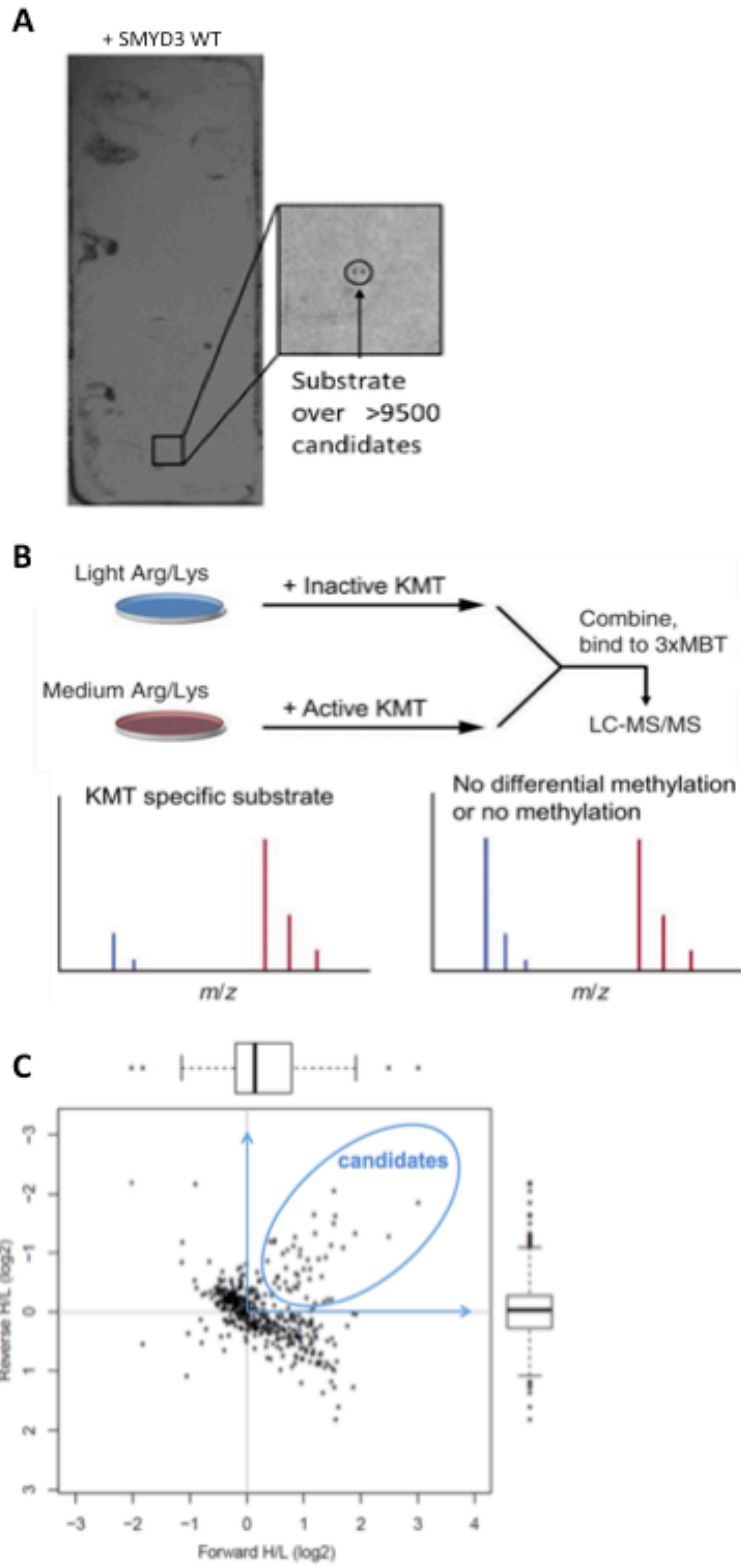


Figure 2: *In vitro* confirmation of methylation of SMYD2 and SMYD3's candidate substrates using radioactive methylation assay with $^3\text{H-SAM}$. (A) Methylation assay of SMYD2 candidate substrates RTF1, BCAR3 and RRP44, incubated with or without recombinant SMYD2 (4 hours exposure time). (B) Methylation assay of SMYD2 candidate substrate SARG compared with RTF1 either wild type or K587R mutant incubated with SMYD2 (4 hours exposure time). (C) Methylation assay of all candidate substrates obtained by GST-pulldown in presence of SMYD3 (4 hours exposure time). (D) Methylation assay of RNF113A and HSP90, incubated with or without the recombinant SMYD3 (4 hours exposure time). (E) Methylation assay of RIOK1 either alone, with SMYD3, or with SETD7 (15 days exposure time). In all figures, top panels are autoradiogram of methylation assay while bottom panels are coomassie stains of proteins in the reaction. Red arrows designate KMTs (SMYD2 or SMYD3) and black arrows designate substrates.

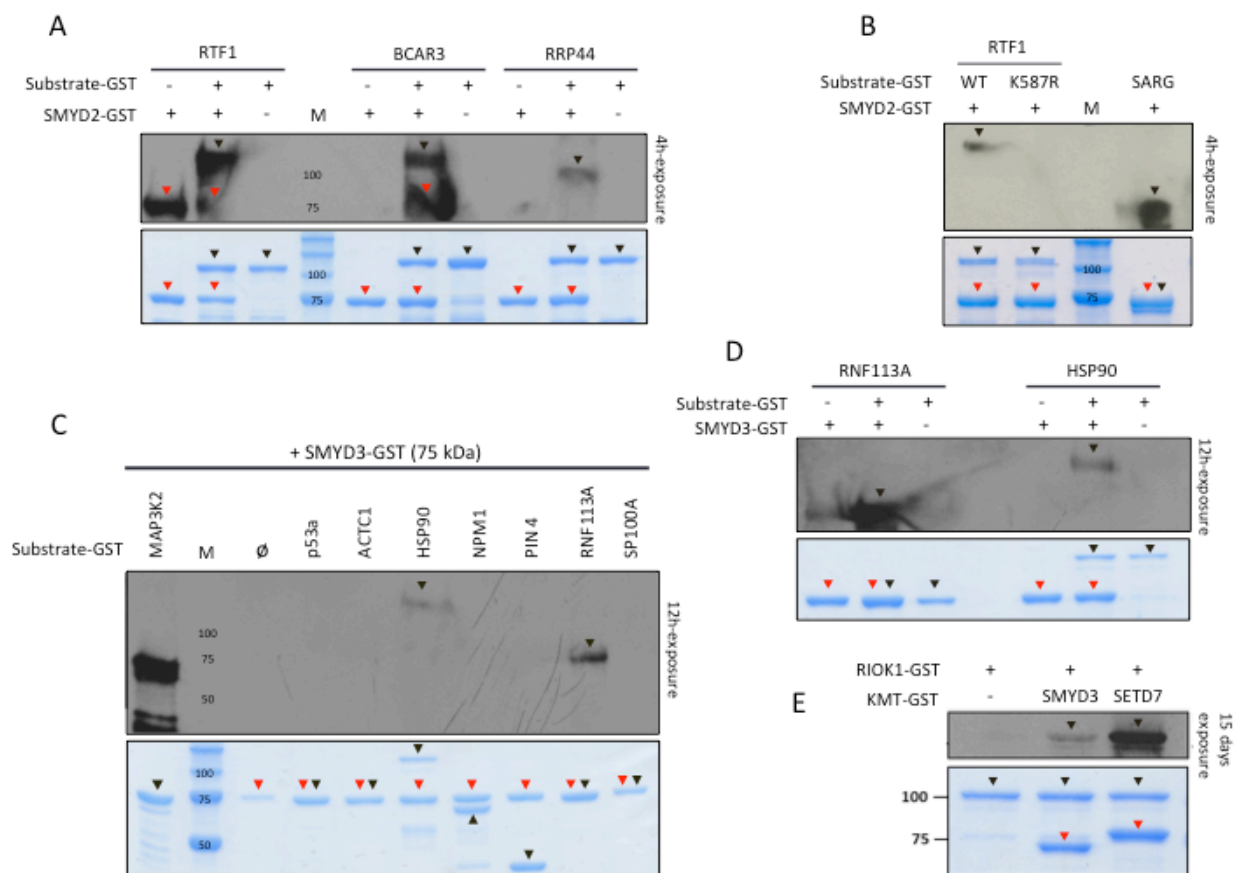


Figure 3: Identification of candidate substrates' methylation site. (A) Identification of RNF113A K20me3 by MS/MS analysis of samples from non-radioactive methylation assay

using recombinant RNF113A and recombinant active SMYD3 with SAM. Confirmation of methylated lysines by *in vitro* radioactive methylation assays of candidate substrates (B) RTF1, (C) BCAR3, (D) RRP44, (E) RNF113A in their wild type and K>R variants incubated with or without the recombinant KMT in presence of H³-SAM. (F) Additional experiment was conducted with K>A (alanine) variants for RNF113A with comparable conditions to confirm the exact site of methylation. In all figures, top panels are autoradiogram of methylation assay while bottom panels are coomassie stains of proteins in the reaction.

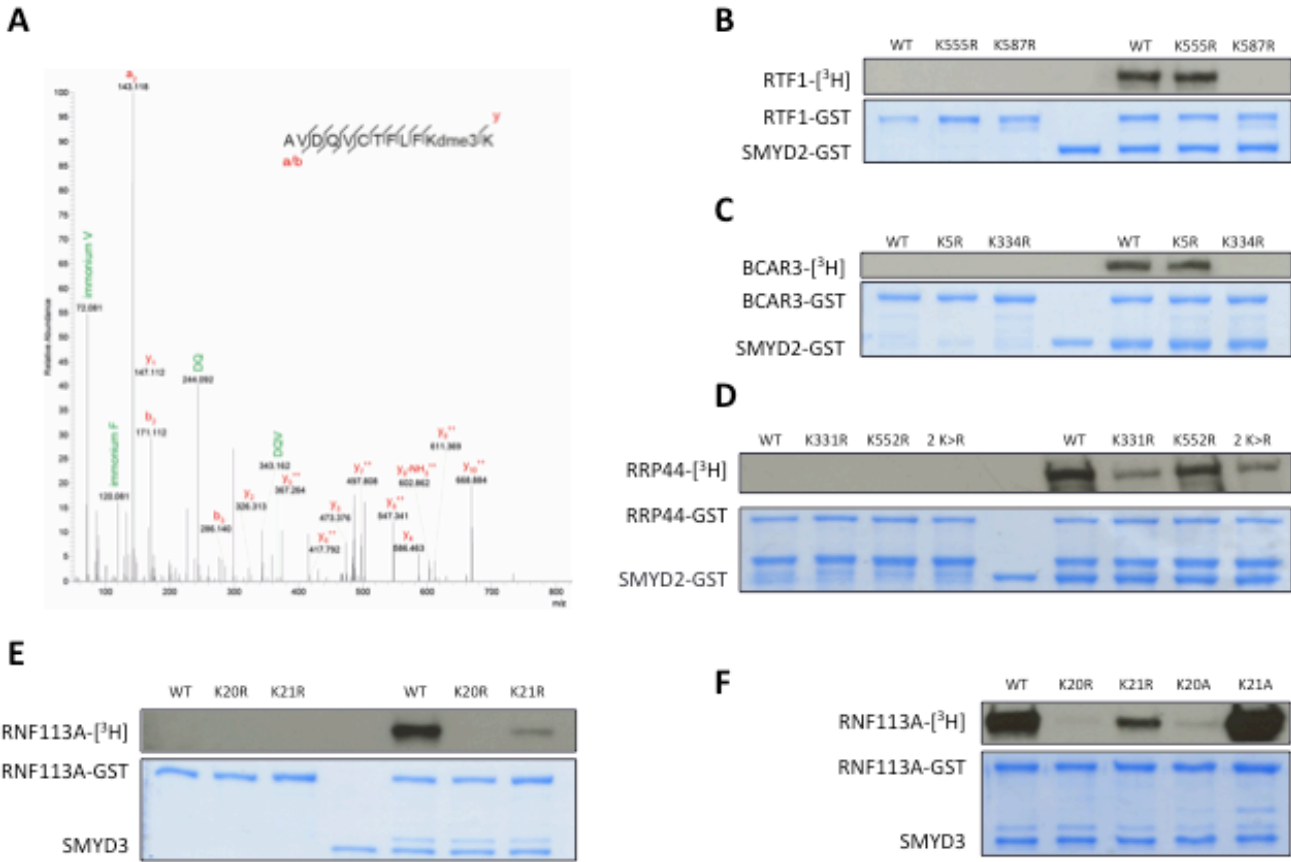


Figure 4: Validation of engineered methylated-lysine containing peptides and methyl specific antibodies. Peptide dotblots of (A) RTF1 K587me0 and K587me1 peptides with immunodetection by streptavidin antibody and RTF1 K587me1 specific antibody. (B) BCAR3 K334me0 and K334me1 with immunodetection by streptavidin antibody and BCAR3 K334me1 antibody. (C) RNF113A K20me0, me1, me2 and me3 peptides with immunodetection by streptavidin antibody and RNF113A K20me3 specific antibody. (D) Peptide dotblots of non-radioactive methylation assays using RNF113A K20me0, me1, me2 and me3 peptides incubated with recombinant SMYD3 and SAM, with immunodetection by streptavidin antibody and RNF113A K20me3 specific antibody. Peptide dotblots of non-radioactive methylation assay of (E) RNF113 WT or K20R mutant in presence of recombinant SMYD3 with immunodetection by K20me3 specific antibody and total RNF113A antibody. (E) RTF1 WT or K587R mutant in presence of recombinant SMYD2 with immunodetection by K587 me1 specific antibody and total RTF1 antibody. (F) BCAR3 WT or K334R mutant in presence of recombinant SMYD2 with immunodetection by K334 me1 specific antibody and total BCAR3 antibody. For panels E, F and G, top two lanes are autoradiogram of methylation assay while bottom panels are coomassie stains of proteins in the reaction.

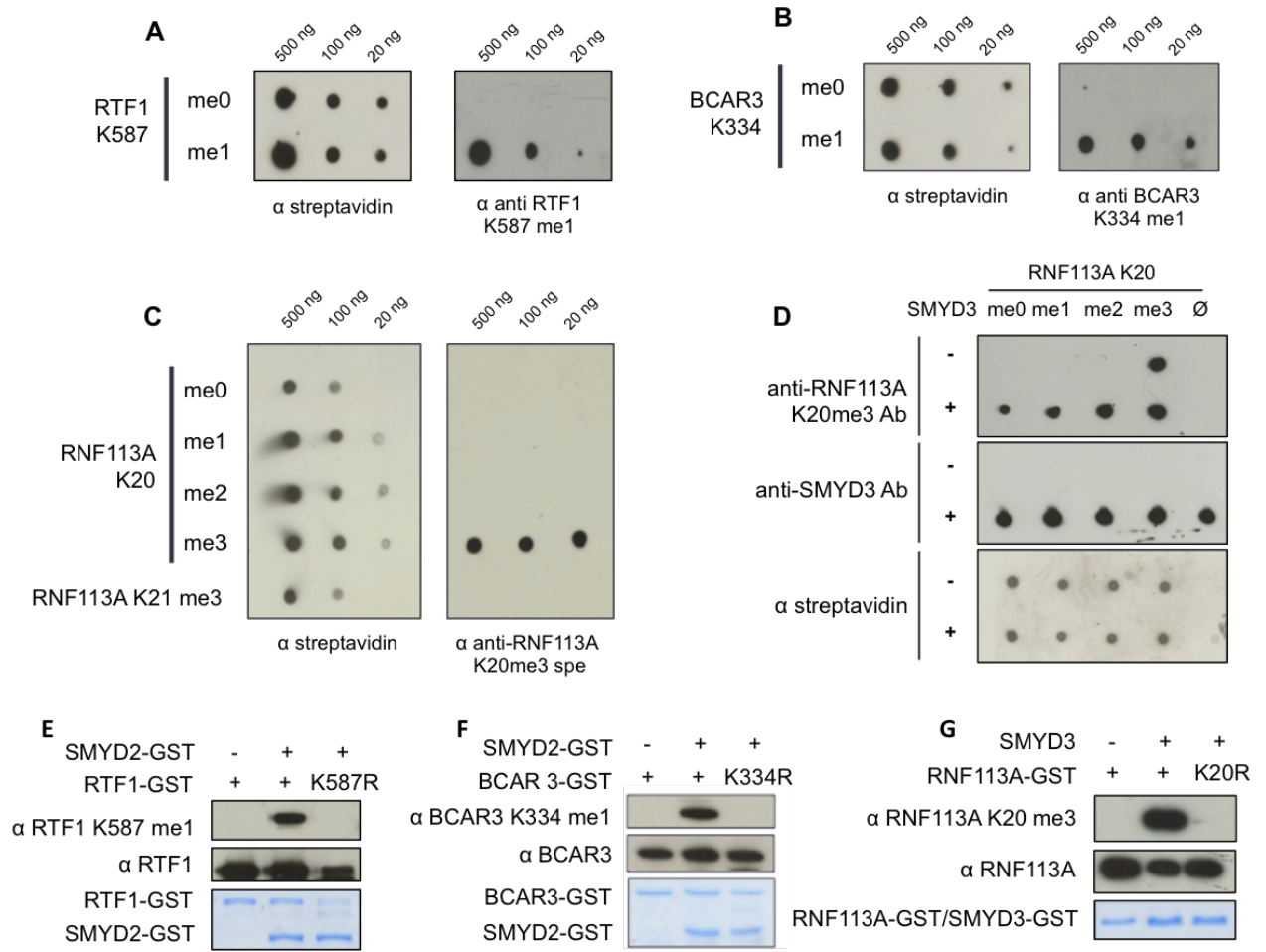


Figure 5: Confirmation of ectopic methylation events in cells. 293T cells were transfected with pCDNA-HA wild type and K>R/A mutants of either (A) RTF1, (B) BCAR3 and (C) RNF113A with or without SMYD3 overexpression, followed by HA immunoprecipitation and immunodetection with commercially available total and methyl specific antibodies.

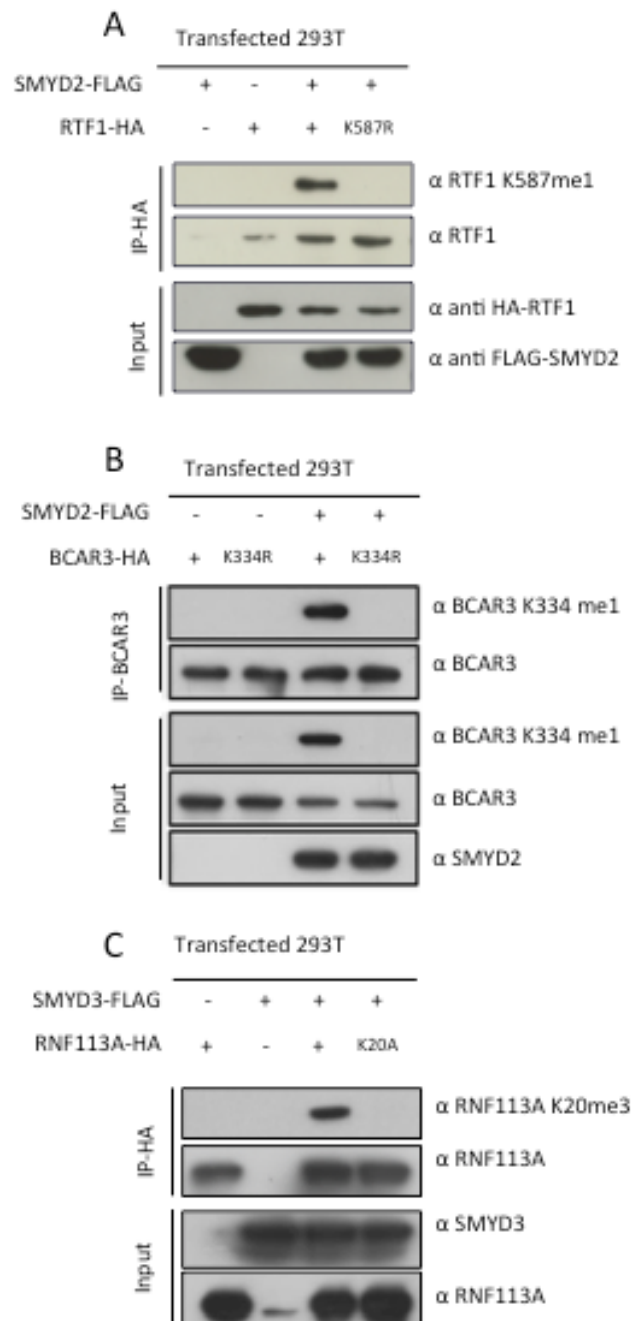
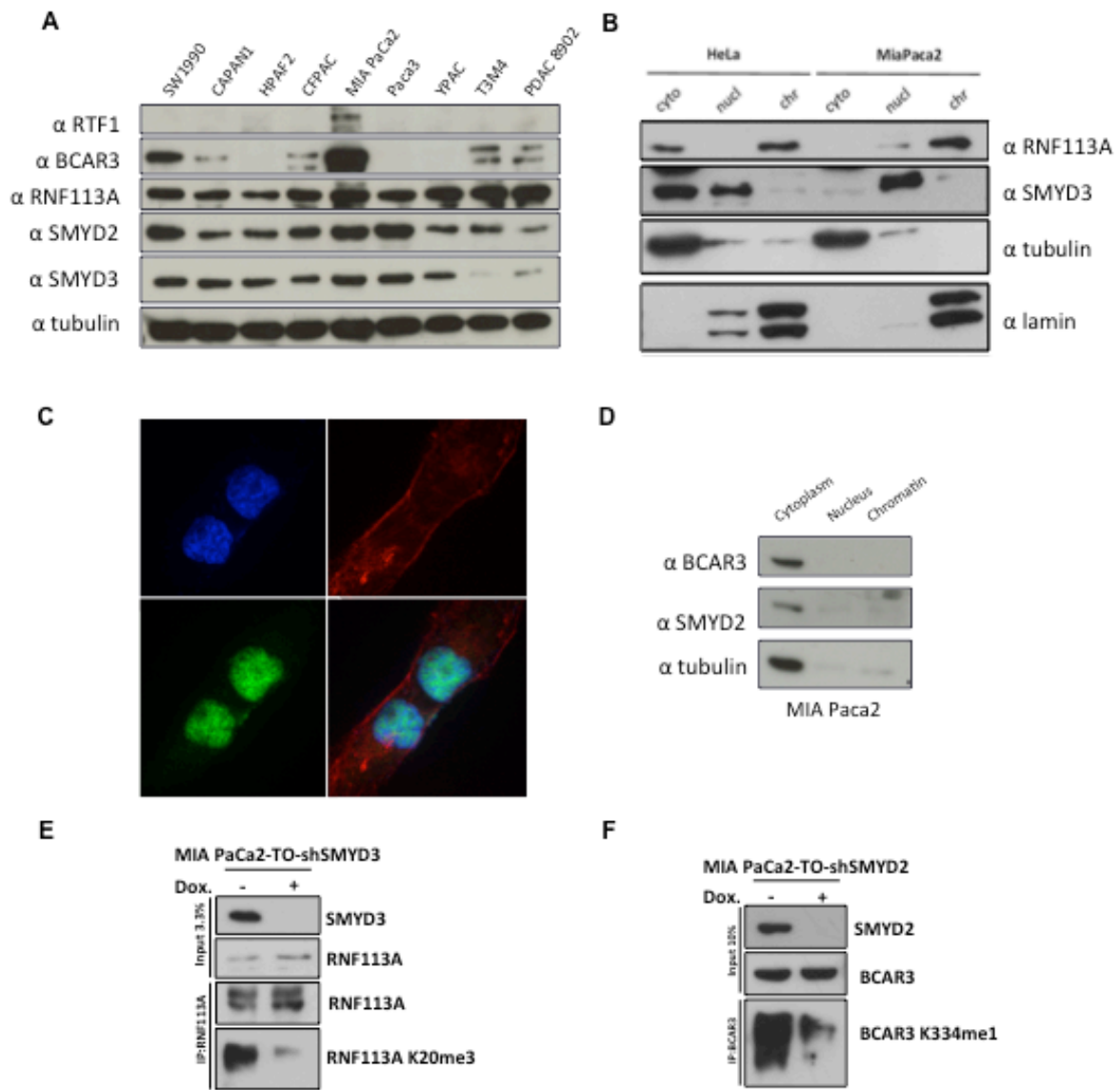


Figure 6: Detection of substrates endogenous methylation in cancer cell lines. (A) Screening of substrates and KMTs' expression profiles in whole cell protein extracts of different PDAC cell lines. (B) Expression of RNF113A and SMYD3 in MIA PaCa-2 and HeLa cells' cytoplasmic, nuclear and chromatin extracts obtained by cell fractionation. (C) Representative picture of RNF113A nuclear expression in MIA PaCa-2 cells by immunofluorescence (upper left: DAPI; upper right: phalloidin; lower left: RNF113A-GFP; lower right: merge of DAPI, phalloidin and RNF113A-GFP). (D) Expression of BCAR3 and SMYD2 in MIA PaCa-2 cells from cytoplasmic, nuclear and chromatin extracts obtained by cell fractionation. (E) Confirmation of endogenous methylation of RNF113A by SMYD3 in MIA PaCa-2 wild type compared with MIAPaCa-2 *Smyd3*^{-/-} cells. (F) Confirmation of endogenous methylation of BCAR3 by SMYD2 in MIA PaCa-2 wild type compared with MIAPaCa-2 *Smyd2*^{-/-} cells. Doxycycline inducible TET-on-shRNA of respectively SMYD3 and SMYD2 were used in (E) and (F) followed by immunodetection using lysine-methyl specific antibodies.



SUPPLEMENTARY DATA

Supplementary Table 1: SMYD2's substrates list obtained by *in vitro* ProtoArray® methylation assays. Methyltransferases constituting false positive due to automethylation are written in grey and italic. Substrates previously described and characterized in literature are written in bold characters. Substrates identified by both *in vitro* ProtoArray® and SILAC 3xMBT-pulldown are marked with an asterisk (*).

Generic Name - Protein Identification
6CKINE - C-C motif chemokine 21
ADD2 - adducin 2 (beta) (ADD2)
ADH4 - alcohol dehydrogenase 4 (class II), pi polypeptide
AGGF1 - Angiogenic factor with G patch and FHA domains 1
AHI1 - Joubertin
ALK - anaplastic lymphoma kinase (Ki-1)
AMMECR1L - AMME chromosomal region gene 1-like
ANKZF1 - ankyrin repeat and zinc finger domain containing 1
APEX1 - APEX nuclease (multifunctional DNA repair enzyme) 1, transcript variant 1
APOBEC3C - apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3C (APOBEC3C), mRNA.
APTX - aprataxin (APTX), transcript variant 4
ARHGEF5 - Rho guanine nucleotide exchange factor 5
ARL6IP4 - ADP-ribosylation factor-like protein 6-interacting protein 4
ARPC1B - actin related protein 2/3 complex, subunit 1B
ASAP - Microtubule-associated protein 9
ASXL1 - additional sex combs like 1 (Drosophila)
AURKA - aurora kinase A, transcript variant 2
AURKB - Serine/threonine-protein kinase 12
BAZ2B - bromodomain adjacent to zinc finger domain, 2B
BBX - bobby sox homolog (Drosophila)
BCKDK - branched chain ketoacid dehydrogenase kinase
BCL7B - B-cell CLL/lymphoma 7B, transcript variant 2
BCL7C - B-cell CLL/lymphoma 7C
BMX - BMX non-receptor tyrosine kinase, transcript variant 2
BMX - Cytoplasmic tyrosine-protein kinase
BRDG1 - signal transducing adaptor family member 1
BTF3 - basic transcription factor 3, transcript variant 2
BUD31 - BUD31 homolog (S. cerevisiae)
BYSL - bystin-like (BYSL)
C11orf31 - Selenoprotein H

C13orf27 - chromosome 13 open reading frame 27
C14orf111 - rRNA-processing protein FCF1 homolog
C17orf85 - Uncharacterized protein C17orf85
C1orf116 - Specifically androgen-regulated gene protein
C1orf55 - UPF0667 protein C1orf55
C22orf33 - chromosome 22 open reading frame 33
C6orf130 - chromosome 6 open reading frame 130
C6orf166 - chromosome 6 open reading frame 166
C6orf201 - Uncharacterized protein C6orf201
C8orf59 - chromosome 8 open reading frame 59
C9orf32 - chromosome 9 open reading frame 32
C9orf9 - chromosome 9 open reading frame 9
CASK - Peripheral plasma membrane protein CASK
CCDC23 - coiled-coil domain containing 23
CCDC55 - coiled-coil domain containing 55, transcript variant 1
CCL19 - chemokine (C-C motif) ligand 19
CDC2 - cell division cycle 2, G1 to S and G2 to M, transcript variant 1
CDC7 - Cell division cycle 7-related protein kinase
CDK5 - cyclin-dependent kinase 5 and p25
CDKL1 - Cyclin-dependent kinase-like 1
CDKL3 - Cyclin-dependent kinase-like 3
CDKN2AIP - CDKN2A interacting protein
CFDP1 - craniofacial development protein 1
<i>CGI-09 - tRNA methyltransferase 6 homolog (S. cerevisiae)</i>
CHEK2 - Serine/threonine-protein kinase Chk2
CHMP2B - chromatin modifying protein 2B
CLIC5 - chloride intracellular channel 5
COPS5 - COP9 constitutive photomorphogenic homolog subunit 5 (Arabidopsis)
CORO6 - coronin 6
CSAG3A - CSAG family, member 3A
CSNK1G1 - casein kinase 1, gamma 1
CSNK1G2 - Casein kinase I isoform gamma-2
CSNK2A2 - casein kinase 2, alpha prime polypeptide
CT45A1 - cancer/testis antigen family 45, member A1
CTNND1 - Catenin delta-1
CUTC - cutC copper transporter homolog (E. coli)
DAPK3 - death-associated protein kinase 3 (DAPK3)
DCUN1D2 - DCN1, defective in cullin neddylation 1, domain containing 2 (S. cerevisiae)
DCX - doublecortex; lissencephaly, X-linked (doublecortin), transcript variant 4
DCX - Neuronal migration protein doublecortin
DDX1 - DEAD (Asp-Glu-Ala-Asp) box polypeptide 1
DDX21 - DEAD (Asp-Glu-Ala-Asp) box polypeptide 21
DDX49 - DEAD (Asp-Glu-Ala-Asp) box polypeptide 49
DDX54 - DEAD (Asp-Glu-Ala-Asp) box polypeptide 54
DDX55 - DEAD (Asp-Glu-Ala-Asp) box polypeptide 55
DDX6 - Probable ATP-dependent RNA helicase

DGCR8 - DiGeorge syndrome critical region gene 8
DKFZp761G2113 - Ataxin-7-like protein 3
DNAJB7 - DnaJ (Hsp40) homolog, subfamily B, member 7
DNAJC8 - DnaJ homolog subfamily C member 8
DNTT - deoxynucleotidyltransferase, terminal, transcript variant 1
EDF1 - endothelial differentiation-related factor 1, transcript variant alpha
EFS - embryonal Fyn-associated substrate, transcript variant 2
EHHADH - Peroxisomal bifunctional enzyme
EIF1AY - eukaryotic translation initiation factor 1A, Y-linked
EIF2A - eukaryotic translation initiation factor 2A, 65kDa
ELOF1 - elongation factor 1 homolog (S. cerevisiae)
EPB41L4A - Band 4.1-like protein 4A
EPHA3 - Ephrin receptor A3, transcript variant 1
FAM116B - family with sequence similarity 116, member B
FAM128A - family with sequence similarity 128, member A
FAM130A2 - family with sequence similarity 130, member A2
FAM131B - family with sequence similarity 131, member B
FAM21C - family with sequence similarity 21, member C
FAM54A - family with sequence similarity 54, member A
FAU - Finkel-Biskis-Reilly murine sarcoma virus
FGD2 - FYVE, RhoGEF and PH domain-containing protein 2
FGF12 - fibroblast growth factor 12
FLJ11184 - hypothetical protein FLJ11184
FSD1 - Fibronectin type III and SPRY domain-containing protein 1
GABARAP - GABA(A) receptor-associated protein
GAS2L3 - GAS2-like protein 3
GCC1 - GRIP and coiled-coil domain-containing protein 1
GLIPR2 - Golgi-associated plant pathogenesis-related protein 1
GTPBP10 - GTP-binding protein 10
HADH - hydroxyacyl-Coenzyme A dehydrogenase
HDGFRP3 - hepatoma-derived growth factor, related protein 3
HIST1H2BD - Histone H2B type 1-D
HMGN1 - high-mobility group nucleosome binding domain 1
HMGN3 - high mobility group nucleosomal binding domain 3, transcript variant 2
HNRPA1 - heterogeneous nuclear ribonucleoprotein A1, transcript variant 1
HP1BP3 - heterochromatin protein 1, binding protein 3
HSP90AA1 - Heat shock protein HSP 90-alpha *
IFIT5 - interferon-induced protein with tetratricopeptide repeats 5
IL20 - interleukin 20
IL21 - interleukin 21
IMPACT - Impact homolog (mouse)
INSR - Insulin receptor
IPPK - inositol 1,3,4,5,6-pentakisphosphate 2-kinase
ITAC - C-X-C motif chemokine 11
JMJD2D - jumonji domain containing 2D
KCNAB1 - potassium voltage-gated channel, shaker-related subfamily, beta member 1

KCTD18 - potassium channel tetramerisation domain containing 18
KIAA0020 - Pumilio domain-containing protein
KIAA0515
KIAA0859, transcript variant 2
KIAA1530 protein
KIF3A - kinesin family member 3A
KIF3B - kinesin family member 3B
KPNA2 - Importin subunit alpha-2
LAD1 - ladinin 1 (LAD1)
LAMC1 - laminin, gamma 1 (formerly LAMB2) (LAMC1)
LARP1 - La ribonucleoprotein domain family, member 1
LIMS2 - LIM and senescent cell antigen-like domains 2
LOC57228 - small trans-membrane and glycosylated protein, transcript variant 2
LRRC6 - leucine rich repeat containing 6
LSM4 - LSM4 homolog, U6 small nuclear RNA associated (S. cerevisiae)
MAP2 - microtubule-associated protein 2, transcript variant 2
MAP3K2 - Mitogen-activated protein kinase kinase kinase 2
MAPKAPK3 - MAP kinase-activated protein kinase 3
MCM10 - Protein MCM10 homolog
MCM2 - minichromosome maintenance complex component 2
METAP2 - methionyl aminopeptidase 2
MPP7 - membrane protein, palmitoylated 7
MYLK - myosin, light polypeptide kinase, transcript variant 5
NEK7 - NIMA (never in mitosis gene a)- related kinase 7
NKRF - NF-kappaB repressing factor (
NOL7 - nucleolar protein 7
NOSTRIN - Nostrin
ODF2 - outer dense fiber of sperm tails 2
PAK4 - p21(CDKN1A)-activated kinase 4, transcript variant 1
PCLO - piccolo (presynaptic cytomatrix protein)
<i>PCMT1 - protein-L-isoaspartate (D-aspartate) O-methyltransferase (PCMT1)</i>
<i>PCMT1 - Protein-L-isoaspartate(D-aspartate) O-methyltransferase</i>
PDAP1 - PDGFA associated protein 1
PHF20L1 - PHD finger protein 20-like 1, transcript variant 3
PIN4 - protein (peptidylprolyl cis/trans isomerase) NIMA-interacting 4
PIP5K2C - phosphatidylinositol-5-phosphate 4-kinase, type II, gamma
PLK1 - Serine/threonine-protein kinase PLK1
POLB - polymerase (DNA directed), beta
PPAN - Suppressor of SWI4 1 homolog
PRC1 - protein regulator of cytokinesis 1, transcript variant 1
PRKAA2 - AMP-activated protein_kinase A2/B1/G1
PRKAG2 - protein kinase, AMP-activated, gamma 2 non-catalytic subunit
PRKRA - Interferon-inducible double stranded RNA-dependent protein kinase activator A
PRR15 - proline rich 15
PRR16 - Proline-rich protein 16
PSMD4 - proteasome 26S subunit, non-ATPase, 4, transcript variant 2

PSRC1 - proline/serine-rich coiled-coil 1, transcript variant 1
PTPN6 - Tyrosine-protein phosphatase non-receptor type 6
PTRH1 - peptidyl-tRNA hydrolase 1 homolog (S. cerevisiae)
PUS7L - Pseudouridylylase 7 homolog-like protein
RABL3 - Rab-like protein 3
RAC1 - Ras-related C3 botulinum toxin substrate 1
RAD51AP1 - RAD51 associated protein 1
RDBP - RD RNA binding protein
RECQL - ATP-dependent DNA helicase Q1
RG9MTD1 - RNA (guanine-9-)-methyltransferase domain-containing protein 1, mitochondrial
RIT1 - Ras-like without CAAX 1
RIT2 - Ras-like without CAAX 2
RP13-36C9.3 - cancer/testis antigen CT45-3
RPL11 - 60S ribosomal protein L11
RPL22 - 60S ribosomal protein L22
RPL23A - 60S ribosomal protein L23a
RPL29 - 60S ribosomal protein L29
RPL30 - 60S ribosomal protein L30
RPL31 - 60S ribosomal protein L31, transcript variant 1
RPL39L - 60S ribosomal protein L39-like
RPS19 - 40S ribosomal protein S19
RPS20 - 40S ribosomal protein S20
RPS25 - 40S ribosomal protein S25
RPS27L - 40S ribosomal protein S27-like protein
RPS28 - 40S ribosomal protein S28
RSBN1L - round spermatid basic protein 1-like
RTF1 - Paf1/RNA polymerase II complex component, homolog (S. cerevisiae)
RTN4 - reticulon 4, transcript variant 3
SCEL - sciellin
SELS - selenoprotein S
SERBP1 - SERPINE1 mRNA binding protein 1
SERBP1 - SERPINE1 mRNA binding protein 1, transcript variant 4
SFRS1 - splicing factor, arginine/serine-rich 1 (splicing factor 2, alternate splicing factor), transcript variant 1
SFRS2 - splicing factor, arginine/serine-rich 2
SFRS5 - Splicing factor, arginine/serine-rich 5
SFRS6 - splicing factor, arginine/serine-rich 6
SIRT6 - sirtuin (silent mating type information regulation 2 homolog) 6 (S. cerevisiae)
SLAIN2 - SLAIN motif family, member 2
<i>SMYD2 - SET and MYND domain-containing protein 2</i>
SNRPA1 - NA
SPANXN4 - Sperm protein associated with the nucleus on the X chromosome N4
SPRR4 - small proline-rich protein 4
SRP19 - signal recognition particle 19kDa
SSH3 - Protein phosphatase Slingshot homolog 3

STAT6 - signal transducer and activator of transcription 6, interleukin-4 induced
STK6 - Serine/threonine-protein kinase 6
SUDS3 - Sin3 histone deacetylase corepressor complex component
SUPT16H - FACT complex subunit SPT16
SURF2 - surfait 2
SYTL2 - synaptotagmin-like 2, transcript variant a
TARBP2 - TAR (HIV-1) RNA binding protein 2, transcript variant 3
TBC1D22B - TBC1 domain family member 22B
TCEA1 - transcription elongation factor A, 1, transcript variant 1
TCEAL1 - Transcription elongation factor A protein-like 1
THEX1 - three prime histone mRNA exonuclease 1
TNIK - TRAF2 and NCK interacting kinase
TP53I13 - Tumor protein p53-inducible protein 13
TPT1 - tumor protein, translationally-controlled 1
TRA2A - Transformer-2 protein homolog
TRIM41 - Tripartite motif-containing protein 41
TRIT1 - tRNA isopentenyltransferase 1
TTK - Dual specificity protein kinase TTK
TTL1 - tubulin tyrosine ligase-like family, member 1, transcript variant 2
UBD - Ubiquitin D
UBD - ubiquitin D (UBD)
UBE2E2 - ubiquitin-conjugating enzyme E2E 2 (UBC4/5 homolog, yeast)
UBE2S - ubiquitin-conjugating enzyme E2S
UBXD3 - UBX domain containing 3
UFD1L - ubiquitin fusion degradation 1 like (yeast), transcript variant 1
UNKL - unkempt homolog (Drosophila)-like
UTP14A - UTP14, U3 small nucleolar ribonucleoprotein, homolog A (yeast)
VAMP4 - vesicle-associated membrane protein 4
VCY - variable charge, Y-linked 1B
VRK1 - vaccinia related kinase 1
WDR5 - WD repeat domain 5, transcript variant 1
<i>WHSC1 - Wolf-Hirschhorn syndrome candidate 1, transcript variant 9</i>
XAF1 - XIAP associated factor-1, transcript variant 1
ZC3H3 - zinc finger CCCH-type containing 3
ZMYND11 - zinc finger, MYND domain containing 11
ZNF323 - Zinc finger protein 323

Supplementary Table 2: SMYD3's substrates list obtained by *in vitro* ProtoArray®

methylation assay. Methyltransferases constituting false positive due to automethylation are written in grey and italic. Substrates previously described and characterized in literature are written in bold characters. Substrates identified by both *in vitro* ProtoArray® and SILAC 3xMBT-pulldown are marked with an asterisk (*).

Generic name	Protein Identification
ARHGAP17	Rho GTPase activating protein 17
ASZ1	Ankyrin repeat, SAM and basic leucine zipper domain-containing protein 1
C9orf41	Chromosome 9 open reading frame 41
CLUAP1	Clusterin associated protein 1
<i>COMT</i>	<i>Catechol-O-methyltransferase</i>
COPS5	COP9 constitutive photomorphogenic homolog subunit 5 (Arabidopsis)
<i>FAM119A</i>	<i>Family with sequence similarity 119, member A</i>
<i>LCMT1</i>	<i>Leucine carboxyl methyltransferase 1</i>
LMNA	Lamin A/C
MAP3K2	Mitogen-activated protein kinase kinase kinase 2 *
NOC2L	Nucleolar complex associated 2 homolog (S. cerevisiae)
NPM1	Nucleophosmin
PCLO	Piccolo presynaptic cytomatrix protein
<i>PCMT1</i>	<i>Protein-L-isoaspartate(D-aspartate) O-methyltransferase</i>
PIN4	Peptidyl-prolyl cis-trans isomerase NIMA-interacting 4
<i>PRMT1</i>	<i>Protein arginine N-methyltransferase 1</i>
<i>PRMT6</i>	<i>Protein arginine N-methyltransferase 6</i>
<i>PRMT8</i>	<i>Protein arginine N-methyltransferase 8</i>
RIOK1	RIO kinase 1 (yeast)
RNF113A	Ring finger protein 113A
<i>SMYD2</i>	<i>SET and MYND domain-containing protein 2</i>
SP100	SP100 nuclear antigen

Supplementary Table 3: SMYD2's substrates list obtained by 3xMBT pulldown in HeLa cells. Substrates previously described and characterized in literature are written in bold characters. Substrates identified by both *in vitro* ProtoArray® and SILAC 3xMBT-pulldown are marked with an asterisk (*). Substrates redundant with Olsen *et al.* study are mentioned on the right column.

Generic name	Protein Identification	Olsen <i>et al.</i>
ACAD9	Acyl-CoA dehydrogenase family member 9, mitochondrial	
ACTBL2	Beta-actin-like protein 2	
ACTC1	Actin, alpha cardiac muscle 1	
AHNAK	Neuroblast differentiation-associated protein	Yes
ARHGEF5	Rho guanine nucleotide exchange factor 5	
ATP5A1	ATP synthase subunit alpha, mitochondrial	Yes
BCAR3	Breast cancer anti-estrogen resistance protein 3	Yes
CAPZB	F-actin-capping protein subunit beta	Yes
CFL1	Cofilin-1	Yes
CUTC	Copper homeostasis protein cutC homolog	
CYFIP1	Cytoplasmic FMR1-interacting protein 1	
DBNL	Drebrin-like protein	Yes
DDX6	Probable ATP-dependent RNA helicase DDX6	Yes
DIAPH1	Protein diaphanous homolog 1	Yes
EEF2	Elongation factor 2	Yes
EFTS	Isoform 2 of Elongation factor Ts, mitochondrial	Yes
GTF3C4	General transcription factor 3C polypeptide 4	Yes
H1	Histone 1	
HSD17B10	3-hydroxyacyl-CoA dehydrogenase type-2	
HSP90b	Heat shock protein HSP 90-beta *	Yes
HSPA1A	Heat shock 70 kDa protein 1A/1B	Yes
KIF5B	Kinesin-1 heavy chain	
KLC1	Kinesin light chain 1	Yes
KPNA2	Importin subunit alpha-2	
KPNA3	Importin subunit alpha-3	
LSM14b	Protein LSM14 homolog B	Yes
MAP1S	Microtubule-associated protein 1S	Yes
MARS	Methionine--tRNA ligase, cytoplasmic	
OLA1	Obg-like ATPase 1	Yes
PDHA1	Pyruvate dehydrogenase E1 component subunit alpha, mitochondrial	Yes
PFKP	6-phosphofructokinase type C	Yes
PHGDH	D-3-phosphoglycerate dehydrogenase	

PRRC2C	Protein PRRC2C	Yes
PSME1	Proteasome activator complex subunit 1	Yes
PSME2	Proteasome activator complex subunit 2	Yes
Rb	Retinoblastoma-associated protein	Yes
RBMX	RNA-binding motif protein, X chromosome	Yes
Rnf214	RING finger protein 214	
RPS27A	Ubiquitin-40S ribosomal protein S27a	
RRP44	Exosome complex exonuclease	Yes
RTF1	RNA polymerase-associated protein RTF1 homolog	Yes
SARG	Specifically androgen-regulated gene protein	Yes
SDHA	Succinate dehydrogenase flavoprotein subunit, mitochondrial	Yes
SLC3A2	4F2 cell-surface antigen heavy chain	
SSB	Lupus La protein	Yes
TFG	Protein TFG	
TOP2A	DNA topoisomerase 2-alpha	Yes
TRIM33	Isoform Beta of E3 ubiquitin-protein ligase	
TSFM	Elongation factor Ts, mitochondrial	
TUBB	Tubulin beta chain	Yes
Tubb4b	Tubulin beta-4B chain	
TXNRD1	Thioredoxin reductase 1, cytoplasmic	Yes
UBR7	Putative E3 ubiquitin-protein ligase	
VARS	Valine--tRNA ligase	
VWA5A	von Willebrand factor A domain-containing protein 5A	
XPO1	Exportin-1	

Supplementary Table 4: SMYD3's substrates list obtained by 3xMBT pulldown in HeLa cells. Substrates previously described and characterized in literature are written in bold characters. Substrates identified by both *in vitro* ProtoArray® and SILAC 3xMBT-pulldown are marked with an asterisk (*).

Generic name	Protein Identification
4F2hc	4F2 cell-surface antigen heavy chain
ACTBL2	Beta-actin-like protein 2
ACTC1	Actin, alpha cardiac muscle 1
AHNAK	Neuroblast differentiation-associated protein AHNAK
ARHGEF5	Rho guanine nucleotide exchange factor 5
ARL4C	ADP-ribosylation factor 1
ATP5A1	ATP synthase subunit alpha, mitochondrial
COPG1	Coatomer subunit gamma-1
HSD17B10	3-hydroxyacyl-CoA dehydrogenase type-2
HSPB1	Heat shock protein beta-1
HSP90b	Heat shock protein HSP 90-beta
IPO7	Importin-7
KPNA3	Importin subunit alpha-3
KPNA4	Importin subunit alpha-4
LSM14B	Protein LSM14 homolog B
MARS	Methionine--tRNA ligase, cytoplasmic
MAP3K2	Mitogen-activated protein kinase kinase kinase 2 *
OLA1	Obg-like ATPase 1
PFKP	6-phosphofructokinase type C
PHGDH	D-3-phosphoglycerate dehydrogenase
RBMX	RNA-binding motif protein, X chromosome
RPL6	60S ribosomal protein L6
RPS27A	Ubiquitin-40S ribosomal protein S27a
RUVBL2	RuvB-like 2
SLIRP	SRA stem-loop-interacting RNA-binding protein, mitochondrial
TFG	Protein TFG
TFRC	Transferrin receptor protein 1
TUBB	Tubulin beta chain
TUBB4B	Tubulin beta-4B chain
UBR7	Putative E3 ubiquitin-protein ligase UBR7
XPO-1	Exportin-1
XPO-2	Exportin-2

**PARTIE II : LA TRIMÉTHYLATION DE
RNF113A PAR SMYD3 INDUIT UNE
AUGMENTATION DE SON ACTIVITÉ E3
UBIQUITINE LIGASE ET PARTICIPE À LA
CHIMIORÉSISTANCE AUX AGENTS
ALKYLANTS**

I. RESUME DE LA PARTIE II

Le développement de mécanismes de résistance aux médicaments chimiothérapeutiques anticancéreux représente un enjeu majeur dans le traitement du cancer. Une meilleure compréhension de ces mécanismes et comment les contrecarrer ou les contourner est nécessaire pour proposer des thérapies alternatives contre le cancer. La lysine méthyltransférase SMYD3 (domaine SET et MYND contenant la protéine 3) est surexprimée dans de nombreux cancers, y compris le cancer du poumon à petites cellules (SCLC), et est généralement associée à un mauvais pronostic. Ici, nous démontrons que SMYD3 triméthyle la ligase E3-ubiquitine RNF113A (protéine de doigt RING 113A) et régule son activité dans la réparation de la désalkylation de l'ADN. Nous avons utilisé une approche protéomique pour identifier RNF113A comme un nouveau substrat de SMYD3. La triméthylation de RNF113 K20 induite par SMYD3 altère son interaction avec la protéine phosphatase PP4 dans les cellules cancéreuses. Ce blocage conduit à un niveau accru de phosphorylation de RNF113A, qui est essentiel pour son activité E3-ubiquitine ligase. Nous avons confirmé dans les lignées cellulaires SCLC que l'événement de méthylation SMYD3-RNF113A potentialise l'activité de RNF113A impliquée dans la réparation de la désalkylation de l'ADN et conduit à la chimiorésistance des cellules aux agents alkylants tels que le méthanesulfonate de méthyle. Par conséquent, nous proposons ici un nouveau mécanisme expliquant les fonctions oncogènes complexes de SMYD3 dans le cancer du poumon, à travers la promotion de la résistance aux dommages alkylants.

ARTICLE

SMYD3 promotes cell resistance to DNA alkylation damage by methylating the E3 ubiquitin ligase RNF113A

Valentina Lukinovic^{1#}, Josh Brickner^{2#}, Gael S. Roth^{1,3#}, Ning Tsao², Clement Oyeniran², Alexandre Casanova¹, Jessica Vayr¹, Yohan Couté⁴, Pascal Jansen⁵, Michiel Vermeulen⁵, Pierre Hainaut¹, Pawel K. Mazur⁶, Nima Mosammaparast² and Nicolas Reynoird^{1‡}

1. Institute for Advanced Biosciences - INSERM U1209 / CNRS UMR 5309 / Université Grenoble-Alpes, France.
6. Department of Pathology & Immunology, Washington University in St. Louis School of Medicine, St. Louis, MO 63110, USA.
7. Clinique universitaire d'Hépatogastroentérologie et Oncologie digestive, CHU Grenoble-Alpes
8. Biosciences and Biotechnology Institute of Grenoble – INSERM U1038 / CEA, France.
9. Department of Molecular Biology, Faculty of Science, Radboud Institute for Molecular Life Sciences, Oncode Institute, Radboud University Nijmegen, 6525 GA, Nijmegen, the Netherlands.
10. Department of Experimental Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA.

these authors contributed equally to this work.

‡ corresponding author: Nicolas Reynoird - Nicolas.reynoird@univ-grenoble-alpes.fr
Institute for Advanced Biosciences - INSERM U1209 / CNRS UMR 5309 / Université Grenoble-Alpes, 38600 la Tronche, France

Keywords : SMYD3, RNF113A, DNA repair, alkylation, small cell lung cancer

ABSTRACT

Acquired resistance to anticancer chemotherapeutic drugs represents a major issue in the treatment of cancer. A better understanding of the underlying mechanisms and how to counteracts or bypass them is required to propose alternative cancer therapies. The lysine methyltransferase SMYD3 (SET and MYND domain containing protein 3) is overexpressed in many cancers, including lung cancer, and is generally associated with poor prognosis. Here, we demonstrate that SMYD3 trimethylates the E3-ubiquitin ligase RNF113A (RING finger protein 113A) and regulates its activity in DNA dealkylation repair. We used proteomics approach to identify RNF113A as a new substrate of SMYD3. SMYD3-dependant trimethylation of RNF113 K20 impaired the interaction with the protein phosphatase PP4 in cells. This blockage leads to an increased level of RNF113A phosphorylation, which is essential for its E3-ubiquitin ligase activity. We confirmed that SMYD3-RNF113A methylation event enhances RNF113A activity involved in DNA dealkylation repair and leads to cell chemoresistance to alkylating agents such as Methyl Methanesulfonate. Therefore, we propose here a new mechanism explaining the complex oncogenic functions of SMYD3 in lung cancer through the promotion of chemoresistance to DNA alkylation damages.

INTRODUCTION

DNA damages repair represents a major issue in the understanding of tumorigenesis after exposure to carcinogens, as well as the development of chemoresistance mechanisms under cytotoxic treatments. Multiple types of DNA damages occur due to endogenous sources such as hydrolysis, oxidation, alkylation, or DNA bases mismatch, as well as exogenous causes including ionizing or ultraviolet radiations and various chemicals agents ¹. To counteract such detrimental damages to DNA for cell homeostasis, several DNA damage repair systems exist and has been well characterized ^{1,2}. For example, mismatch repair (MMR) is involved in the repair of erroneous insertion, deletion, and mis-incorporations of DNA bases ³. Other systems are involved in DNA single strand lesions such as nucleotide excision repair (NER) or base excision repair (BER), which intervenes in the repair of chemical damages of bases such as deamination, oxidation, and alkylation ⁴. These different repair systems are important for a cell fate and are often altered during cancers ⁵. Indeed, DNA alterations present in early oncogenesis generate driven mutations which promote tumor development. Moreover, alterations of DNA damage repair systems can also be part of the chemoresistance developed after cancer therapies ⁶. However, inducing DNA damages is still the most frequently used therapeutic way to kill highly proliferative cancer cells and treat cancer. For example, alkylation DNA damages represent a classical therapeutic strategy to kill cancer cells with a large panel of existing alkylating agents used in chemotherapy ⁷. In response to alkylation damages, several systems are activated in cells to counteract these modifications, such as the O-6-methylguanine-DNA methyltransferase (MGMT), an alkyltransferase that removes alkylations ², or the activating signal cointegrator complex (ASCC) ^{8,9}.

Recently, the RING finger protein 113A (RNF113A) was described as essential for the functioning of the ASCC complex in DNA dealkylation repair ⁹. This E3-ubiquitin ligase was first described in a rare autosomal recessive disease called trichothiodystrophy, in which the non-sense mutation c.901 C> T, p.Q301 * is responsible of multiple phenotypic consequences such as intellectual impairment, cerebellar hypoplasia, loss of consciousness,

facial dysmorphism, abnormal capillary implantation, endocrinopathies, and immunoglobulin deficiency ¹⁰. Moreover, X-linked trichothiodystrophy patients are hypersensitive to alkylating agents with defective formation of ASCC foci in cells, due to dysfunction in the alkylation damage signal transduction ⁹. DNA alkylation activates RNF113A, that act as a transducer of alkylation signal and induces the activation of ASCC2 which will recruit DNA repair enzymes such as ASCC3 or ALKBH3 (alkylated DNA Repair Protein AlkB Homolog 3) ⁹. The specific sensor of this signaling, as well as upstream regulation of RNF113A to properly transduce DNA dealkylation repair system have yet to be understood. Finely tuned regulation of RNF113A by post-translational modifications is one of the possible systems that could modulate its capacity to coordinate DNA dealkylation repair.

Post-translational modifications (PTM) of proteins are involved in many physiological and pathological processes including cancer ¹¹. Growing data pinpoint lysine methylation, a dynamic and finely tuned PTM, as essential in oncogenesis by regulating chromatin functions upon histone modifications. However, increasing evidences suggest that lysine methylation signaling is not restricted only to histones, with a major role for this PTM in different cellular processes ¹². Lysine methylation events are catalyzed by lysine-methyltransferases (KMTs) which constitute a large group of more than 200 enzymes divided in two main families: SET (Su(var)3-9, Enhancer-of-zeste and Trithorax) domain-containing KMTs and 7 β -strand domain-containing KMTs ¹³⁻¹⁵. Among SET KMTs, SMYD3 (SET and MYND domain containing protein 3) is overexpressed in many cancers and involved in various oncogenic pathways ^{16,17}. SMYD3 oncogenic role was recently demonstrated in the canonical Ras/Raf/Mek/Erk signaling pathway, through MAP3K2 (Mitogen-Activated Protein Kinase Kinase 2) trimethylation in RAS-dependent lung and pancreatic adenocarcinomas ¹⁸. However, as exemplified by other intensively studied KMTs such as EZH2 (Enhancer of zeste homolog 2) involved in many different signalings, SMYD3 oncogenic activity probably occurs via multiple pathways. Notably, SMYD3 is associated with tumor cell proliferation and its overexpression is correlated to poor prognosis in many cancers where RAS pathway does not constitute a driver signaling ¹⁹.

In this study, we demonstrate a new crosstalk between two post-translational modifications of RNF113A important for its function. Methylation of RNF113A by SMYD3 indirectly regulates RNF113A phosphorylation level, required for its E3-ubiquitin ligase activity in DNA alkylation repair. Thus, we propose a new mechanism of lung cancers resistance to alkylation damage, through increased activity of RNF113A upon SMYD3 methylation.

RESULTS

SMYD3 dependent SCLC development

Besides its already described function in RAS-driven lung adenocarcinoma through the methylation of MAP3K2¹⁸, SMYD3 is also overexpressed in small cell lung cancer (SCLC), a cancer that is not driven by Ras-signaling pathway. Therefore, we developed a SCLC mouse model with conditional SMYD3 knock-out. Small Cell Lung Cancer mouse model was obtained with an inducible triple knock out of Rb1, Rbl2 and p53 triggered by Cre recombinase at 8 weeks to obtain Rb^{-/-};Rbl2^{-/-};p53^{-/-} TKO mice (Figure S1). To study the impact of SMYD3 on tumor growth, we then crossed these mice with SMYD3 conditional knock-out mice, and we compared tumor progression between Smyd3 WT vs Smyd3^{-/-} mice. At 26 weeks, absence of SMYD3 drastically decreased the development of SCLC with a significant difference in tumor area or tumor volume quantified by mCT imaging and immunohistochemistry (Figure 1A and 1B). SCLC from TKO;Smyd3 mice contained less proliferative cells and more apoptotic cells (Figure 1A and C). As a result, SMYD3 knock-out mice's survival was extended by approximately one third, an impressive feature knowing the aggressiveness of SCLC (Figure 1D). Therefore, SMYD3 expression is important for SCLC progression, a cancer where SMYD3 oncogenic activity is unlikely to be related to its already described function in the RAS pathway.

Identification of RNF113A as a new substrate of SMYD3

To identify new substrates of SMYD3 that could explain its oncogenic activity in SCLC, we performed an unbiased high throughput approach based on methylation assay on a ProtoArray® platform (Figure S2A). Among approximately 9500 potential candidates, we identified 22 hits as methylated in presence of SMYD3, including the previously characterized substrate MAP3K2 (Table 1). While several hits were discarded as false positives being automethylated KMTs, several candidates were further tested for validation using *in vitro* methylation assay. Using MAP3K2 substrate as control, we confirmed

RNF113A as a new candidate successfully methylated by SMYD3 (Figure 2A). Further MS/MS analyses on methylated RNF113A revealed the presence of trimethylation on lysine K20 (Figure 2B). We noted that the enzyme works in a processive manner as all methylated states of lysine K20 could be observed by mass spectrometry (Figure S2B). To confirm this lysine as the single site of methylation, *in vitro* methylation assays were performed with RNF113A mutant K20A and revealed that SMYD3 only methylates RNF113A on its lysine K20 (Figure 2C). In order to validate the presence of RNF113A methylation in cells, we raised a RNF113A K20me3 methyl specific antibody and validated its specificity (Figure S2C). Additionally, we observed that SMYD3 uses either unmodified, mono or dimethylated RNF113K20 as template to reach trimethylation level (Figure S2D). Next, under more physiologic cellular context, we confirmed the methylation of ectopic RNF113A by SMYD3 in human 293T cells and not for the RNF113A K20A mutant (Figure S2E). Finally, we were able to detect endogenous RNF113A K20 trimethylation in cancer cells, and to demonstrate its loss upon inducible repression of SMYD3 (Figure 2D). Thus far, we discovered a new physiological substrate of SMYD3 and demonstrated that RNF113A is trimethylated on its lysine K20 both *in vitro* and in cancer cells by SMYD3.

RNF113A methylation repels the phosphatase PP4

RNF113A is a protein involved in splicing as well as in DNA dealkylation repair ^{9,20}. Interestingly, a recent report also linked RNF113A to lung cancer resistance to Cisplatin ²¹. This prompted us to decipher any potential regulation of RNF113A activity upon SMYD3 methylation. Lysine methylation signaling predominantly affects the binding capacity of a protein. Therefore, to identify RNF113A methyl-sensitive interactors that are specifically repelled or attracted by SMYD3 methylation, we performed a SILAC (Stable Isotope Labeling of Amino Acids in Cell Culture) peptide pulldown using biotinylated peptides containing either unmodified or trimethylated versions of RNF113A lysine K20. Contrary to the general belief that methylation event recruits specific methyl-readers, we did not discovered methyl-specific RNF113A binders but rather a list of proteins strongly associated with K20me0 and repelled

by RNF113A K20me3. Interestingly, among the top hits identified in three independent experiments, we discovered 4 proteins belonging to the serine/threonine-protein phosphatase complex PP4: PPP4C, PPP4R2, PPP4R3A and PP4R3B (Figure 3A, Figure S3A). This was particularly appealing as we previously identified a similar mechanism with another substrate of SMYD3, as MAP3K2 methylation repels the phosphatase PP2a¹⁸. The phosphatase complex PP4 is composed of 3 subunits: PPP4C, the catalytic subunit, the chaperone PPP4R2 allowing the cohesion of the complex, and either one of the substrate specific binders PPP4R3A or PPP4R3B (Figure S3B). Notably, PPP4R3A was the most strongly associated with RNF113A, consistent with its function for PP4 substrate's recognition. Interestingly, a recent study of PPP4R3A identified the specific binding motif of this phosphatase subunit as "FxxP", where the first x is preferentially a lysine, perfectly matching the "FKKP" sequence of RNF113A with the first K being the lysine K20²². Therefore, trimethylation of lysine K20 by SMYD3 within the FxxP recognition motif is likely to interfere with the binding capacity of the phosphatase complex. Indeed, we confirmed that recombinant PPP4R3A directly binds to RNF113A peptide and that the trimethylation of lysine K20 abrogates its association (Figure 3B). We noted that the recombinant PPP4C subunit was able to bind to both versions of RNF113A peptide, however this interaction was probably artefactual as ectopic PPP4C produced in 293T cells did not bind RNF113A peptide at all, as expected for the catalytic subunit of PP4 (Figure 3B). Moreover, endogenous PPP4R3A from different cancer cells extracts bound to unmethylated RNF113A and was repelled by either di- or trimethylation of lysine K20 (Figure 3C and S3C). Finally, we demonstrated that this specific binding could occur in the context of the full RNF113A protein and not just a peptide surrounding K20, as PPP4R3A co-immunoprecipitated with ectopic RNF113A in 293T cells (Figure 3D). Interestingly, while mutation of RNF113A lysine K20 into alanine did not affect this binding, the replacement with a bulkier amino acid such as phenylalanine blocked the interaction, mimicking the effect of lysine K20 trimethylation by SMYD3 (Figure 3D). We further confirmed this observation by pulldown assay of PPP4R3A from 293T extract with recombinant RNF113A, as PPP4R3A bound to RNF113A WT, K20A

and K20R mutants, but not K20F mutant (Figure S3D). Therefore, we demonstrated that the phosphatase PP4 binds to RNF113A through its PPP4R3A subunit, and that RNF113A K20 trimethylation by SMYD3, or K20F mutation, efficiently blocks this interaction.

RNF113A phosphorylation is required for its E3-ubiquitin ligase activity

As we identified a loss of interaction between RNF113A and the phosphatase PP4, we postulated that RNF113A is phosphorylated and that its phosphorylation is an important requirement for its activity. Indeed, while performing RNF113A immunodetection after electrophoresis, we noticed that the protein migrated higher than its theoretical molecular weight, and pre-incubation of cellular extract with commercially available phosphatases restored its expected weight (Figure 4A and Figure S4A). Moreover, PhosphoSitePlus database classifies RNF113A as a heavily phosphorylated protein ²³ and we confirmed several potential phosphorylated residues of RNF113A by MS/MS analysis of RNF113A purified from HeLa cells (Figure S4B). Moreover, As RNF113A was recently described as an E3-ubiquitin ligase stimulated by DNA alkylation damages, we aimed to demonstrate that its level of phosphorylation was linked to its activity of ubiquitination. First, we confirmed by His-Ub pulldown that RNF113A is specifically activated by the alkylating agent methyl methanesulfonate (MMS), but no other DNA damaging agents (Figure 4B). Then, we performed *in vitro* poly-Ub chain assay, where an active E3 ligase incubated with recombinant E1 and E2 ligases should form ubiquitin chains that can be observed as smear by electrophoresis. As recombinant RNF113A is inactive ⁹, we purified the E3 ligase from engineered HeLa cells. Remarkably, while purified RNF113A is capable to form ubiquitin chains in this system, its pre-incubation with the FastAP Thermosensitive Alkaline Phosphatase completely abrogated its E3 ligase activity *in vitro* (Figure 4C). Finally, we confirmed that the PP4 phosphatase dephosphorylates RNF113A by monitoring the phosphorylation level of RNF113A after incubation with the purified catalytic PP4C subunit (Figure 4D). Therefore, we demonstrated that RNF113A is phosphorylated in cancer cells

and that this phosphorylation level, which is regulated by the phosphatase PP4, is essential for its E3-ubiquitin ligase activity upon alkylation damages.

SMYD3 indirectly regulates phosphorylation of RNF113A and its E3 ligase activity

We previously observed that active RNF113A is phosphorylated, and that its methylation by SMYD3 repels the phosphatase complex PP4. We also demonstrated that the K>F mutation on RNF113A lysine K20 mimics the blockage of PP4 interaction. Therefore, we took advantage of the RNF113A K20F mutant to study the impact of SMYD3 methylation on RNF113A activity, induced by the incapacity of PP4 to dephosphorylate and inactivate RNF113A. We hypothesized that RNF113A K20F mutant is constitutively phosphorylated once stimulated, leading to a constant active form of RNF113A. To test this, we used different experimental systems to monitor RNF113A ubiquitination or autoubiquitination activity. First, we used tandem ubiquitin binding element (TUBE) conjugated beads to isolate endogenously ubiquitinated proteins from HeLa cells, and observed that RNF113A is ubiquitinated upon MMS treatment (Figure 5A and Figure S5A). We confirmed this result in a different system, by using overexpression of His-tagged Ubiquitin in cells and enrichment of ubiquitinated proteins by Ni-NTA beads (Figure S5B). To address the possibility that this ubiquitination signal might not be RNF113A autoubiquitination but rather the result of other E3-ligases, we deleted the catalytic RING domain of RNF113A and observed a complete loss of ubiquitination signal upon MMS treatment, both by TUBE assay and His-Ub pulldown (Figure 5A and Figure S5C). Then, we compared RNF113A WT and K20F mutant autoubiquitination, and observed that while the basal activity of RNF113A K20F mutant is slightly increased compared to wild type, the non-unactivable mutant is significantly more ubiquitinated upon MMS treatment (Figure 5B and Figure S5C). This result suggests a specific activation of RNF113A after alkylation damages, either by a direct phosphorylation of the protein or by an increased population of phosphorylated RNF113A in cancer cells. This also suggests that PP4, and indirectly SMYD3 through RNF113A methylation, plays an important function in regulating RNF113A activity by impacting its phosphorylation level. To

support this hypothesis, we monitored the RNF113A ligase activity and its phosphorylation at different time points after MMS treatment. Strikingly, we observed a fast and transient activation of RNF113A 30 minutes after MMS treatment that rapidly decreased at 1 hour and returned to basal level at 2 hours (Figure 5C). Moreover, this peak of RNF113A activity perfectly matched a transient increase of RNF113A phosphorylation at 30min followed by a clear decrease afterwards, while the methylation of RNF113A remained constant (Figure 5D). As we did not detect change of total RNF113A protein level between the different time points, this demonstrates that RNF113A activity relies on its phosphorylation, which is stimulated upon MMS-induced alkylation damages. Our observations also demonstrate that the cycle of RNF113A phosphorylation and dephosphorylation is an active process, down-regulated by the phosphatase PP4. Next, we compared the phosphorylation level of both wild-type and K20F mutant of RNF113A with or without MMS stimulation, and discovered a significantly higher phosphorylation of the mutant under both basal and MMS-stimulated contexts (Figure 5E). As expected, we observed that this difference in phosphorylation level correlates with a strikingly higher ligase activity of RNF113A K20F mutant compared to wild-type after MMS stimulation (Figure 5F). This observation suggests that compared to RNF113A WT, RNF113A K20F mutant that mimics SMYD3 methylation can be activated upon MMS stimulation but cannot be inactivated afterwards by PP4. Confirming our hypothesis, we noticed that the FastAP phosphatase was able to dephosphorylate both wild-type and K20F RNF113A, while PP4 could only dephosphorylate efficiently the wild-type form of RNF113A, further cementing the importance of the crosstalk between SMYD3 methylation and PP4 on the regulation of the RNF113A E3 ligase activity (Figure S5D). Therefore, we demonstrated that while RNF113A's activation upon MMS treatment correlates with an increase of phosphorylation, its rapid inactivation and dephosphorylation depends on the efficiency of PP4 activity that can be altered by SMYD3 methylation of RNF113A.

RNF113A regulation by SMYD3 is associated with cell resistance to alkylating agents

RNF113A E3 ligase activity plays a central role in DNA dealkylation repair by the ASCC complex⁹. To test the validity of our model of RNF113A activity being regulated by a crosstalk between phosphorylation and methylation modifications by PP4 and SMYD3, we studied its impact on resistance to alkylation damage response. Indeed, because alkylation-damage agents are an important aspect of chemotherapies, and RNF113A had recently been linked to lung cancer Cis-platin chemoresistance²¹, we wondered if RNF113A's regulation by SMYD3 could also participate in cancer cells resistance to alkylation-based therapy. We took advantage of the SCLC cell line DMS-114, originating from a patient before chemotherapy and with low expression of RNF113A and SMYD3, to study the resistance to MMS treatment after engineering of RNF113A and SMYD3. We observed a significant acquired resistance of DMS-114 cells with ectopic expression of RNF113A and SMYD3 to both alkylating agents MMS and Carmustine (Figure 6A and 6B).

Therefore, in addition to our extensive molecular analysis of RNF113A regulation by post-translational modifications crosstalk induced by SMYD3 and PP4, we demonstrated here for the first time the relevance of SMYD3 and RNF113A in SCLC. RNF113A ligase activity is essential for response to alkylation damages, and by indirect stimulation of its activity, SMYD3 enhances the chemoresistance promoted by RNF113A to alkylating agents in SCLC (Figure 7).

DISCUSSION

Our work demonstrates that SMYD3 participates in small cell lung cancer progression and resistance to alkylation-based therapy. Here, we identified a novel substrate of SMYD3, the E3 ubiquitin ligase RNF113A, and demonstrated that SMYD3 trimethylates its lysine K20. This post-translational event increases RNF113A phosphorylation level and indirectly regulates its function in DNA dealkylation repair. We carried out an in-depth molecular analysis of RNF113A regulation by post-translational modifications crosstalk with consequences on cell resistance to damaging agents. This is the first time that a KMT as SMYD3 is involved in DNA dealkylation repair, and the first description of a methylation mark regulating RNF113A activity.

As SCLC is highly aggressive and lethal with few therapeutic options, it is of great importance to study further if SMYD3 and RNF113A could serve as prognostic markers for SCLC progression or its capacity to acquire chemotherapy resistance. Therefore, further studies are required to fully decipher the molecular mechanism of RNF113A post-translation regulation, and to assess its impact on RNF113A-dependant chemoresistance.

First, how RNF113A K20 trimethylation interferes with RNF113A-PP4 interaction is unclear. Methylation, due to its small steric hindrance, does not cause changes in protein structure nor affects the charge of modified protein residues. However, this event can potentially affect hydrogen bonding capacity of lysines. Therefore, lysine methylation is a regulator of protein-protein interactions, often through changes in the abovementioned bonding capacity. A better structural analysis of the interaction between RNF113A and PP4 will give better insights on how SMYD3 methylation can impede it. Then, while we clearly demonstrated that RNF113A phosphorylation is important for its activity, the specific site of phosphorylation, the kinase responsible for it and how this kinase senses alkylation damages are still undescribed. Moreover, while it seems that RNF113A is already phosphorylated at basal level, its activation upon MMS treatment clearly correlates with an increase of RNF113A phosphorylation level. Increase of RNF113A phosphorylation level can either be due to

specific new phosphorylation marks on single RNF113A proteins, or to a higher ratio of phospho-RNF113A/RNF113A among its general population. The link between RNF113A phosphorylation and its E3-ubiquitin ligase activity remains unclear as well. We demonstrated that deletion of the RING domain abolishes RNF113A, and it is possible that RNF113A structural conformation hides its RING domain, and that phosphorylation events counteract its auto-inhibition and promote an open conformation.

RNF113A promotes ASCC dealkylation repair activity through its E3 ubiquitin ligase activity, and we postulate that SMYD3 methylation increases dealkylation repair and promotes cancer cell resistance to alkylation damages. However, it is still unclear how enhanced RNF113A activity affects the ASCC machinery. ASCC2 contains a CUE domain belonging to the ubiquitin-binding domain superfamily ²⁴, and RNF113A ubiquitin signaling could participate in the ASCC complex assembly and functional activation. Therefore, SMYD3 methylation, by increasing the relative functional population of RNF113A or the duration of RNF113A activity, could promote ASCC complex assembly and enhance its efficiency to better repair alkylation damages.

Besides, while we demonstrated that SMYD3 clearly regulates RNF113A E3 ligase activity, it is unclear if RNF113A methylation can modulate its other functions. Recently, RNF113A has been described as an important subunit of the splicing machinery, where its E3 ligase activity is not required. Recently, this second characterized function has been linked to cisplatin resistance in lung adenocarcinoma ²¹. Thus, while Brickner and colleagues demonstrated the specificity of RNF113A-ASCC response to alkylation damages, RNF113A may play a central role in cancer's acquired resistance to multiple drugs, and the impact of RNF113A trimethylation by SMYD3 on chemoresistance, as well as the effect on its role in the spliceosome, are an important issues to explore. Additional experiments are required, especially *in vivo* and in human samples, to assess the impact of this new signaling on tumor response to alkylating agents. Monitoring SMYD3 and RNF113A expression profiles as well as studying the impact of different mutants modulating its methylation and ubiquitination level

will be an important analysis to assess cancer cells resistance to damaging agents through the SMYD3-RNF113a axis.

SCLC model is particularly interesting to study the different consequences of RNF113A K20 trimethylation on DNA damage response, as cisplatin-based chemotherapy is the reference in first line, and the alkylating agent-based therapy using cyclophosphamide was the previous standard treatment^{25,26}. Moreover, etoposide-cisplatin (EP) was not superior to cyclophosphamide-adriamycin-vincristine (CAV) with comparable overall survival and objective tumor response rate, and has mostly been preferred due to better tolerance profile²⁷. Nowadays, CAV is still a second line option in cisplatin-resistant patients with comparable outcomes to other options such as topotecan²⁸. SCLC still represents a major challenge in lung cancer field without recent significant improvement in therapeutics. Better understanding the different pathways involved in chemoresistance to cyclophosphamide and cisplatin may help to improve SCLC patients' treatment through better selection of the optimal therapeutic option.

Besides, new therapeutic approaches combining classical cytotoxic drugs with innovative therapies targeting DNA damage response may be developed based on this field of research and help treatment of many patients as these cisplatin or alkylating agents are used in a broad number of cancers. Finally, further translational researches are needed to study the prognostic value, as well as the predictive value of SMYD3 and RNF113A in the response to treatments.

This work demonstrates a new signaling involving SMYD3 in cancer through the trimethylation of the E3-ubiquitin ligase RNF113A which promotes chemoresistance, a major challenge in cancer treatment. SMYD3 and RNF113A may constitute new predictive markers as well as new therapeutics targets to help the medical decision and optimize cancer patients' care.

REFERENCES

- 1 Hakem, R. DNA-damage repair; the good, the bad, and the ugly. *The EMBO journal* 27, 589-605, doi:10.1038/emboj.2008.15 (2008).
- 2 Helleday, T., Petermann, E., Lundin, C., Hodgson, B. & Sharma, R. A. DNA repair pathways as targets for cancer therapy. *Nature Reviews Cancer* 8, 193-204, doi:10.1038/nrc2342 (2008).
- 3 Li, G. M. Mechanisms and functions of DNA mismatch repair. *Cell research* 18, 85-98, doi:10.1038/cr.2007.115 (2008).
- 4 Kim, Y.-J. & Wilson, D. M., 3rd. Overview of base excision repair biochemistry. *Curr Mol Pharmacol* 5, 3-13, doi:10.2174/1874467211205010003 (2012).
- 5 Kiwerska, K. & Szyfter, K. DNA repair in cancer initiation, progression, and therapy-a double-edged sword. *J Appl Genet* 60, 329-334, doi:10.1007/s13353-019-00516-9 (2019).
- 6 Sakthivel, K. M. & Hariharan, S. Regulatory players of DNA damage repair mechanisms: Role in Cancer Chemoresistance. *Biomed Pharmacother* 93, 1238-1245, doi:10.1016/j.biopha.2017.07.035 (2017).
- 7 Kondo, N., Takahashi, A., Ono, K. & Ohnishi, T. DNA damage induced by alkylating agents and repair pathways. *J Nucleic Acids* 2010, 543531, doi:10.4061/2010/543531 (2010).
- 8 Dango, S. *et al.* DNA unwinding by ASCC3 helicase is coupled to ALKBH3-dependent DNA alkylation repair and cancer cell proliferation. *Mol Cell* 44, 373-384, doi:10.1016/j.molcel.2011.08.039 (2011).
- 9 Brickner, J. R. *et al.* A ubiquitin-dependent signalling axis specific for ALKBH-mediated DNA dealkylation repair. *Nature* 551, 389-393, doi:10.1038/nature24484 (2017).
- 10 Corbett, M. A. *et al.* A novel X-linked trichothiodystrophy associated with a nonsense mutation in RNF113A. *J Med Genet* 52, 269-274, doi:10.1136/jmedgenet-2014-102418 (2015).
- 11 Sims, R. J. & Reinberg, D. Is there a code embedded in proteins that is based on post-translational modifications? *Nature Reviews Molecular Cell Biology* 9, 815-820, doi:10.1038/nrm2502 (2008).
- 12 Hamamoto, R., Saloura, V. & Nakamura, Y. Critical roles of non-histone protein lysine methylation in human tumorigenesis. *Nature reviews. Cancer* 15, 110-124, doi:10.1038/nrc3884 (2015).
- 13 Tschiersch, B. *et al.* The protein encoded by the *Drosophila* position-effect variegation suppressor gene *Su(var)3-9* combines domains of antagonistic regulators of homeotic gene complexes. *The EMBO journal* 13, 3822-3831 (1994).
- 14 Martin, J. L. & McMillan, F. M. SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. *Current Opinion in Structural Biology* 12, 783-793, doi:https://doi.org/10.1016/S0959-440X(02)00391-3 (2002).

- 15 Schubert, H. L., Blumenthal, R. M. & Cheng, X. Many paths to methyltransfer: a chronicle of convergence. *Trends in Biochemical Sciences* 28, 329-335, doi:10.1016/s0968-0004(03)00090-2 (2003).
- 16 Bottino, C., Peserico, A., Simone, C. & Caretti, G. SMYD3: An Oncogenic Driver Targeting Epigenetic Regulation and Signaling Pathways. *Cancers* 12, 142, doi:10.3390/cancers12010142 (2020).
- 17 Yi, X., Jiang, X. J. & Fang, Z. M. Histone methyltransferase SMYD2: ubiquitous regulator of disease. *Clin Epigenetics* 11, 112, doi:10.1186/s13148-019-0711-4 (2019).
- 18 Mazur, P. K. *et al.* SMYD3 links lysine methylation of MAP3K2 to Ras-driven cancer. *Nature* 510, 283-287, doi:10.1038/nature13320 (2014).
- 19 Giakountis, A., Moulos, P., Sarris, M. E., Hatzis, P. & Talianidis, I. Smyd3-associated regulatory pathways in cancer. *Semin Cancer Biol* 42, 70-80, doi:10.1016/j.semcancer.2016.08.008 (2017).
- 20 Gatti da Silva, G. H., Jurica, M. S., Chagas da Cunha, J. P., Oliveira, C. C. & Coltri, P. P. Human RNF113A participates of pre-mRNA splicing in vitro. *J Cell Biochem*, doi:10.1002/jcb.28163 (2018).
- 21 Shostak, K. *et al.* The X-linked trichothiodystrophy-causing gene RNF113A links the spliceosome to cell survival upon DNA damage. *Nat Commun* 11, 1270, doi:10.1038/s41467-020-15003-7 (2020).
- 22 Ueki, Y. *et al.* A Consensus Binding Motif for the PP4 Protein Phosphatase. *Mol Cell* 76, 953-964 e956, doi:10.1016/j.molcel.2019.08.029 (2019).
- 23 Hornbeck, P. V. *et al.* PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res* 43, D512-520, doi:10.1093/nar/gku1267 (2015).
- 24 Komander, D. & Rape, M. The ubiquitin code. *Annu Rev Biochem* 81, 203-229, doi:10.1146/annurev-biochem-060310-170328 (2012).
- 25 Green, R. A., Humphrey, E., Close, H. & Patno, M. E. Alkylating agents in bronchogenic carcinoma. *The American Journal of Medicine* 46, 516-525, doi:https://doi.org/10.1016/0002-9343(69)90071-0 (1969).
- 26 Einhorn, L. H., Fee, W. H., Farber, M. O., Livingston, R. B. & Gottlieb, J. A. Improved Chemotherapy for Small-Cell Undifferentiated Lung Cancer. *JAMA* 235, 1225-1229, doi:10.1001/jama.1976.03260380019019 (1976).
- 27 Roth, B. J. *et al.* Randomized study of cyclophosphamide, doxorubicin, and vincristine versus etoposide and cisplatin versus alternation of these two regimens in extensive small-cell lung cancer: a phase III trial of the Southeastern Cancer Study Group. *Journal of Clinical Oncology* 10, 282-291, doi:10.1200/JCO.1992.10.2.282 (1992).
- 28 von Pawel, J. *et al.* Topotecan Versus Cyclophosphamide, Doxorubicin, and Vincristine for the Treatment of Recurrent Small-Cell Lung Cancer. *Journal of Clinical Oncology* 17, 658-658, doi:10.1200/JCO.1999.17.2.658 (1999).

- 29 DuPage, M., Dooley, A. L. & Jacks, T. Conditional mouse lung cancer models using adenoviral or lentiviral delivery of Cre recombinase. *Nat Protoc* 4, 1064-1072, doi:10.1038/nprot.2009.95 (2009).
- 30 Nakada, S., Chen, G. I., Gingras, A. C. & Durocher, D. PP4 is a gamma H2AX phosphatase required for recovery from the DNA damage checkpoint. *EMBO Rep* 9, 1019-1026, doi:10.1038/embor.2008.162 (2008).

TABLES

Table 1: SMYD3's substrates list obtained by *in vitro* ProtoArray® methylation assay.

Methyltransferases constituting false positive due to automethylation are written in grey and italic. Substrates previously described and characterized in literature are written in bold characters.

Generic name	Protein Identification
ARHGAP17	Rho GTPase activating protein 17
ASZ1	Ankyrin repeat, SAM and basic leucine zipper domain-containing protein 1
C9orf41	Chromosome 9 open reading frame 41
CLUAP1	Clusterin associated protein 1
<i>COMT</i>	<i>Catechol-O-methyltransferase</i>
COPS5	COP9 constitutive photomorphogenic homolog subunit 5 (Arabidopsis)
<i>FAM119A</i>	<i>Family with sequence similarity 119, member A</i>
<i>LCMT1</i>	<i>Leucine carboxyl methyltransferase 1</i>
LMNA	Lamin A/C
MAP3K2	Mitogen-activated protein kinase kinase kinase 2
NOC2L	Nucleolar complex associated 2 homolog (S. cerevisiae)
NPM1	Nucleophosmin
PCLO	Piccolo presynaptic cytomatrix protein
<i>PCMT1</i>	<i>Protein-L-isoaspartate(D-aspartate) O-methyltransferase</i>
PIN4	Peptidyl-prolyl cis-trans isomerase NIMA-interacting 4
<i>PRMT1</i>	<i>Protein arginine N-methyltransferase 1</i>
<i>PRMT6</i>	<i>Protein arginine N-methyltransferase 6</i>
<i>PRMT8</i>	<i>Protein arginine N-methyltransferase 8</i>
RIOK1	RIO kinase 1 (yeast)
RNF113A	Ring finger protein 113A
<i>SMYD2</i>	<i>SET and MYND domain-containing protein 2</i>
SP100	SP100 nuclear antigen

FIGURES

Figure 1. SMYD3-dependent SCLC progression.

(A) Comparison of SCLC development from TKO mice harboring either wild-type or Smyd3^{-/-} knock out. Tumors were monitored using hematoxylin and eosin (HE) and cells proliferation/apoptosis were assessed by phosphor-Histone H3 (pH3) and cleaved Caspase 3 (cl. Caspase3) markers respectively. (B) Representative micro computed tomography images of 20 weeks old mice lungs. the endpoint of survival studies on upper quadrants. (C) Quantification of tumor area per lung, pH3 positive cells and cleaved caspase-3 positive cells per tumor area in representative SCLC samples from TKO and TKO Smyd3^{-/-} mice. (D) Kaplan-Meier survival analysis of TKO (n=12, median survival=181 days) and TKO;Smyd3 (n=12, median survival=266 days) mutant mice, time post infection. P=0.0001 by log-rank test for significance. For all statistical analyses, p value represents: * <0.01; **<0.001; ***<0.0001.

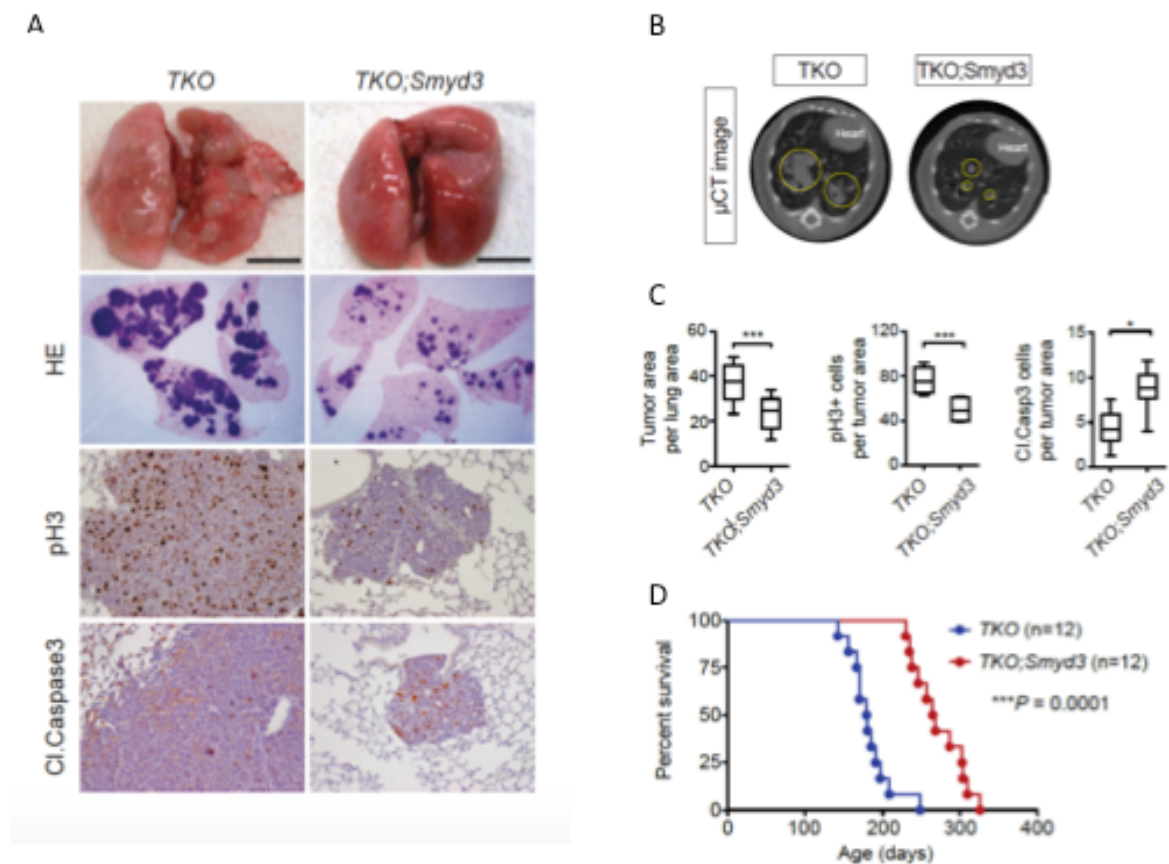


Figure 2. Identification of RNF113A as a new substrate of SMYD3.

(A) Methylation assay using radiolabeled SAM and recombinant SMYD3 with candidates identified by ProtoArray® with MAP3K2 as substrate control. Methylation events are detected by autoradiography and Coomassie panel represents loading control. (B) Identification of RNF113A K20me3 by MS/MS analysis of RNF113A methylated *in vitro* by SMYD3. Note that Deuterated S-adenosyl-methionine was used as methyl donor. (C) *In vitro* methylation assay with recombinant SMYD3 and RNF113A WT, K20A or K21A mutants. Methylation events are detected by autoradiography and Coomassie panel represents loading control. (D) Detection of endogenous RNF113A methylation by SMYD3 in cancer cells, after immunoprecipitation of total RNF113A and immunodetection with RNF113AK20me3 antibody.

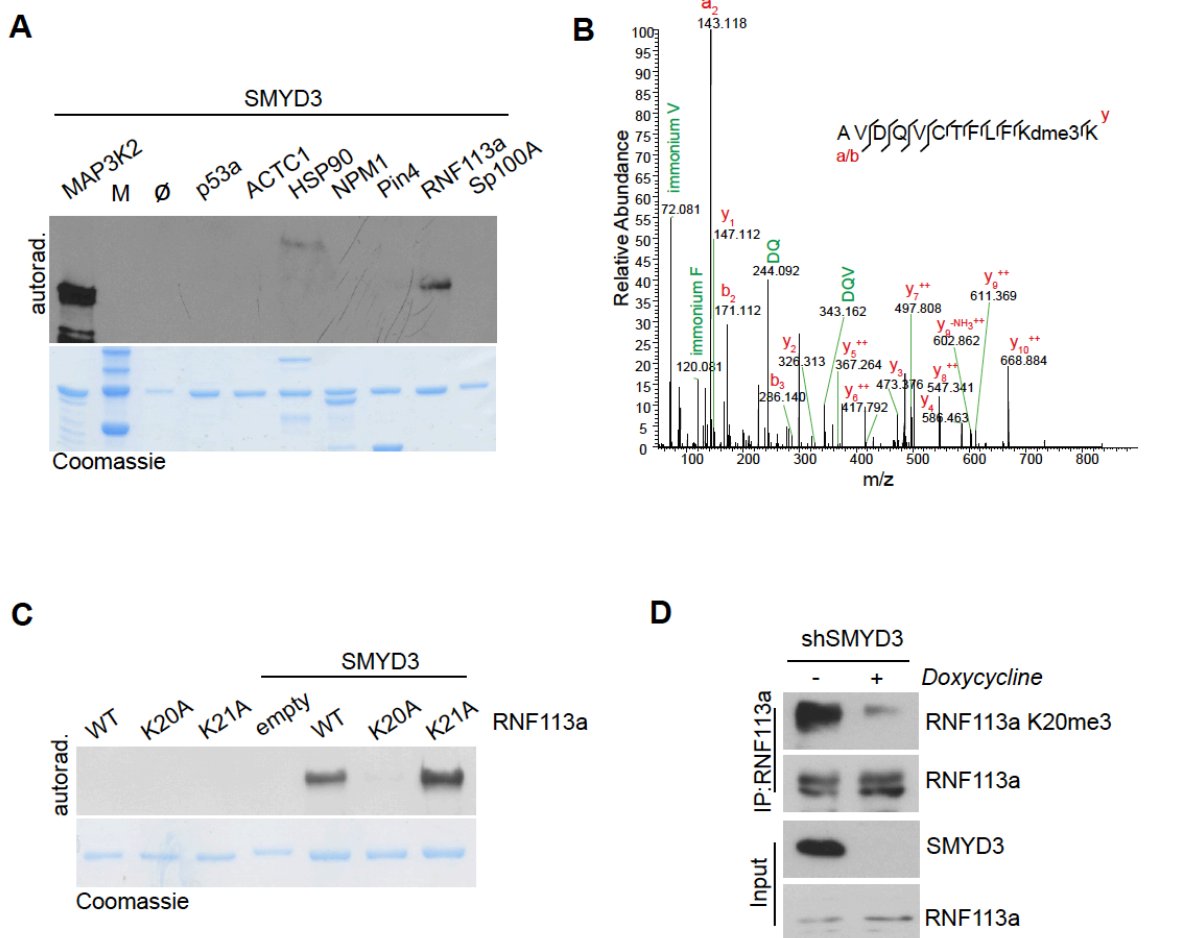


Figure 3. SMYD3 methylation of RNF113A repels the phosphatase PP4.

A) Scatter plot representation of SILAC peptide pulldown quantitative proteomics analysis of RNF113A K20me0 and/or RNF113A K20me3 peptides interactors. Forward H/L experiment designates the relative quantification of Heavy labeled proteins bound to RNF113A K20me3 peptide versus Light labeled proteins bound to RNF113A K20me0 peptide. Reverse H/L experiment represents a swap of labeled proteins with the peptides. Specific interactors of RNF113A K20me0 reside in the lower left quadrant. The four PP4A complex subunits are highlighted in red. (B) Peptide pulldown assay with recombinant GST-tagged PP4 subunits. (C) Peptide pulldown assay with detection of endogenous PPP4R3a in DMS114 SCLC cell line extracts. (D) Co-immunoprecipitation of endogenous PPP4R3a from 293T cells with ectopic wild-type, K20 or K20F mutants HA-RNF113A

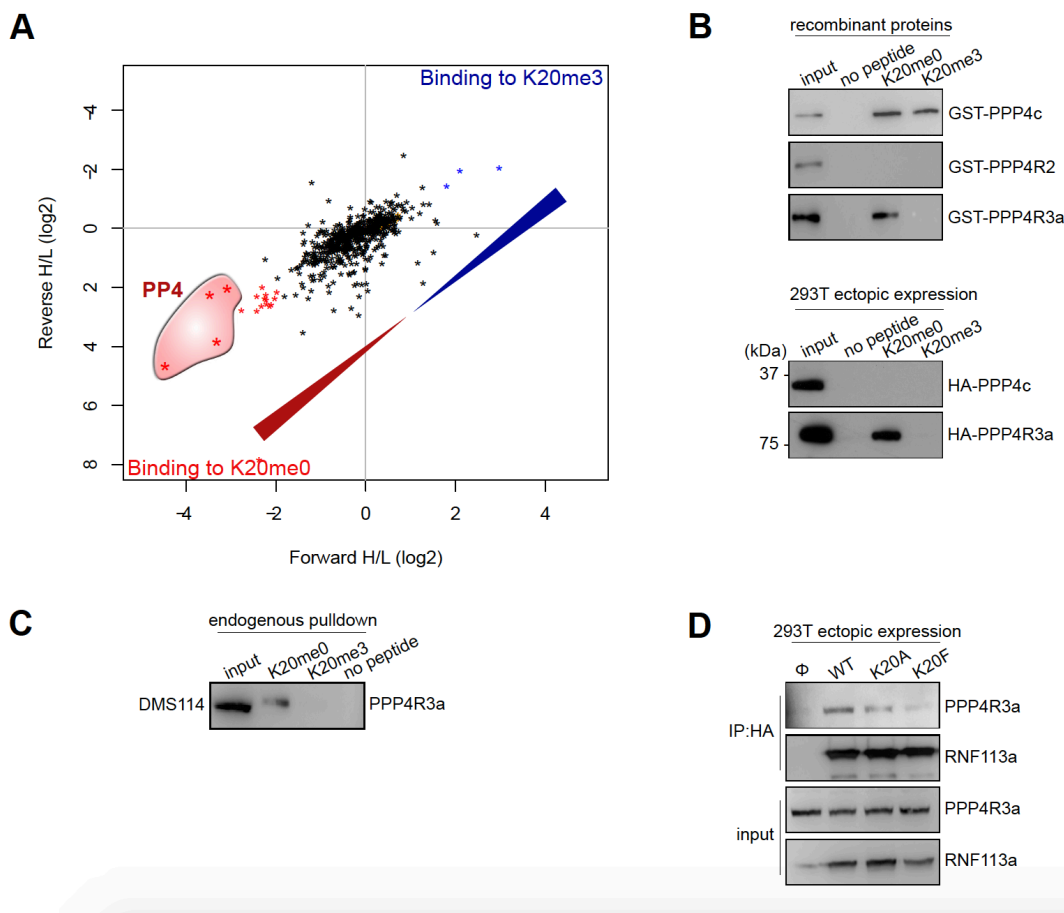


Figure 4. RNF113A phosphorylation is required for its activity.

(A) Immunodetection of RNF113A with or without dephosphorylation of whole cell extract by λ phosphatase or FastAP thermosensitive alkaline phosphatase. (B) Monitoring of RNF113A ubiquitination by His-Ub enrichment in 293T cells after treatments with indicated damaging agents. (C) *In vitro* E3 ubiquitin ligase activity of purified RNF113A is detected as smears of poly-Ub chains, with or without cells treatment with MMS and dephosphorylation step of protein extract before RNF113A purification. Bottom panel presents quantification of RNF113A WT activity upon MMS treatment and dephosphorylation for several biological replicates. (D) Stable HA-RNF113A engineered HeLa whole cell extract with or without MMS treatment was subjected to dephosphorylation assays with FastAP or purified PP4 phosphatase, and total and phospho-RNF113A were monitored by immunodetection after HA-RNF113A immunoprecipitation.

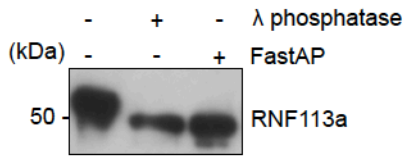
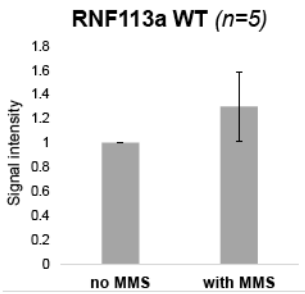
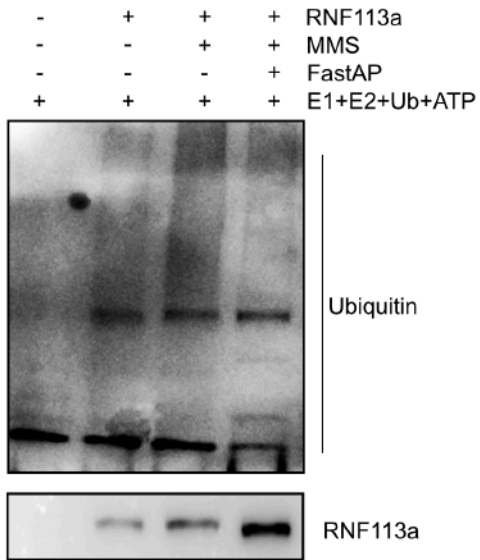
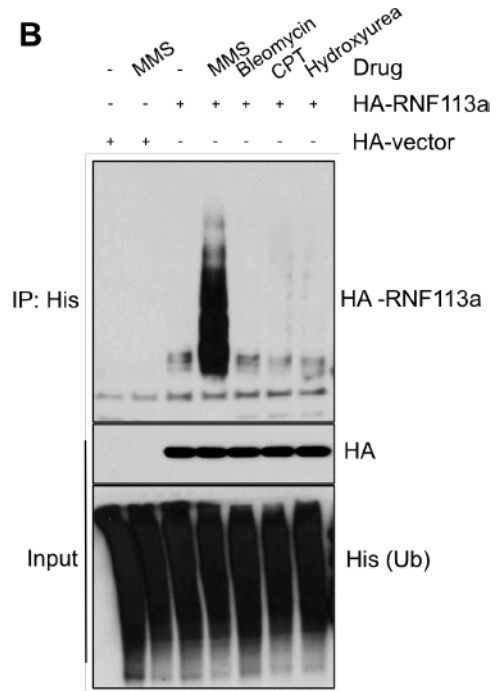
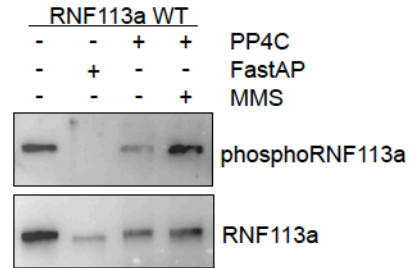
A**C****B****D**

Figure 5. RNF113A transient phosphorylation is regulated by SMYD3.

(A) Detection of auto-ubiquitinated RNF113A after TUBE pulldowns from HeLa stably expressing RNF113A WT or RNF113A Δ RING, with or without DNA alkylation damages induced by MMS. Control of DNA damage was monitored using immunodetection of phosphoH2AX. (B) Detection of auto-ubiquitinated RNF113A after TUBE pulldowns from HeLa stably expressing RNF113A WT or RNF113A K20F mutant, with or without DNA alkylation damages induced by MMS. (C) *In vitro* poly-Ub chain assay performed with purified RNF113A from HeLa cells after different MMS treatment durations (D) Immunodetection of total and modified RNF113A (phosphorylation and methylation) immunoprecipitated from HeLa cells stably expressing RNF113A WT, after different MMS treatment durations. (E) Immunodetection of RNF113A phosphorylation level using immunoprecipitated RNF113A WT or K20F mutant from engineered HeLa cells with or without MMS treatment. (F) E3 ligase activity detected by poly-Ub chain assays using purified RNF113A WT or K20F from engineered HeLa cells, with or without MMS treatment and/or dephosphorylation step by FastAP.

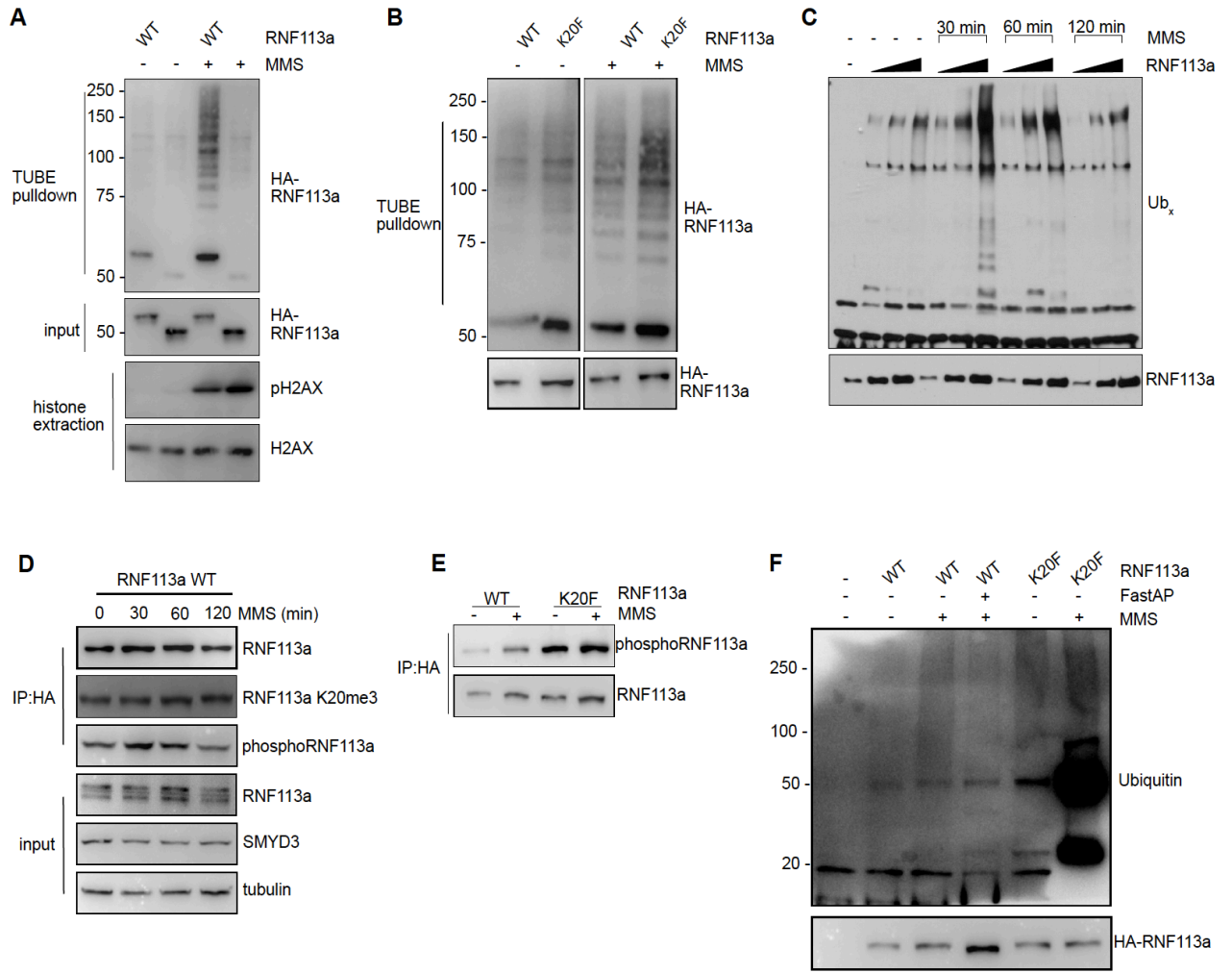


Figure 6. SMYD3 modulates RNF113A-dependant alkylation damage response and cell resistance to alkylating chemotherapy. Cell sensitivity comparison of SCLC DMS-114 cells overexpressing SMYD3 and RNF113A after treatment with either two alkylating agents (A) MMS and (B) Carmustine.

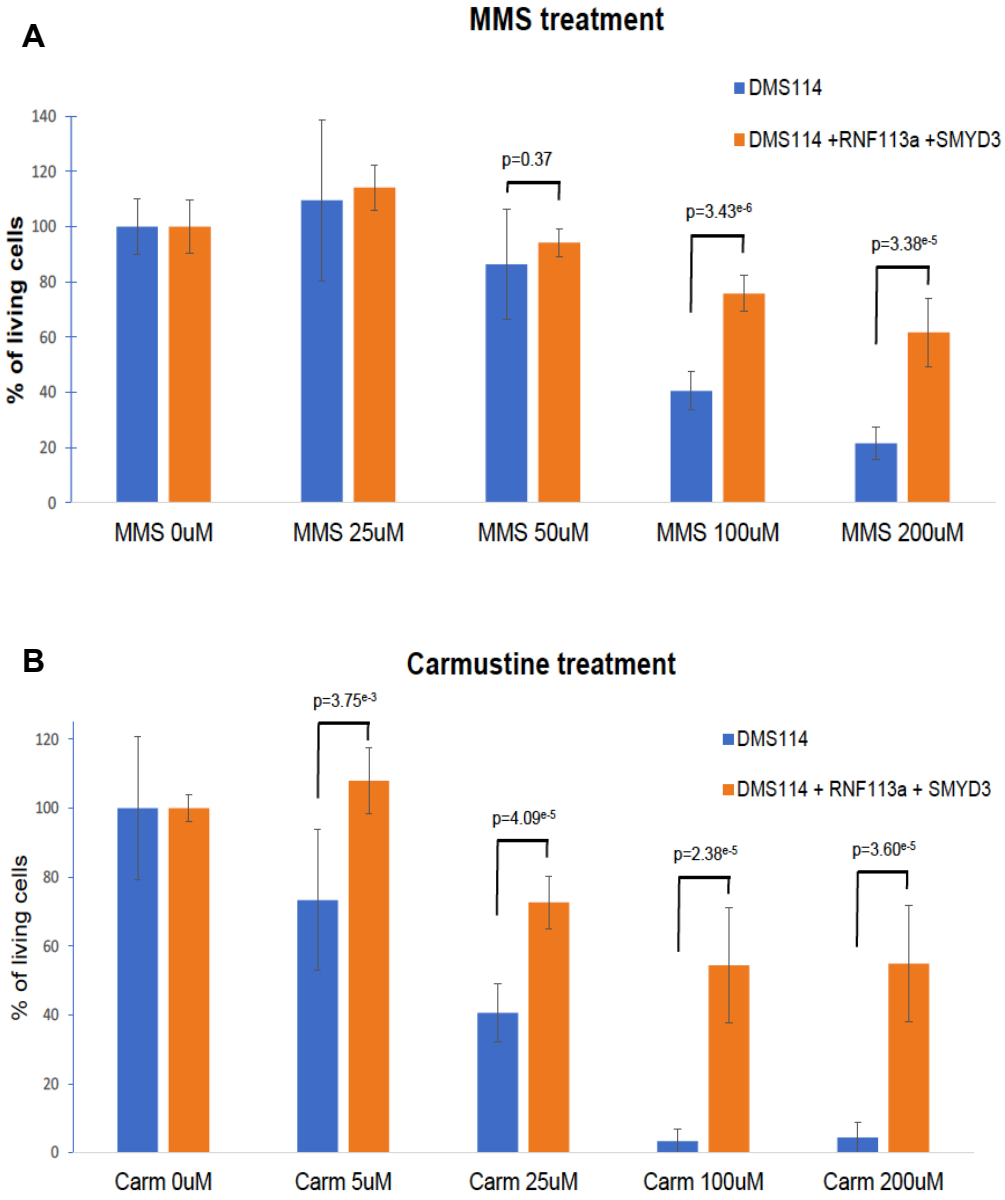
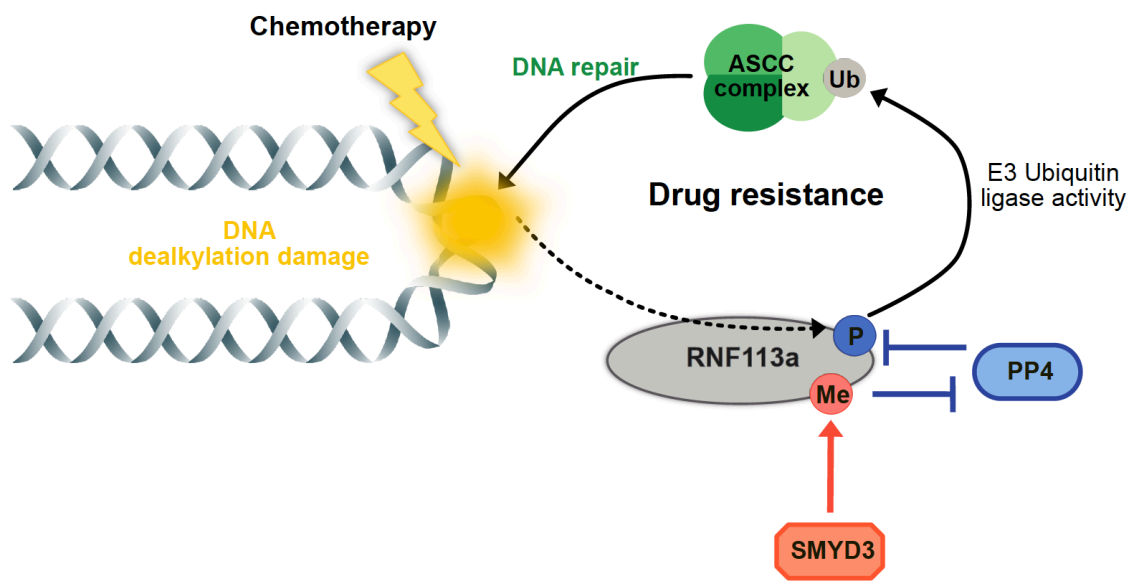


Figure 7. Model of SMYD3/PP4 crosstalk signaling that regulates RNF113A activity in DNA alkylating damage response and repair.



MATERIAL AND METHOD

Lung cancer mouse model.

Mice harboring conditional knock out Rb1, Rbl2 and p53 or Rb1, Rbl2, p53 and Smyd3 were treated at 8 weeks with adenovirus expressing Cre by intratracheal infection as previously described²⁹ to obtain Rb^{-/-};Rbl2^{-/-};p53^{-/-} (TKO) or Rb1^{-/-};Rbl2^{-/-};p53^{-/-};Smyd3^{-/-} (TKO;smyd3) mice. Tumours were analyzed and quantified at 20 weeks post-infection by micro computed tomography (μ CT) and tumour burden was analysed 26 weeks after lentiviral infection.

Protein array.

ProtoArray[®] (Invitrogen, version 5.0) was incubated overnight with the tritium radiolabeled cofactor ³H-SAM and either recombinant GST-tagged SMYD3 or GST alone. Methylation was then revealed by autoradiography.

Methylation assay.

In vitro methylation assays were completed by using 1 to 2 μ g of recombinant proteins or peptides which were incubated with 1 μ g of recombinant SMYD3 and 0.1 mM S-adenosyl-methionine (SAM, Sigma) or 2 μ Ci SAM[³H] (American Radiolabeled Chemicals) in buffer containing 250 mM Tris-HCl (pH 8.0), 50% glycerol, 100 mM KCl, 25 mM MgCl₂ at 30°C overnight. The reaction mixture was resolved by SDS-PAGE, followed by autoradiography, Coomassie stain or mass spectrometry analysis.

Mass spectrometry analyses.

Non-radioactive *in vitro* methylation assays using deuterated S-adenosyl-methionine were analyzed by in-gel tryptic digestion followed by LC-MS/MS using an OrbiTrap VELOS pro. Results were analyzed by with MaxQuant version 1.2.2.5 software (Cox and Mann, 2008).

Peptide pull-down and SILAC.

Streptavidin Sepharose beads (5 – 10 μ l; GE Healthcare) were saturated with 7.5 μ g of specific biotinylated peptides for 2 h at 4 °C under rotation in peptide buffer 50 mM Tris (pH 8), 150 mM NaCl, 0.5 % NP40, 0.5 mM DTT, 10 % glycerol, anti-proteases, and beads were afterwards washed in the same buffer. Beads were then incubated with either 1 μ g of recombinant proteins or 1 mg of cytoplasmic cell extract for 4 h at 4 °C under rotation in peptide buffer. Beads were then washed 3 times in the same buffer and resuspended in Laemmli buffer for immunoblot analysis.

For SILAC peptide pull-down, cytoplasmic extracts were prepared from cells cultivated in culture media containing either Lysines K0 and Arginines R0 ('light', with regular C12 and N14 isotopes) or Lysines K6 and Arginines R10 ('heavy', with C13 and N15 isotopes). A 2-way experiments was performed, the 'forward' condition combining RNF113A-K20me0 peptide with light extract and RNF113A-K20me3 peptide with heavy extract, the 'reverse' condition combining RNF113A-K20me3 peptide with light extract and RNF113A-K20me0 peptide with heavy extract. Beads of each pair of peptide pull-down were then pooled together, washed and extracts were resuspended in Laemmli buffer and resolved by SDS-PAGE. In-gel trypsin digestion was performed and peptides purified using C18 stage tips (Fisher) before mass spectrometry analysis to quantify the ratio of each potential binders for K20me0 and K20me3 peptides, in forward and reverse condition.

Expression and purification of recombinant proteins.

Recombinant proteins were purified from transformed BL21 pLys cells. Cells were resuspended in lysis buffer containing 50 mM Tris pH 7.5, 150 mM NaCl, 0.05 % NP-40, 0.25 mg mL⁻¹ lysozyme, 0.5 mM PMSF and protease inhibitors, and additionally lysed by sonication. GST-tagged proteins were purified using Glutathione Sepharose® 4B beads (GE Healthcare) and either eluted in 10 mM reduced L-glutathione (Sigma) or cleaved from the GST tag using purified Precision enzyme.

Cell culture, transfections and cell viability assays.

HeLa and 293T cells were grown in Dulbecco's modified Eagle's medium (GIBCO) supplemented with 10% fetal bovine serum (Dutcher), 100 units ml⁻¹ penicillin/streptomycin. DMS114 were cultured in RPMI medium (GIBCO) supplemented with 10% fetal bovine serum (Dutcher), 100 units ml⁻¹ penicillin/streptomycin. All cells were cultured at 37°C in a humidified incubator with 5% CO₂ and regularly tested for absence of mycoplasma contamination (Lonza kit).

For transient expression cells were transfected with TransIT®-LT1 transfection reagent (Mirus) and collected 36 h after transfection. For stable expression, cells were transfected with retroviral pMSCV (HA/Flag-tagged RNF113A WT, K20F and ΔRING) and pCMV (SMYD3) constructs using packaging pVSVg and pUMCV vectors, followed respectively by 5 µg ml⁻¹ blasticidin or 400 µg ml⁻¹ neomycin selection for one week. For inducible knockdown, cells were transfected with lentiviral shRNA constructs using the packaging vectors pVSVg and pΔ8.2, followed by 2 µg ml⁻¹ puromycin selection for one week. For survival assay of DNA damaging agents, DMS-114 cells were cultured overnight in 96-well plate in 100 µl media. Cells were then exposed to medium containing the indicated concentration of MMS (Sigma) or Carmustine (Sigma) for 24 h at 37°C. The media was then replaced with normal media and cell viability was assessed using the PrestoBlue® assay (Thermo Fischer Scientific) 48 hours later.

Ubiquitin ligase assays.

Ubiquitination ligase assays were completed in reaction buffer 25 mM Tris-HCl pH 7.6, 5 mM MgCl₂, 100 mM NaCl, 1 mM DTT containing 2 mM ATP and 10 µM ubiquitin in final volume 30 µl. Ubiquitin activating enzyme 1 (UBE1; Boston Biochemicals) was used at 0.6 µM and Ubiquitin-conjugating Enzyme (UBCH5c /Ube2D3; Boston Biochemicals) 0.5 -1 µM. HA-Flag-tagged-RNF113A WT and K20F were purified from HeLa and added to each reaction mixture. The reaction mixtures were incubated at 37°C for 2 h. Reactions were stopped by addition of Laemmli buffer and analyzed by Western blot.

TUBE pulldown.

HA-Flag-tagged RNF113A WT, K20F and Δ RING were stably expressed in HeLa cells. For TUBE pulldown DNA alkylation damage was stimulated with MMS treatment for indicated amount of time (usually 4 h if not indicated). Cells are collected and lysed in TUBE lysis buffer (50 mM Tris-HCl, pH 7.5, 1 mM EGTA, 1 mM EDTA, 1% (vol/vol) Triton X-100, 0.27 M Sucrose, 0.2 mM PMSF, 100 mM iodoacetamide, protease and phosphatase inhibitors) for 60 min at 4 °C. Prior to this step, the efficiency of DNA damage was validated by collecting some of the cell pellet and analyze of pH2AX signal after histone extraction. Following cell lysis and centrifugation, whole cell lysate was added to TUBE2 beads (Recombinant Human Ubiquitin 1 Tandem UBA Agarose, R&D Systems). Cell lysate and the beads were incubated overnight at 4 °C. Beads were extensively washed (3 times) in high salt TAP buffer (50 mM Tris-HCl, pH 7.9, 300 mM KCl, 5m M MgCl₂, 0.2 mM EDTA, 0.1% NP-40, 10% glycerol, 2 mM mercaptoethanol, 0.2 mM PMSF) and twice in low salt TAP buffer (50 mM Tris-HCl, pH 7.9, 0 mM KCl, 5 mM MgCl₂, 0.2 mM EDTA, 0.1% NP-40, 10% glycerol, 2 mM mercaptoethanol, 0.2 mM PMSF). Proteins were eluted in laemmli buffer and TUBE pulldowns were analyzed by western blot.

Co-immunoprecipitation with recombinant proteins.

Recombinant proteins (PPP4R3a and RNF113A mutants) were purified as described above. More specifically, PPP4R3a was cleaved from its GST tag by Precision enzyme and dialyzed. GST-tagged RNF113A mutants were bound to Glutathione Sepharose beads for minimum 1 h and PPP4R3a was added and samples were incubated overnight at 4 °C. After three washes in wash buffer, proteins were eluted in Laemmli buffer. Dialysis and washes were completed in buffer containing 150 mM NaCl, 50 mM Tris pH 8, 0.1% NP40, 0,5 mM EDTA, 0.5 mM DTT. Direct interaction between proteins was analyzed by immunoblotting.

Co-immunoprecipitation after ectopic expression of proteins.

Immunoprecipitation of HA-tagged RNF113A WT, K20A and K20F was completed after transient expression in 293T cells for 36 h. The cells were resuspended in lysis buffer (50

mM Tris-HCl pH 8, 137 mM NaCl, 1 mM EDTA, 0.5 % NP40, 10% glycerol, anti-proteases and phosphatase inhibitors, 1 mM PMSF). The cell lysate was cleared by centrifugation and added to previously washed anti-HA resin in the same buffer. After overnight incubation at 4 °C with rotation, HA-resin with bound proteins was washed three times in the same buffer. Proteins were eluted with laemmli buffer and analyzed by western blot.

Immunoprecipitation after denaturation.

293T cells were transfected with His-tagged ubiquitin and HA-tagged RNF113A. Cells were treated for 1 h with 1 mM MMS, and media was then replaced with fresh media with proteasome inhibitor MG-132 4 h before collecting the cells. Part of the cells was collected before addition of the lysis buffer to analyze the protein expression level (input). Cells were resuspended in lysis buffer (6 M guanidinium-HCl, 0.1 M Na₂HPO₄/NaH₂PO₄, 10 mM Tris-HCl, pH 8, 0.005 M imidazole, and 0.01 M β-mercaptoethanol) and lysed by sonication. After centrifugation, cell lysate was added to previously equilibrated Ni₂₊-NTA beads agarose beads. Cell lysate was incubated with beads overnight at 4 °C with rotation. Beads with bound proteins were extensively washed as follows: once in lysis buffer, once in wash buffer (8 M urea, 0.1 M Na₂HPO₄/NaH₂PO₄, 10 mM Tris-HCl (pH 6.8), 0.005 M imidazole, 0.01 M β-mercaptoethanol) and twice in wash buffer with 0.1% Triton X-100. Proteins were eluted for 45 min at RT on a shaker with elution buffer (0.2 M imidazole, 0.15 M Tris-HCl (pH 6.8), 30% glycerol, 0.72 M β-mercaptoethanol, 5% SDS). Laemmli buffer was added and immunoprecipitation was analyzed by western blot.

Protein purification for in vitro ubiquitination assay.

RNF113A WT or variants used in ubiquitination ligase assays were purified from HeLa stably expressing Flag/HA-tagged RNF113A. Cell pellet is resuspended in Flag lysis buffer (50mM Tris-HCl (pH 8.0), 137mM NaCl, 1mM EDTA, 1% Triton X 100, 10% Glycerol, 0.1 mM PMSF, protease and phosphatase inhibitors) and lysed for 30 min at 4°C. After centrifugation clear cell lysate was added to Anti-FLAG-M2 affinity gel (SIGMA Aldrich). Binding of the protein

was allowed for 4 hours or overnight at 4°C. Afterwards, beads were extensively washed with cold wash buffer (50 mM Tris-HCl (pH 8.0), 450 mM NaCl, 1 mM EDTA, 1% Triton X 100, 10% Glycerol, 0.1 mM PMSF), Flag lysis buffer and TAP buffer (50 mM Tris (pH 7.9), 100 mM KCl, 5 mM MgCl₂, 0.2 mM EDTA, 0.1% NP-40, 10% glycerol, 0.2 mM PMSF). Elution was completed for 4 hours or overnight with TAP buffer and 0.5 mg ml⁻¹ 3x Flag peptide (SIGMA Aldrich) in desired volume.

Purification of PPP4 from mammalian cells.

Purification of PP4 catalytic subunits was previously described³⁰. 293T cells were transfected with HA-tagged PPP4c subunit. Transfected cells were lysed in immunoprecipitation buffer (0.5% NP-40, 50 mM HEPES-KOH (pH 8.0), 100 mM KCl, 2 mM EDTA, 10% glycerol, 1 mM PMSF, 1 mM DTT, protease and phosphatase inhibitors). After centrifugation, cleared cell lysate was incubated with HA-resin for 5 h at 4°C with rotation. Beads with bound proteins were washed twice in immunoprecipitation buffer and twice in wash buffer (50 mM Tris-Cl (pH 7.2), 100 mM NaCl, 0.4 mM EDTA, 1 mM DTT). Phosphatase subunits were eluted by three elutions, each 1 h at 4°C in wash buffer with 0.5 mg ml⁻¹ HA-peptide. Final elutions were supplemented with glycerol to final concentration of 50 %, snap-frozen and kept at -80 °C.

In vitro dephosphorylation assay.

In vitro dephosphorylation of RNF113A using FastAP Thermosensitive Alkaline Phosphatase (ThermoFisher) was completed in reaction mixture containing FastAP reaction buffer, 10 U of FastAP phosphatase on whole cell lysate or immunoprecipitated proteins. Reaction mixture was incubated for 1 h at 37°C. For *in vitro* dephosphorylation of RNF113A by PP4 phosphatase, purified PPP4c was first incubated in reaction buffer (50 mM Tris-Cl (pH 7.2), 0.1 mM CaCl₂, 5 mM MnCl₂ and 0.2 mg ml⁻¹ BSA) for 10 min at 30°C. This reaction mixture was then added to RNF113A WT and K20F purified as described above (Protein purification for *in vitro* ubiquitination assay) and phosphatase reactions were incubated at 30°C for 1 h. Changes in phosphorylation level was analyzed by western blotting.

Immunoblot analysis.

Proteins were resolved by SDS-PAGE, transferred to PVDF membrane and analyzed by immunoblot. Antibodies used: RNF113A (HPA000160, Atlas Antibodies); SMYD3 (home-made) ; streptavidin / peroxidase conjugated anti-biotin (200-032-211, Jackson ImmunoResearch); PPP4R3 (A300-840A, Bethyl Laboratories) ; GST (MA000031M0m, CusAbio technology); HA (C29F4, Cell Signaling Technology) ; Ubiquitin (3933, Cell Signaling Technology) ; His (MA000011M0m, CusAb) ; pH2AX (05-636 , Milipore) ; H2AX (A700-059-T , Bethyl Laboratories) ; tubulin (sc-5286 SantaCruz); RNF113A K20me3 (home-made) and Phospho-CDK Substrate Motif [(K/H)pSP] (Cell Signaling Technology).

SUPPLEMENTARY FIGURES

Figure S1.

(A) Schematics of the inducible triple knock out $Rb^{-/-}Rbl2^{-/-};p53^{-/-}$ (TKO) Small Cell Lung Cancer mouse model with putative *Smyd3* knock out. (B) Mouse protocol for tumor development and follow-up analyses.

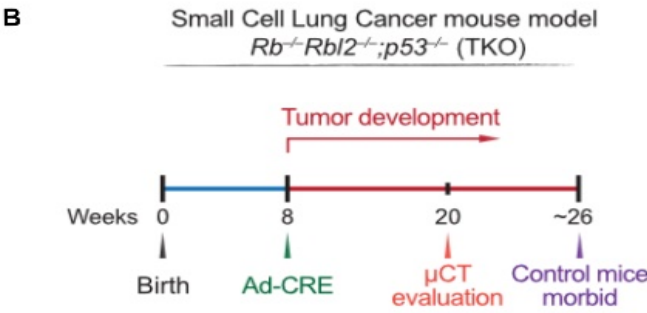
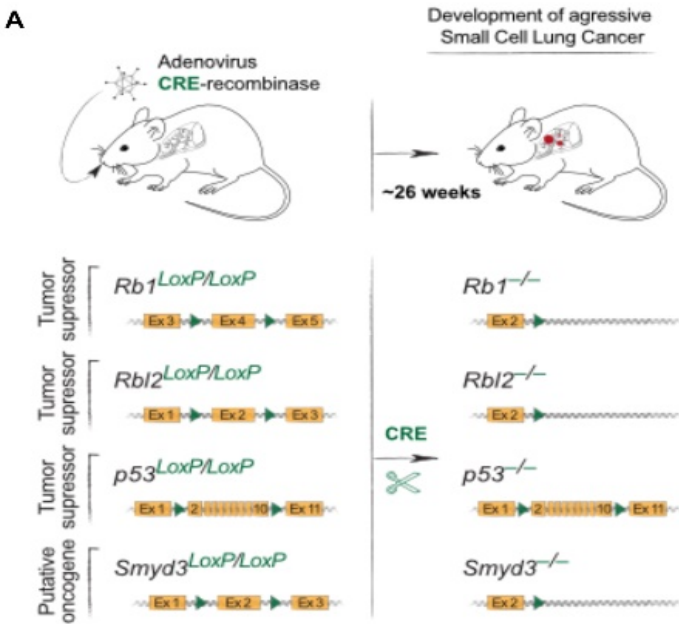


Figure S2.

(A) Protein array methylation assay with recombinant SMYD3 and radiolabeled S-adenosyl-methionine. Magnification depicts the signals identified in square 43 of the Protoarray®, corresponding to the detailed spotted proteins and controls. (B) Identification of RNF113A K20me1 (top panel) and RNF113A K20me2 (bottom panel) by MS/MS analysis of RNF113A methylated *in vitro* by SMYD3. Note that Deuterated S-adenosyl-methionine was used as methyl donor. (C) RNF113A K20me3 antibody's specificity tested by dotblot using RNF113A K20 peptides with different methyl states and RNF113A K21me3 and MAP3K2 K260me3 peptides. Streptavidin panel is a loading control for the different peptides. (D) *In vitro* methylation assay using RNF113A K20 peptides with different methyl states and recombinant SMYD3, detected using RNF113A K0me3 antibody on dotblot. Streptavidin panel is a loading control for the different peptides. (E) Detection of RNF113A methylation in 293T cells using RNF113A K20me3 antibody after ectopic expression of SMYD3 and RNF113A WT or K20A mutant.

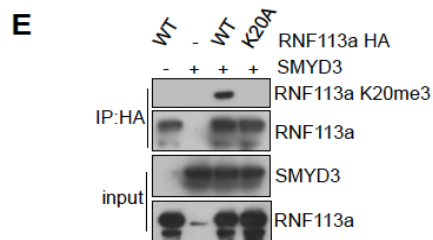
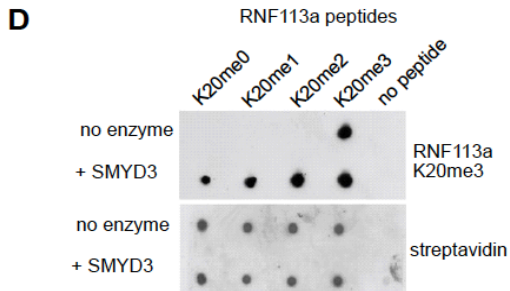
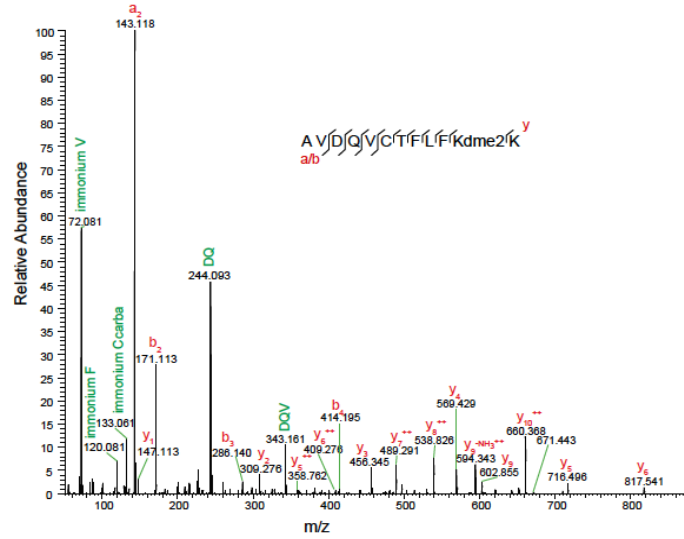
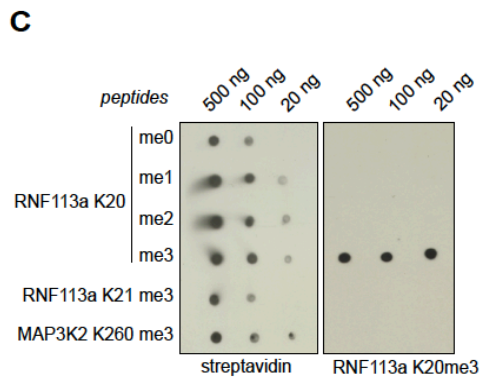
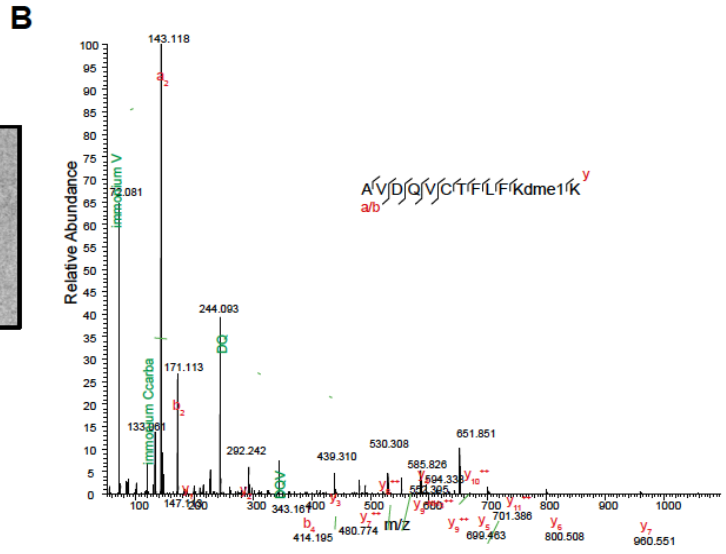
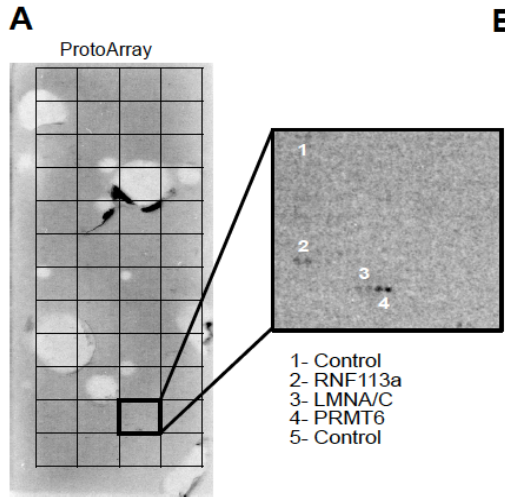


Figure S3.

(A) Magnification of the SILAC-based quantitative proteomic analysis presented in Figure 3A, with identification of major hits. (B) Model depicting the post-translational modifications crosstalk of RNF113A induced by the phosphatase complex PP4 and the lysine methyltransferase SMYD3. (C) Peptide pulldown as in Figure 3C, with intermediate mono- and dimethylation state of RNF113A K20. (D) Pulldown assay using recombinant GST-RNF113A wild-type or K20A and K20F mutants as bait for PPP4R3a.

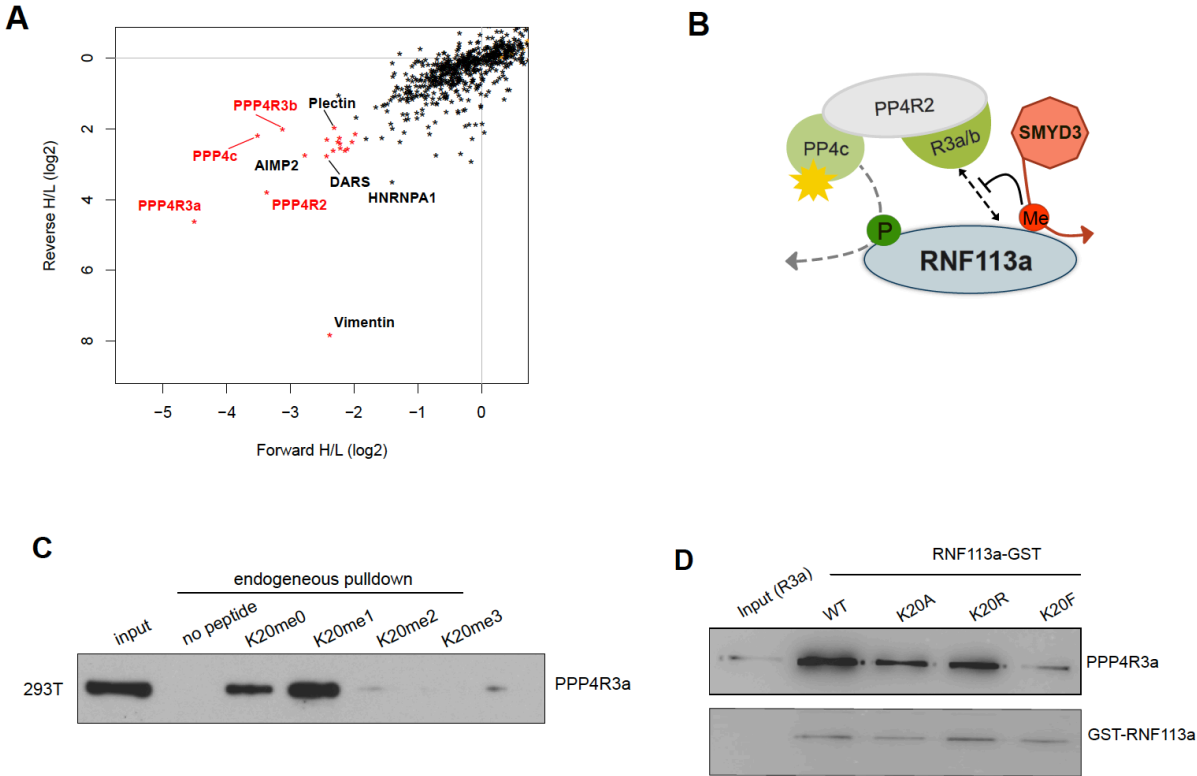


Figure S4.

(A) Immunodetection of RNF113A with or without prior CIP treatment of whole cell extract, obtained after Phos-Tag SDS-PAGE electrophoresis. (B) MS/MS identification of multiple phosphorylated residues on RNF113A purified from HeLa cells.

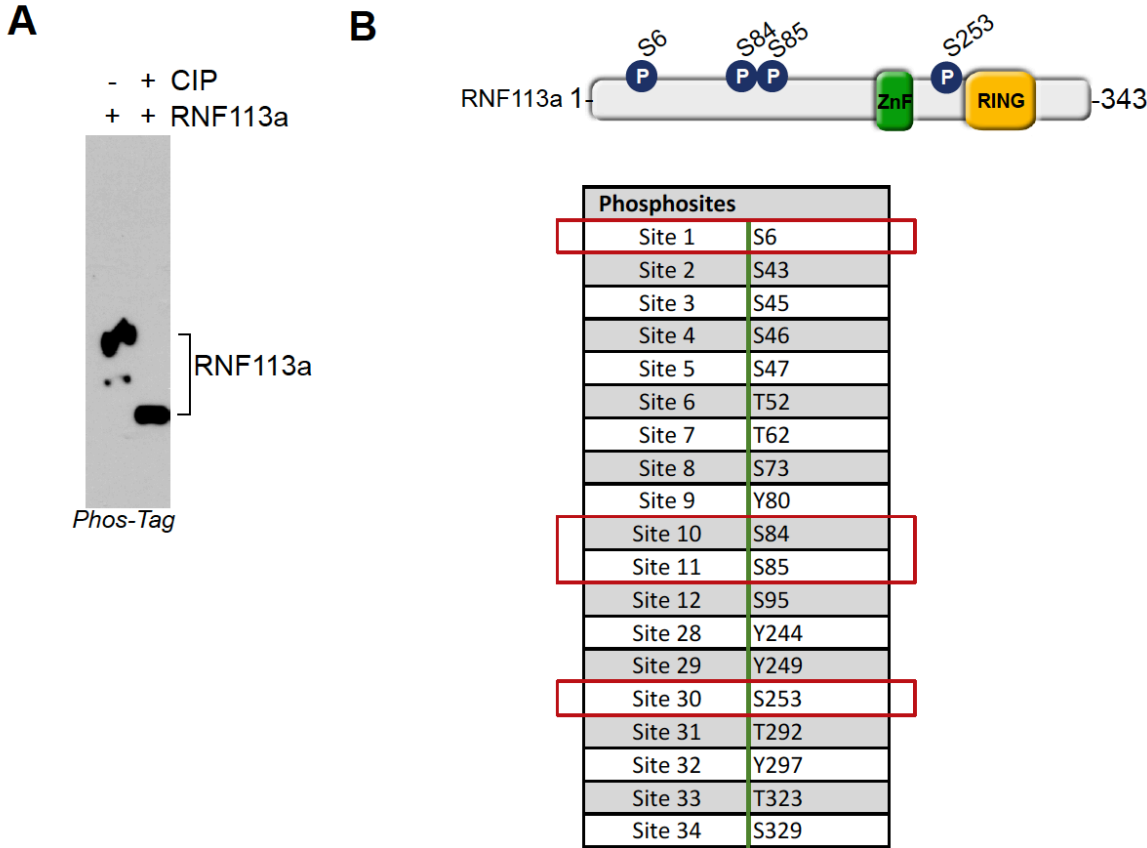
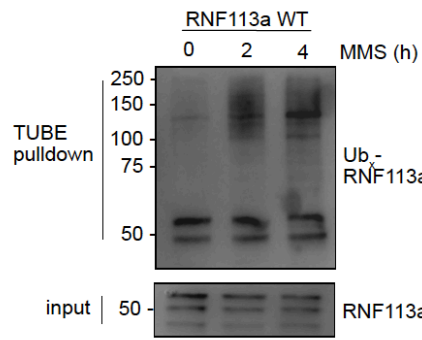
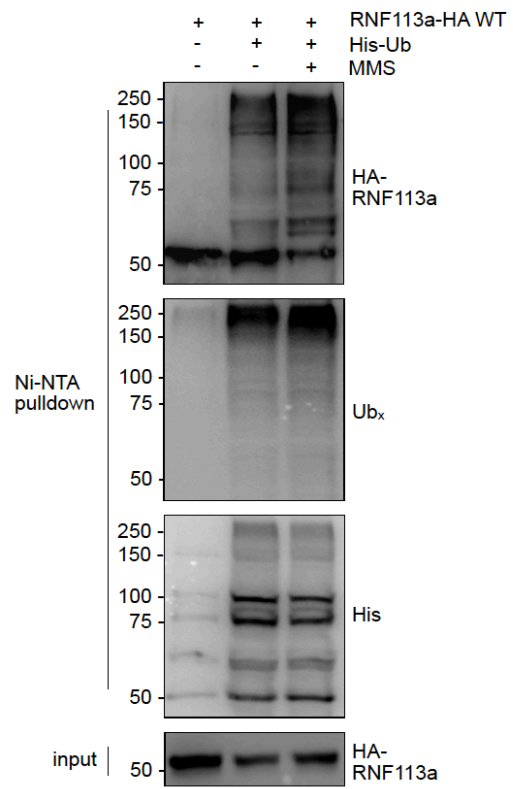
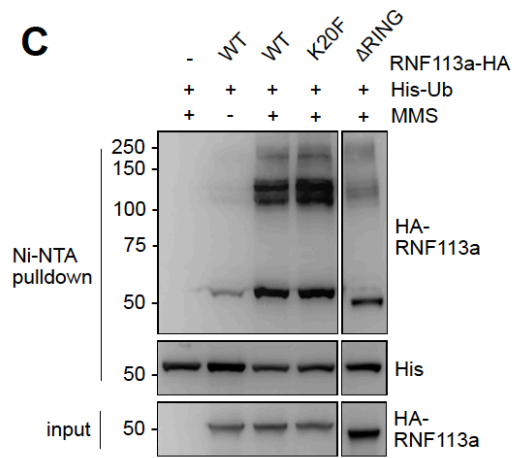
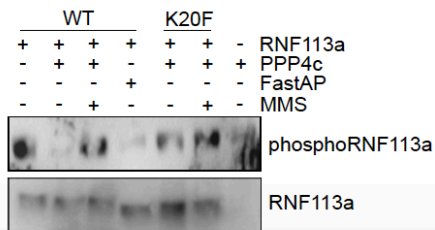


Figure S5.

(A) Detection of RNF113A ubiquitination after 0, 2h and 4h of MMS treatment, by TUBE pulldown of ubiquitinated proteins from HeLa cells engineered to stably express RNF113A WT (B) Detection of RNF113A ubiquitination with or without MMS treatment, after Ni-NTA pulldown from 293T cells expressing ectopic RNF113A WT and His-tagged Ubiquitin. (C) Detection of RNF113A auto-ubiquitination after similar Ni-NTA pulldown as in B, using 293T cells expressing ectopic either RNF113A WT, K20F or DRING mutants. (D) Monitoring of RNF113A phosphorylation level after immunoprecipitation of total RNF113A from HeLa cells stably expressing RNF113A WT or K20F mutant. Cells were treated or not with MMS, and whole cell extract was treated or not with either FastAP and PPP4c purified from 293T cells.

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

DISCUSSION GÉNÉRALE

Les lysines méthyltransférases SMYD2 et SMYD3 sont impliquées dans de nombreux processus de carcinogenèse. Leur surexpression est fréquente dans la plupart des types de cancer où elle a souvent une valeur pronostique. Les travaux récents impliquant SMYD2 et SMYD3 dans le contrôle de voies signalétiques p53 ou RAS démontrent que ces enzymes sont des effecteurs essentiels de processus emblématiques qui constituent les « signatures moléculaires » (ou *hallmarks*) du cancer ^{1,2}. Néanmoins, le faible nombre de substrats caractérisés à l'heure actuelle, notamment pour SMYD3, fait suspecter que de nombreuses signalisations restent à découvrir. Par ailleurs, la majorité des signalisations décrites à ce jour restent mal caractérisées, avec des liens parfois flous entre l'expression de la KMT et l'activité de ses supposées cibles dans un contexte cellulaire donné. La recherche de nouveaux substrats utilise principalement des méthodes protéomiques à haut débit qui présentent souvent un niveau de précision et de pertinence biologique discutable, et qui doivent conduire à une grande prudence dans l'interprétation des résultats. La compréhension de la régulation des événements de méthylation des lysines et leurs conséquences biologiques représentent un enjeu de taille dans la compréhension du fonctionnement de nombreuses signalisations cellulaires et pour l'identification de nouvelles modalités pour le ciblage thérapeutique.

Identification de nouveaux substrats : contraintes techniques et critères de choix

Dans notre étude, nous avons identifié plusieurs nouveaux substrats et confirmé *in vitro* et en cellules que SMYD2 monométhyle les substrats BCAR3 et RTF1, respectivement sur leur lysine K334 et K587, et que RNF113A est triméthylée par SMYD3 sur sa lysine K20. Pour identifier ces substrats, nous avons utilisé deux techniques différentes, la technique de méthylation radioactive sur puce à protéines ou ProtoArray® et la technique du SILAC-3xMBT-*pull-down*. La technique ProtoArray® permet de rechercher une méthylation induite par une KMT d'intérêt sur de nombreuses protéines recombinantes analysées en parallèle et sans *a priori*. Elle permet ainsi d'étudier des signalisations indépendamment du type

cellulaire ou d'une situation physiologique ou pathologique donnée. Cette méthode est peu fiable pour SMYD2, qui est capable de méthyle de nombreux substrats *in vitro* de façon peu spécifique, sans confirmation systématique en cellule, générant ainsi un nombre élevé de faux positifs. En effet, en présence de SMYD2, 251 protéines présentaient un signal de méthylation en ProtoArray®, ce qui complexifiait la sélection d'un nombre limité de candidats pour vérification approfondie. De plus, seulement 2 substrats déjà décrits dans la littérature, HSP90 et MAPKAPK3, ont été retrouvés parmi notre liste de candidats ^{3,4}. A l'inverse, SMYD3 semble plus spécifique *in vitro* car, en excluant les auto-méthylations des méthyltransférases présentes sur la puce, seulement 10 candidats ont été détectés comme étant méthylés. Sur ces 10 candidats, 5 ont pu être produits en bactérie, et la méthylation a été démontrée pour 2 d'entre eux (RNF113A et RIOK1), confirmant la nécessité de d'étapes supplémentaires de validation des cibles présomptives identifiées par criblage *in vitro*. Le caractère *in vitro* de cette technique induit de nombreux biais lors de la mise en contact très artificielle de 2 protéines ; la conformation tridimensionnelle et la capacité d'interaction de ces protéines peuvent en effet être profondément altérées, du fait de l'absence de PTMs et des différents paramètres biochimiques (viscosité, pH, pression, concentration des protéines au niveau intracellulaire) et biologiques (localisation intracellulaire, moment du cycle cellulaire, complexes protéiques). Les résultats du ProtoArray® doivent donc faire l'objet d'une analyse prudente et nécessitent plusieurs étapes de confirmation directe *in vitro*, puis en cellules.

La technique du SILAC-3xMBT-*pulldown* consiste à marquer différenciellement le protéome de deux cultures cellulaires, en présence ou non de la KMT d'intérêt, puis à enrichir les protéines contenant des lysines mono- et di-méthylées ⁵. Cette méthode est donc *a priori* peu adaptée pour une triméthyltransférase comme SMYD3. Elle a néanmoins permis la mise en évidence de substrats méthylés par SMYD3 car cette dernière agit de manière processive, ce qui génère des substrats à l'état mono-, di-, ou tri-méthylé. Ainsi cette méthode a identifié MAP3K2, un substrat déjà caractérisé dans la littérature, et a identifié 5 autres substrats, parmi lesquels 3 ont été produits en protéine recombinante, et seule

HSP90 a montré un signal de méthylation de faible intensité *in vitro* en présence de SMYD3. Le 3xMBT-*pulldown* semble plus adapté pour une KMT telle que SMYD2, car nous avons pu confirmer plusieurs évènements de méthylation identifiés par cette technique. En effet, sur les 6 protéines que nous avons sélectionnées parmi la liste des 55 candidats obtenus, 5 ont été produites en protéine recombinante, et la méthylation de 4 d'entre elles a été confirmée en test de méthylation *in vitro* (RTF1, BCAR3, RRP44 et SARG) puis en cellules pour RTF1, BCAR3 et RRP44.

De nombreuses autres méthodes sont utilisées dans la littérature pour identifier de nouvelles cibles des méthyltransférases, notamment pour SMYD2. Olsen et coll. ont établi une liste de substrats de SMYD2 par immunoprécipitation de peptides porteurs d'une lysine monométhylée grâce à un anticorps pan-monométhyle. Ils ont utilisé des extraits de cellules de carcinome épidermoïde de l'œsophage surexprimant SMYD2, traitées ou non avec un inhibiteur de SMYD2. Cette approche est intéressante car elle permet une identification directe de la lysine méthylée des substrats candidats, sous réserve que le site de méthylation soit compatible avec la méthode de digestion par la trypsine utilisée. Olsen et coll. ont détecté plus de 300 candidats potentiellement méthylés par SMYD2, dont 32 étaient également présents dans notre liste de candidats obtenus par SILAC-3xMBT-*pulldown*. Parmi les 6 substrats sélectionnés pour confirmation de méthylation *in vitro* (BCAR3, LSM14B, RRP44, RTF1, SARG, ACTC1), tous étaient également présents parmi les substrats potentiels de Olsen et coll., à l'exception de ACTC1. Ce dernier était, de notre point de vue, un substrat prometteur car identifié à la fois pour SMYD2 et SMYD3, et décrit dans la littérature comme intervenant dans le métabolisme musculaire cardiaque, ce qui pouvait renforcer les données suggérant un rôle physiologique des deux KMTs dans ce domaine ⁶. Ahmed et coll. ont proposé une méthode d'identification bioinformatique basée sur des données *in silico* intégrant des paramètres structuraux, biochimiques, cellulaires et protéomiques. Ils ont ainsi sélectionné un panel de motifs peptidiques proches des sites d'interactions déjà décrits pour d'autres substrats. Ce type de processus peut permettre d'identifier des nouvelles interactions mais reste cependant ciblé sur des motifs issus

d'interactions déjà connues. Dans le cas d'une KMT comme SMYD2, qui est capable de méthyler ses substrats sur des lysines au sein de séquences très diverses, cette approche peut générer un grand nombre d'interactions artéfactuelles ⁷. Ainsi, ils ont obtenu 14 substrats dont 6 ont été confirmés en cellules. Aucun de ces substrats n'était commun avec notre liste de substrats potentiels.

Enfin, Cornett et coll. ont utilisé une approche de protéomique fonctionnelle originale qui se base sur une banque de 47 millions de peptides et qui s'affranchit des étapes de digestion nécessaires en spectrométrie de masse, ce qui peut permettre la mise en évidence de substrats non identifiables en MS/MS. Cette méthode, basée sur la reconnaissance d'un panel exhaustif de peptides linéaires, ne tient pas compte des contraintes biologiques cellulaires ni de la conformation tridimensionnelle de la structure complète des protéines qui conditionne la présentation d'un motif d'interaction. Elle peut également générer des peptides non retrouvés dans les cellules, donc sans valeur biologique ⁸.

Ces différentes méthodes sont, pour la plupart, libérées des contraintes liées aux conditions cellulaires, ce qui impacte potentiellement la pertinence biologique des résultats, avec un risque d'artéfacts produits *in vitro* ou *in silico*. En effet, le ratio substrats validés/faux positifs de ces techniques de protéomique à haut débit est globalement assez décevant, raison pour laquelle leur principal rôle est d'orienter et de suggérer des candidats de manière préliminaire. Leur utilisation impose donc une analyse fine de leurs résultats bruts et des étapes de validation *in vitro* puis en cellules. Cela requiert également le développement d'outils spécifiques à la signalisation étudiée, tels que des anticorps méthyl-spécifiques.

Du candidat presomptif au substrat valide : stratégies et priorités

Après confirmation *in vitro* et en lignées cellulaires des événements de méthylation, un certain nombre d'étapes restent à valider pour permettre la caractérisation d'une nouvelle signalisation, telles que la réalisation d'un SILAC *peptide-pulldown* permettant l'identification

des interacteurs spécifiques de l'état méthylé ou non du substrat, et l'étude des différentes voies de signalisation impactées.

Parmi les principaux substrats que nous avons confirmés de manière fiable en cellules, RTF1 est un substrat particulièrement prometteur du fait de son rôle physiologique essentiel. Le rôle des PTMs dans la régulation de l'activité de RTF1 a été peu exploré jusqu'à présent, ce qui souligne l'intérêt d'étudier l'impact potentiel de K587me1 sur ses différentes fonctions. Cette marque de méthylation pourrait en effet impacter le fonctionnement de la RNA polymérase 2, l'hématopoïèse, les modifications d'histones telles que l'ubiquitinylation de l'histone H2B et la méthylation H3K4me3, avec des conséquences possibles sur la pluripotence des cellules souches^{9,10}. Sa caractérisation n'a néanmoins pas été poursuivie au-delà de la confirmation de la méthylation en cellules par manque de temps et devant l'apparition d'obstacles tels qu'une fiabilité insuffisante des peptides produits rendant les résultats du SILAC incertains, de même que l'impossibilité de réprimer de manière stable son expression en lignée cellulaire. En effet, la répression de RTF1 induisait la mortalité de l'ensemble des cellules transfectées par nos différents shRNA. Sa perte d'expression est probablement incompatible avec la survie cellulaire du fait de ses implications multiples dans la régulation de la chromatine et dans d'autres mécanismes transcriptionnels et post-transcriptionnels en association au *Paf1 complex* (PAF1C). Une solution pourrait être sa surexpression sous sa forme sauvage et mutante K587R, suivie d'une répression avec un shRNA ciblant les régions UTR afin de ne réprimer que l'expression du transcript endogène. Cela permettrait donc d'étudier l'impact de la méthylation RTF1 K587me1, dont les conséquences phénotypiques sont probablement plus fines qu'une répression totale de RTF1, et donc potentiellement compatibles avec la survie cellulaire. RTF1 reste donc un sujet d'étude pour notre laboratoire et fera l'objet d'un projet de recherche à part entière dans un futur très proche.

La signalisation SMYD2 - BCAR3 K334 est un enjeu de recherche extrêmement prometteur, puisque ce substrat est impliqué dans l'invasion et la migration des cellules tumorales dans le cancer du sein, dans lequel SMYD2 est surexprimée^{11,12}. Notre approche par SILAC

peptide-pulldown a révélé que BCAR3 interagit avec des protéines impliquées dans le cytosquelette cellulaire (données non publiées et confidentielles). Ces résultats servent de point de départ pour une étude approfondie au sein de l'équipe du Dr Reynoird. Cette étude vise à confirmer les interacteurs méthyl-spécifiques identifiés et à étudier l'impact de BCAR3 K334me1 sur les signalisations d'aval. Pour cela, des lignées de cancer du sein seront établies avec répression et surexpression stable de SMYD2, et de BCAR3 sous sa forme sauvage, et mutée K334R, afin d'étudier les conséquences phénotypiques de cette marque de méthylation sur la morphologie des cellules et sur leur capacité à migrer. Par ailleurs, ces lignées seront utilisées pour réaliser des xénogreffes afin d'étudier l'impact de la méthylation de BCAR3 par SMYD2 sur la prolifération tumorale et le potentiel métastatique des tumeurs.

La signalisation SMYD2-RRP44 représente également un intérêt majeur du fait de l'appartenance de RRP44 à l'exosome et de sa participation à la dégradation de l'ARN et à l'épissage alternatif ¹³. Cela suggère que la méthylation de RRP44 pourrait contribuer à la régulation de l'expression de nombreux gènes. A ce jour, nous n'avons pas identifié clairement le ou les sites de méthylation de RRP44 car la mutation K331R ne supprime pas complètement le signal de méthylation, ce qui suggère qu'une autre lysine est méthylée par SMYD2. Il est donc nécessaire de produire plusieurs segments protéiques recombinants afin de réaliser de nouveaux tests de méthylation *in vitro* pour identifier, dans un premier temps, la portion de RRP44 dans laquelle se situe le second site méthylé. Ces fragments seront ensuite envoyés en analyse MS/MS, dont la méthode de digestion devra être modifiée afin d'augmenter les chances d'obtenir un peptide contenant la lysine méthylée. Ce résultat devra être ensuite confirmé par de nouvelles expériences de mutagenèse dirigée.

Les autres marques de méthylation que nous avons identifiées ne font l'objet que de données très préliminaires, mais pourraient impacter différentes voies de signalisation d'intérêt. Ainsi, la méthylation de SARG par SMYD2 pourrait influencer la sensibilité des tumeurs prostatiques aux androgènes ¹⁴. La méthylation de RIOK1 par SMYD3 pourrait jouer un rôle dans la régulation des liaisons entre protéines et la maturation de l'ADN, et favoriser la prolifération cellulaire dans certains cancers, tels que le cancer colorectal, ou le cancer

gastrique dans lequel RIOK1 est surexprimée ^{15,16}. Enfin, la méthylation de HSP90 par SMYD3 pourrait modifier ses fonctions de chaperonne impliquées dans de nombreux processus physiologiques et pathologiques ¹.

Caractérisation des conséquences de la triméthylation de rnf113a par smyd3

La signalisation sur laquelle je me suis focalisé au cours de ma thèse est la triméthylation de RNF113A K20 par SMYD3. Cet événement de méthylation induit une perte d'interaction de RNF113A avec la phosphatase PP4, notamment la sous-unité R3A/B qui lie RNF113A non méthylée. Ceci aboutit à une perte de la déphosphorylation de RNF113a, augmentant donc de manière constitutive le niveau de phosphorylation de RNF113A. Cette phosphorylation permanente stimule son activité d'ubiquitine ligase impliquée dans la réparation de l'ADN après exposition à des agents alkylants (Figure 1). Cette signalisation offre une nouvelle explication possible au rôle pro-oncogène de SMYD3 dont la surexpression est associée à un mauvais pronostic dans plusieurs cancers ¹. Nous avons initialement décidé d'étudier SMYD2 et SMYD3 dans l'adénocarcinome pancréatique du fait des récentes données ^{4,17}, de l'enjeu de santé publique que représente cette pathologie, et de ma pratique clinique principalement axée sur les cancers digestifs et notamment bilio-pancréatiques. Néanmoins, la signalisation que nous avons privilégiée portant plus spécifiquement sur un substrat de SMYD3, nous nous sommes finalement orientés vers un cancer dont la pathogénèse n'est pas dépendante d'une mutation activatrice de l'oncogène Ras. En effet, la présence d'une mutation activatrice de Ras est retrouvée dans plus de 90% des adénocarcinomes pancréatiques ¹⁸, ce qui suggère que, dans cette pathologie, la signalisation SMYD3-MAP3K2 constitue une part importante de l'implication de SMYD3 dans l'oncogenèse et la prolifération cellulaire, qui peut rendre délicate l'étude d'autres voies de signalisation impactées par SMYD3. Ainsi, en nous affranchissant de la signalisation Ras/Raf/Mek/Erk, nous augmentons nos chances de mettre en évidence les conséquences phénotypiques secondaires à la triméthylation de RNF113A. Nous avons décidé de choisir comme contexte

pathologique les cancers bronchopulmonaires à petites cellules (CBPPC). Dans cette pathologie agressive, la survie à 5 ans est de 7% et les options thérapeutiques sont très restreintes, ce qui en fait un enjeu de santé publique ¹⁹. RNF113A et SMYD3 sont fortement exprimées dans les CBPPC, elles représentent donc des nouvelles signalisations d'intérêt potentielles, dont la caractérisation devrait aider à mieux comprendre la tumorigenèse de ces cancers et aider au développement de nouvelles stratégies thérapeutiques. Des modèles murins de CBPPC déjà développés et maîtrisés avec des KO constitutionnels *Smyd3*^{-/-} chez nos collaborateurs du "MD Anderson Anticancer Center" à Houston, confirment l'importance de SMYD3 dans la prolifération tumorale et le potentiel métastatique de ces tumeurs. Les analyses *in vivo* de xénogreffes de cellules de CBPPC *Smyd3*^{-/-} et/ou *Rnf113a*^{-/-} sont en cours pour étudier les conséquences du KO de SMYD3 et/ou de RNF113A sur la prolifération tumorale. Nous allons également étudier le caractère pronostique de l'expression de SMYD3 et de RNF113A dans cette pathologie à partir des prélèvements conservés et des données cliniques collectées dans le cadre de la tumorothèque du CHU Grenoble-Alpes (N° AC-2014-2094), dont les prélèvements ont été largement utilisés dans le cadre d'études internationales sur la caractérisation histomoléculaire des carcinomes pulmonaires (^{20,21}).

La spécificité de l'implication de RNF113A dans la réparation des lésions d'alkylation via son interaction avec l'ASCC semblait démontrée par l'étude de Mosammaparast et coll. ²². Néanmoins, il a très récemment été démontré que l'expression de RNF113A est augmentée suite à l'exposition à des drogues de chimiothérapie utilisant d'autres mécanismes, tels que les inhibiteurs de la topoisomérase II avec l'étoposide, ou des sels de platine. Dans cette étude, RNF113A semble participer à la chimiorésistance des adénocarcinomes pulmonaires traités par cisplatine notamment. Cette notion renforce l'intérêt de notre modèle de CBPPC dans lequel la première ligne de traitement est une association de cisplatine et d'étoposide

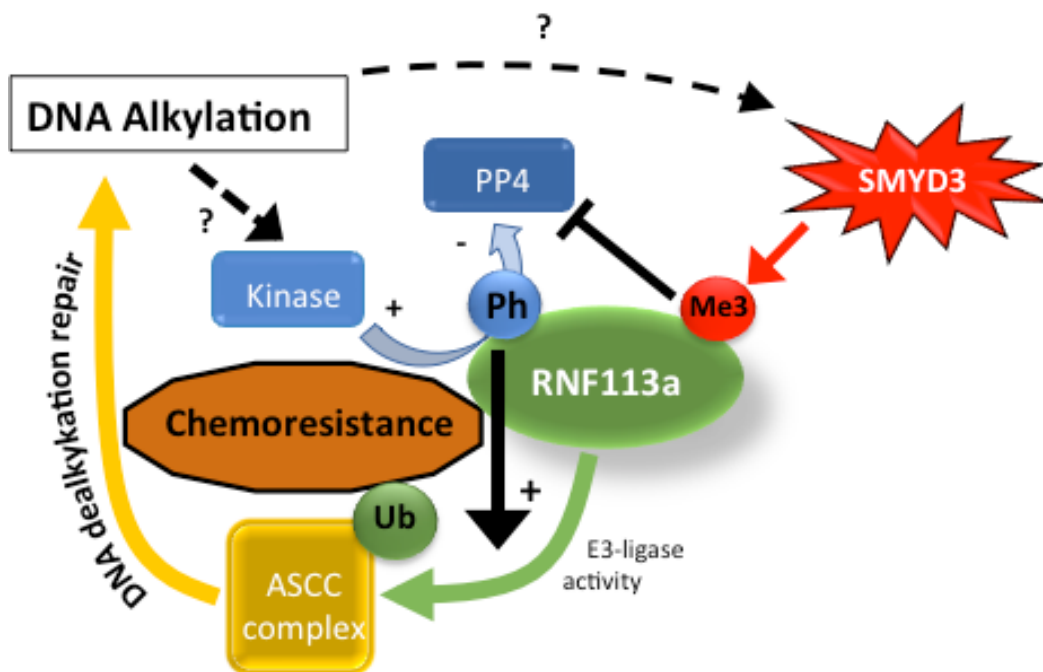


Figure 1 : Représentation schématique de la signalisation SMYD3-RNF113A et de son rôle dans la chimiorésistance aux agents alkylants. En méthylant RNF113A sur sa lysine K20, SMYD3 induit une augmentation du niveau de phosphorylation de RNF113A par perte de son interaction avec sa phosphatase PP4. Cela aboutit à une augmentation de l'activité de E3-ubiquitine ligase de RNF113A impliquée dans la réparation des lésions d'alkylation de l'ADN.

Concernant le rôle de la signalisation SMYD3-RNF113A dans la réparation des lésions d'alkylation, un certain nombre d'éléments restent cependant à explorer. En particulier, le mécanisme par lequel la survenue de lésions d'alkylation active RNF113A et l'impact potentiel de ces dernières sur l'activité de SMYD3 ne sont pas clairement compris. Le délai d'action semble trop court pour passer par des mécanismes transcriptionnels et implique probablement la reconnaissance de la marque d'alkylation par un domaine de RNF113A ou un de ses interacteurs. RNF113A contient un motif en doigt de zinc CCCH capable de se lier à l'ARN. L'ARN étant également modifié en cas d'exposition aux agents alkylants, il est possible que ce dernier ait un rôle dans l'activation de RNF113A, comme suggéré par des données non publiées récentes du laboratoire du Dr Mosammamaparast avec qui nous collaborons.

Le modèle du CBPPC a été sélectionné pour des raisons médico-scientifiques et pragmatiques comme précédemment mentionné, mais la signalisation SMYD3-RNF113A est

très probablement impliquée dans de nombreux autres cancers, et sa caractérisation de même que l'étude de son impact pronostique constituent donc des enjeux majeurs pour la recherche en oncologie en général. Dans cette étude, nous abordons l'alkylation comme mode d'action cytotoxique pharmacologique, mais elle constitue également un mécanisme de carcinogénèse très important. Les lésions d'alkylation de l'ADN correspondent à la formation d' O⁶-methylguanine (O⁶-meG) au niveau de l'ADN. Elles peuvent être réparées en théorie par différents systèmes, tels que la réparation par excision de bases (*Base Excision Repair*, BER) ou la réparation des mésappariements (*Mismatch repair*, MMR), notamment via l'hétérodimère hMSH2-hMSH6 ²⁴. Ces lésions d'alkylation sont associées à une augmentation du nombre de mutations ponctuelles, au niveau d'oncogènes tels que HRAS (mutations activatrices) ou de suppresseurs de tumeurs tels que TP53 (mutations inactivatrices). Il semble donc essentiel d'étudier les liens entre l'alkylation comme mécanisme de carcinogénèse et le niveau d'expression de RNF113A.

Cette notion soulève plusieurs questions :

- La surexpression de RNF113A constitue-t-elle un facteur protecteur, qui peut limiter les lésions d'alkylation carcinogènes ?
- Une perte de fonction de RNF113A peut-elle favoriser l'émergence de tumeurs par défaut de réparation de l'ADN lors des processus de carcinogénèse ?
- Les tumeurs présentant un niveau faible d'expression de RNF113A présentent-elles une charge mutationnelle plus élevée, et cela induit-il des modifications de leur microenvironnement immunitaire (infiltrat lymphocytaire CD8, taux d'expression du PD1 (*programmed cell death 1*) ?
- Enfin, l'expression de RNF113A est-elle modifiée dans les tumeurs présentant un déficit de réparation de l'ADN, tel qu'un défaut d'expression d'une des protéines du MMR dans le cadre du Syndrome de Lynch, ou en contexte de déficience d'autres systèmes, tels que le BER ou la O⁶-Methylguanine-DNA méthyltransférase (MGMT)?

L'étude du niveau d'alkylation peut être effectuée en lignées cellulaires avec surexpression ou répression de SMYD3 et RNF113A et mesure du pourcentage de O⁶-meG par

pyroséquençage. Néanmoins, ces lignées étant déjà fortement modifiées et cancéreuses, une étude de la carcinogénèse *in vivo* semble plus pertinente. L'utilisation de modèles de souris exposés à des agents carcinogènes alkylants, réprimés ou non pour RNF113A permettrait d'étudier l'impact de la répression de RNF113A sur la survenue de cancers induits par des carcinogènes alkylants, avec l'hypothèse d'une augmentation de ce risque. Par exemple, une exposition au diethylnitrosamine (DEN), un carcinogène induisant de multiples altérations, dont l'alkylation, et qui induit la formation de tumeurs coliques et hépatiques chez la souris ²⁵ pourrait être réalisée chez des souris exprimant normalement RNF113A et chez des souris avec KO conditionnel de RNF113a pour comparer la cinétique d'apparition de ces lésions.

Enfin, une validation translationnelle est essentielle, basée sur un suivi longitudinal de l'expression de SMYD3 et RNF113A par immunohistochimie (IHC) sur des prélèvements en début de prise en charge (avant traitement), puis après exposition à un agent alkylant. Cela permettrait de mieux définir la valeur prédictive de réponse ou de non-réponse aux traitements, ainsi que l'impact pronostique du niveau d'expression pré-thérapeutique de SMYD3 et RNF113A. L'analyse après exposition permettrait également de comparer le niveau d'expression entre les répondeurs et les non-répondeurs au traitement, et son évolution au sein de chaque groupe (pré- vs post-thérapeutique). Ce type d'étude permettrait de répondre à plusieurs questions :

- Une forte expression initiale de SMYD3 et/ou RNF113A est-elle associée à un mauvais pronostic et à une mauvaise réponse à la chimiothérapie ? Ainsi en fonction des différents niveaux d'expression, des groupes histo-pronostiques pourraient être définis (1 : SMYD3 bas - RNF113A bas, 2 : SMYD3 bas - RNF113A élevé, 3 : SMYD3 élevé - RNF113A bas, 4 : SMYD3 élevé - RNF113A élevé).

- Observe-t-on une augmentation de l'expression de SMYD3 et/ou RNF113A au cours de la chimiothérapie chez les non-répondeurs ? Ce paramètre est intéressant pour étudier l'activation potentielle de cette voie de chimiorésistance, après expositions aux agents alkylants pharmacologiques. Ce type de suivi longitudinal intégrant des marqueurs

immunohistochimiques reste néanmoins plus compliqué à mettre en œuvre pour des raisons éthiques.

L'analyse pré-thérapeutique de l'expression de SMYD3 et RNF113A est également intéressante car elle permettrait de comparer le niveau d'expression entre des patients présentant un cancer dans un contexte d'exposition à des agents carcinogènes alkylants, tels que les nitrosamines (cigarettes, viandes salées ou saumurées, etc) et des sujets présentant la même pathologie sans ces facteurs de risque. Concernant son rôle pronostique et son implication dans la chimiorésistance, à l'heure actuelle très peu de données sont disponibles sur RNF113A, et les résultats sont discordants. Une meilleure caractérisation de son impact pronostique et prédictif de réponse pourrait ainsi affiner le choix des traitements des patients atteints de cancer, en sélectionnant les patients plus sensibles aux agents alkylants. Ce type de corrélation a déjà été observé concernant d'autres systèmes de réparation des lésions d'alkylation telle que la MGMT dans les tumeurs neuroendocrines (TNE) ou dans le glioblastome ²⁶. En effet, une expression élevée de la MGMT semble prédictive de chimiorésistance au temozolomide (TEM), un agent alkylant largement utilisé dans ces indications. Ainsi, j'ai initié une collaboration avec le Pr Thomas Walter (Hôpital Edouard Herriot, Lyon), spécialisé dans les TNE digestives, dans le but d'étudier la valeur prédictive du niveau d'expression de SMYD3 et de RNF113A, dans la réponse au TEM. Pour ce faire, nous allons analyser en IHC l'expression de SMYD3 et de RNF113A sur des échantillons tumoraux pré-thérapeutiques fixés en paraffine de 80 patients porteurs d'une TNE pancréatique et traités par TEM. Pour ces patients, de nombreuses données cliniques, biologiques, radiologiques et anatomopathologiques ont été colligées, dont notamment le niveau d'expression de la MGMT en IHC, ce qui nous permettra d'étudier les corrélations entre les deux voies de réparation des lésions d'alkylation de l'ADN.

De plus, nous aurons accès aux échantillons de l'essai de phase 2 MGMT-NET (NCT03217097) actuellement en cours d'inclusion. Ce dernier étudie la valeur prédictive de réponse tumorale objective, de la méthylation du promoteur de la MGMT dans les TNE traitées par TEM ou une chimiothérapie contenant de l'oxaliplatine. S'agissant d'une étude

ancillaire dans le cadre d'un essai randomisé, multicentrique, prospectif, cela nous permettra de vérifier la validité de nos résultats avec un haut niveau de preuve.

Par ailleurs, des modèles *in vitro* et *in vivo* plus pertinents, utilisant des cellules tumorales issues de patients, doivent également être développés, tels que les organoïdes et les xénogreffes de type PDX (*patient-derived xenograft*). Ce type de modèle permet d'étudier les différences interindividuelles de sensibilité, aux agents alkylants avant d'initier un traitement et d'étudier les liens entre cette chimiosensibilité et l'expression de SMYD3 et RNF113A. Sur un plan plus médical, l'objectif théorique de ces modèles est de permettre une médecine personnalisée, même si le délai de mise en place n'est pas toujours compatible avec une prise en charge optimale des patients dans les pathologies rapidement évolutives telles que le CBPPC.

Une autre hypothèse à explorer dans l'étude des variabilités interindividuelles de sensibilité aux agents carcinogènes alkylants est l'existence de polymorphismes génétiques impactant la fonction de SMYD3 et/ou de RNF113A. De tels polymorphismes pourraient expliquer que malgré des niveaux d'expression comparable de ces deux protéines, la sensibilité aux carcinogènes alkylants, ou aux agents chimiothérapeutiques alkylants, diffère d'un sujet à l'autre, ou d'une tumeur à l'autre. Des polymorphismes ont en effet déjà été mis en évidence pour SMYD3, tels que des polymorphismes de répétition du motif CCGCC dans la région de liaison au facteur de transcription E2F-1 en 5' de SMYD3. Dans ce cas, plus le nombre de répétitions du motif est grand, plus l'interaction entre la région promotrice de SMYD3 et E2F-1 est importante, induisant une augmentation de l'activité transcriptionnelle de SMYD3 ²⁷.

Conclusions : de la découverte a la perspective clinique

En conclusion, ce travail de recherche constitue un effort structuré pour explorer un domaine encore très peu balisé de la signalétique moléculaire et révéler au grand jour de nouveaux processus moléculaires impliqués dans la cancérogenèse. Ce travail devrait non seulement améliorer la compréhension de processus complexes mais également aider à développer de

nouveaux biomarqueurs pour le pronostic, la prédiction et le ciblage thérapeutique. Plus précisément, nous avons identifié et confirmé de manière fiable 4 nouveaux substrats pour SMYD2, et 3 pour SMYD3, générant autant de pistes de projets de recherche avec pour but la caractérisation des différentes voies de signalisation impactées par ces événements de méthylation. Ce travail a servi de base pour des programmes au sein de notre laboratoire visant à explorer de manière méthodique les événements de signalisation liés à RTF1 K587me1, BCAR3 K334me1 et RNF113A K20me3. Les deux premiers de ces événements font actuellement l'objet de projets d'étude au sein de l'équipe de recherche, avec un niveau d'avancement important pour BCAR3 (projet développé par A. Casanova, doctorant encadré par le Dr Reynoird). Le troisième événement permet d'éclairer un nouvel aspect du rôle oncogénique de SMYD3, indépendant de son activité déjà décrite dans la voie RAS-MAPK, à savoir, la promotion de l'activité ubiquitine ligase de RNF113A. Notre hypothèse est que cette voie participe à la chimiorésistance aux agents alkylants dans les cancers broncho-pulmonaires à petites cellules. Nous sommes actuellement en train de développer des lignées cellulaires stables présentant différents niveaux d'expression de SMYD3, et de RNF113A sous sa forme sauvage et mutante K20R afin de confirmer le rôle de la méthylation de RNF113A dans la régulation de la réparation des lésions alkylantes de l'ADN, *in vitro* et en xéno greffe. La validation translationnelle de SMYD3 et RNF113A en tant que marqueurs prédictifs de réponse aux agents alkylants est également en cours dans une cohorte de patients porteurs d'une tumeur neuroendocrine pancréatique.

Enfin, la caractérisation de ce nouveau mécanisme de chimiorésistance devrait également permettre le développement de thérapies innovantes ciblant la réparation de l'ADN, afin de renforcer l'activité anti-tumorale des chimiothérapies cytotoxiques dite classiques.

L'ensemble de ces études devraient donc permettre une meilleure compréhension de la physiopathologie des cancers, une meilleure identification des patients bénéficiant ou non des agents pharmacologiques alkylants et devrait *in fine* améliorer la prise en charge des patients atteints de cancer.

REFERENCES

- 1 bottino, c., peserico, a., simone, c. & caretti, g. Smyd3: an oncogenic driver targeting epigenetic regulation and signaling pathways. *Cancers* **12**, 142, doi:10.3390/cancers12010142 (2020).
- 2 yi, x., jiang, x. J. & fang, z. M. Histone methyltransferase smyd2: ubiquitous regulator of disease. *Clin epigenetics* **11**, 112, doi:10.1186/s13148-019-0711-4 (2019).
- 3 hamamoto, r., toyokawa, g., nakakido, m., ueda, k. & nakamura, y. Smyd2-dependent hsp90 methylation promotes cancer cell proliferation by regulating the chaperone complex formation. *Cancer lett* **351**, 126-133, doi:10.1016/j.canlet.2014.05.014 (2014).
- 4 reynoird, n. *Et al.* Coordination of stress signals by the lysine methyltransferase smyd2 promotes pancreatic cancer. *Genes dev* **30**, 772-785, doi:10.1101/gad.275529.115 (2016).
- 5 moore, k. E. *Et al.* A general molecular affinity strategy for global detection and proteomic analysis of lysine methylation. *Mol cell* **50**, 444-456, doi:10.1016/j.molcel.2013.03.005 (2013).
- 6 tracy, c. *Et al.* The smyd family of methyltransferases: role in cardiac and skeletal muscle physiology and pathology. *Curr opin physiol* **1**, 140-152, doi:10.1016/j.cophys.2017.10.001 (2018).
- 7 ahmed, h., duan, s., arrowsmith, c. H., barsyte-lovejoy, d. & schapira, m. An integrative proteomic approach identifies novel cellular smyd2 substrates. *J proteome res* **15**, 2052-2059, doi:10.1021/acs.jproteome.6b00220 (2016).
- 8 cornett, e. M. *Et al.* A functional proteomics platform to reveal the sequence determinants of lysine methyltransferase substrate selectivity. *Sci adv* **4**, eaav2623-eaav2623, doi:10.1126/sciadv.aav2623 (2018).
- 9 van oss, s. B. *Et al.* The histone modification domain of paf1 complex subunit rtf1 directly stimulates h2b ubiquitylation through an interaction with rad6. *Mol cell* **64**, 815-825, doi:10.1016/j.molcel.2016.10.008 (2016).
- 10 cao, q. F. *Et al.* Characterization of the human transcription elongation factor rtf1: evidence for nonoverlapping functions of rtf1 and the paf1 complex. *Mol cell biol* **35**, 3459-3470, doi:10.1128/mcb.00601-15 (2015).
- 11 cross, a. M. *Et al.* Breast cancer antiestrogen resistance 3-p130(cas) interactions promote adhesion disassembly and invasion in breast cancer cells. *Oncogene* **35**, 5850-5859, doi:10.1038/onc.2016.123 (2016).
- 12 li, l. X. *Et al.* Lysine methyltransferase smyd2 promotes triple negative breast cancer progression. *Cell death dis* **9**, 326, doi:10.1038/s41419-018-0347-x (2018).
- 13 mette, m. F., aufsatz, w., winden, j. V. D., matzke, m. A. & matzke, a. J. M. Transcriptional silencing and promoter methylation triggered by double - stranded rna. *The embo journal* **19**, 5194-5201, doi:10.1093/emboj/19.19.5194 (2000).
- 14 steketee, k., ziel-van der made, a. C., van der korput, h. A., houtsmuller, a. B. & trapman, j. A bioinformatics-based functional analysis shows that the specifically androgen-regulated gene

- sarg contains an active direct repeat androgen response element in the first intron. *J mol endocrinol* **33**, 477-491, doi:10.1677/jme.1.01478 (2004).
- 15 hong, x. *Et al.* Targeting posttranslational modifications of riok1 inhibits the progression of colorectal and gastric cancers. *Elife* **7**, doi:10.7554/elife.29511 (2018).
- 16 read, r. D. *Et al.* A kinome-wide rnai screen in drosophila glia reveals that the rio kinases mediate cell proliferation and survival through torc2-akt signaling in glioblastoma. *Plos genet* **9**, e1003253, doi:10.1371/journal.pgen.1003253 (2013).
- 17 mazur, p. K. *Et al.* Smyd3 links lysine methylation of map3k2 to ras-driven cancer. *Nature* **510**, 283-287, doi:10.1038/nature13320 (2014).
- 18 ryan, d. P., hong, t. S. & bardeesy, n. Pancreatic adenocarcinoma. *N engl j med* **371**, 1039-1049, doi:10.1056/nejmra1404198 (2014).
- 19 gazdar, a. F., bunn, p. A. & minna, j. D. Small-cell lung cancer: what we know, what we need to know and the path forward. *Nature reviews. Cancer* **17**, 725-737, doi:10.1038/nrc.2017.87 (2017).
- 20 mcleer-florin, a. *Et al.* Alk fusion variants detection by targeted rna-next generation sequencing and clinical responses to crizotinib in alk-positive non-small cell lung cancer. *Lung cancer* **116**, 15-24, doi:10.1016/j.lungcan.2017.12.004 (2018).
- 21 scoccianti, c. *Et al.* Prognostic value of tp53, kras and egfr mutations in nonsmall cell lung cancer: the euelc cohort. *Eur respir j* **40**, 177-184, doi:10.1183/09031936.00097311 (2012).
- 22 brickner, j. R. *Et al.* A ubiquitin-dependent signalling axis specific for alkbh-mediated dna dealkylation repair. *Nature* **551**, 389-393, doi:10.1038/nature24484 (2017).
- 23 shostak, k. *Et al.* The x-linked trichothiodystrophy-causing gene rnf113a links the spliceosome to cell survival upon dna damage. *Nat commun* **11**, 1270, doi:10.1038/s41467-020-15003-7 (2020).
- 24 margison, g. P., santibáñez koref, m. F. & povey, a. C. Mechanisms of carcinogenicity/chemotherapy by o6-methylguanine. *Mutagenesis* **17**, 483-487, doi:10.1093/mutage/17.6.483 (2002).
- 25 sarris, m. E., moulos, p., haroniti, a., giakountis, a. & talianidis, i. Smyd3 is a transcriptional potentiator of multiple cancer-promoting genes and required for liver and colon cancer development. *Cancer cell* **29**, 354-366, doi:10.1016/j.ccell.2016.01.013 (2016).
- 26 fan, c. H. *Et al.* O6-methylguanine dna methyltransferase as a promising target for the treatment of temozolomide-resistant gliomas. *Cell death dis* **4**, e876, doi:10.1038/cddis.2013.388 (2013).
- 27 tsuge, m. *Et al.* A variable number of tandem repeats polymorphism in an e2f-1 binding element in the 5' flanking region of smyd3 is a risk factor for human cancers. *Nature genetics* **37**, 1104-1107, doi:10.1038/ng1638 (2005).

ANNEXES



Lysine methylation signaling in pancreatic cancer

Gaël S. Roth^{a,b,c}, Alexandre G. Casanova^a, Nathanaël Lemonnier^a,
and Nicolas Reynoird^a

Purpose of review

Despite better knowledge of its genetic basis, pancreatic cancer is still highly lethal with very few therapeutic options. In this review, we discuss the potential impact of epigenetic therapies, focusing on lysine methylation signaling and its implication in pancreatic cancer.

Recent findings

Protein lysine methylation, a key mechanism of posttranslational modifications of histone proteins, has emerged as a major cell signaling mechanism regulating physiologic and pathologic processes including cancer. This finely tuned and dynamic signaling mechanism is regulated by lysine methyltransferases (KMT), lysine demethylases (KDM) and signal transducers harboring methyl-binding domains. Recent evidence demonstrates that overexpression of cytoplasmic KMT and resulting enhanced lysine methylation is a reversible event that enhances oncogenic signaling through the Ras and Mitogen-Activated Protein Kinases pathway in pancreatic cancer, opening perspectives for new anticancer chemotherapeutics aimed at controlling these activities.

Summary

The development of potent and specific inhibitors of lysine methylation signaling may represent a hitherto largely unexplored avenue for new forms of targeted therapy in cancer, with great potential for yet hard-to-treat cancers such as pancreatic cancer.

Keywords

epigenetic therapy, lysine demethylase, lysine methylation signaling, lysine methyltransferase, pancreatic cancer

INTRODUCTION

Pancreatic ductal adenocarcinoma (PDAC) is the most frequent type of pancreatic cancer and one of the most lethal malignancies with a 5-year survival rate under 5% for all stages taken together [1], and its mortality is projected to surpass that of breast and colorectal cancers by 2030 in the United States [2]. PDAC mostly develops along a dysplasia-carcinoma sequence, characterized by acinar-to-ductal metaplasia followed by progression of pancreatic intraepithelial neoplasia (PanIN) from low (I) to high (II–III) grade until the emergence of an invasive adenocarcinoma. Genetic alterations in human PDAC occur in concordance to this progression sequence: KRAS mutation (70–90% of cases) is already detectable in PanIN grade 1, followed by TP53 (60–70%), CDKN2A (>50%) and SMAD4 (50%) at later stages (PanIN grade 2–3, PDAC) [3,4]. Thus, a relatively small and well-characterized set of signaling pathways is involved in pancreatic carcinogenesis. However, targeting those common signaling pathways for therapy is extremely challenging and most of the targeted therapies used in other cancers failed to show a benefit in PDAC

patients. New therapeutic approaches are urgently needed, and targeting epigenetic and posttranslational abnormalities involved in PDAC progression appears to be a promising alternative to conventional therapies. Several acetylation and methylation signaling inhibitors are currently studied in preclinical and clinical phases in different cancers. In this review, we will present the rationale and the therapeutic potential of targeting lysine methylation signaling in PDAC, based on the recent findings and the potent treatments already under clinical trials in other cancers.

^aInstitute for Advanced Biosciences, 'Molecular Pathologies of Cancer and Biomarkers' Lab, INSERM U1209 – CNRS UMR5309, Université Grenoble Alpes, ^bClinique Universitaire d'Hépatogastroentérologie, Pôle Digidune, CHU Grenoble-Alpes and ^cFaculté de médecine, Université Grenoble Alpes, Grenoble, France

Correspondence to Nicolas Reynoird, Institute for Advanced Biosciences, 'Molecular Pathologies of Cancer and Biomarkers' Lab, CNRS UMR5309 – INSERM U1209 – Grenoble Alpes University, Site santé, allée des Alpes, 38700 la Tronche, France. Tel: +33 4 76 54 95 76; e-mail: Nicolas.reynoird@univ-grenoble-alpes.fr

Curr Opin Oncol 2017, 29:000–000

DOI:10.1097/CCO.0000000000000421

Cancer biology

KEY POINTS

- PDAC emerges as one of the deadliest cancers with very few and modestly efficient therapeutic options.
- Epigenetic and PTM signalings carry high hopes for the development of new targeted therapies in cancer.
- Because of high specificity and reversibility, lysine methylation signaling emerges as one favorable therapeutic possibility for PDAC.
- EHMT2, EZH2, DOT1L, SMYD2 and SMYD3 lysine methyltransferases inhibition have already shown interesting preclinical results in PDAC.

Methylation signaling in cell homeostasis

Methylation is a covalent posttranslational modification (PTM) that contributes to all aspects of cell physiology and is a primary source of functional diversity in mammalian cells [5,6]. This event is catalyzed by methyltransferases, which recruit a methyl moiety from the cofactor S-adenosyl-L-methionine (SAM) and add it onto a specific substrate. There are approximately 200 methyltransferases in human, and SAM is the second highest abundant cofactor in cell after adenosine triphosphate, which suggests an essential role for methylation signaling in cellular biology. Its deregulation is a major contributor to cancer [7], as methylation

participates to the epigenetic regulation of genes expression but also directly regulates multiple cellular pathways by affecting the three building blocks in cellular organisms: DNA, RNA and proteins. DNA methylation (5-methylcytosine) occurs preferentially on CpG islands, frequently contained within gene promoters and leads to gene repression. Hence, hypermethylation of tumor-suppressor genes is a common feature in cancer. RNA methylation (mainly N6-methyladenosine) affects all kinds of RNA, impacting RNA stability and mRNA translation, and its deregulation has recently been linked to various diseases [8]. Finally, protein methylation occurs on two different amino acids: arginine and lysine. Protein methylation is a complex and finely tuned signaling, as arginine residues can either be monomethylated or symmetrically and asymmetrically dimethylated, and lysine residues can either be monomethylated, dimethylated or trimethylated.

Lysine methylation is a dynamic and reversible modification promoting specific signaling, thanks to three different actors. Two highly regulated sets of enzymes, called lysine methyltransferases (KMTs) and lysine demethylases (KDMs), respectively, catalyze the addition and removal of lysine methylation (Fig. 1) [9]. A third set of proteins can transduce a specific response for each methylation event, by direct recognition of the modified lysine thanks to specific methyl-binding domains (MBDs) in a manner specified by the extent and context of the methylation event [10]. The importance of lysine

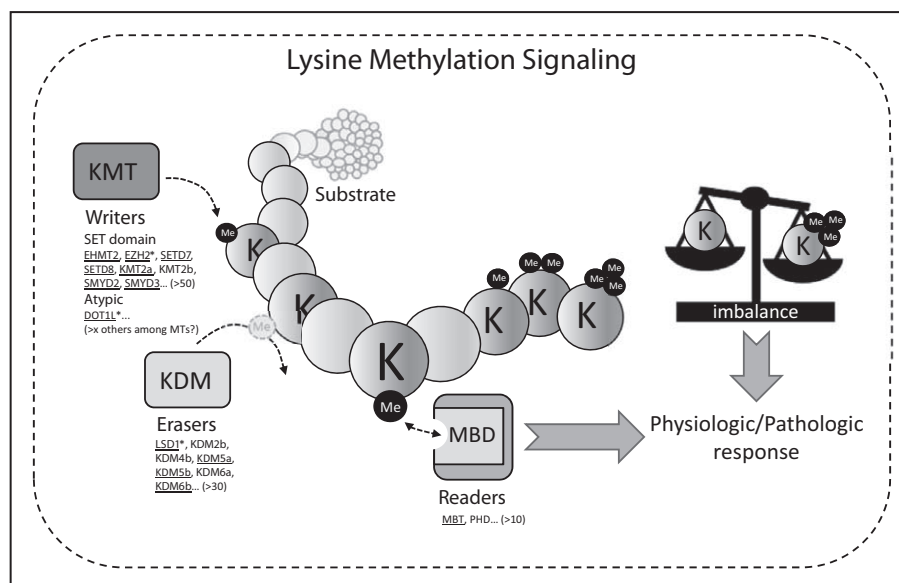


FIGURE 1. Overview of lysine methylation signaling and its actors. Representation of lysine methylation signaling and its actors. A few specific lysine methyltransferases (KMTs), lysine demethylases (KDMs) and methyl-binding proteins (MBD) are given as example, and written accordingly to specific criteria: in *Italic*, enzymes involved in pancreas physiology; in **bold**, enzymes involved in pancreatic cancer; underlined, enzymes with existing inhibitors; with asterisk *, enzymes inhibition in clinical trial.

methylation can be estimated by the number of its actors compared to other PTMs. First, among the approximately 200 methyltransferases in human, at least 50 may methylate lysines. More than 50 methyltransferases harbor the typical lysine methylation catalytic SET domain (Su(var)3-9, Enhancer-of-zeste and Trithorax), and some putative methyltransferases without SET domain have also been found to methylate lysines. Moreover, there are more than 30 estimated lysine demethylases in human, and at least 11 different MBDs have been characterized and are found in hundreds of different proteins. Thus, lysine methylation signaling has all the assets to have a broad action on cellular homeostasis. Lysine methylation is a major contributor of the so-called histone code, known to regulate chromatin state and genes expression. For instance, PR/SET domain 9 prevails an active chromatin state and gene expression by trimethylating the lysine 4 of histone H3, whereas the trimethylation of lysine 27 of H3 catalyzed by enhancer of zeste homolog 2 (EZH2) and the Polycomb repressive complex 2 represses gene expression and promotes heterochromatin formation. However, there are growing evidences of nonhistone proteins methylation events, including several proteins with clear roles in cancer such as p53, RB, RelA or MAP3K2 [11–14]. It is likely that many proteins involved in the pathogenesis of cancer are methylated and the signaling system defined by lysine methylation has recently emerged as a potential key regulator of tumorigenesis pathways, as demonstrated by the increasing number of KMT alterations identified in cancer [15,16]. However, despite the fundamental role for this modification in biology and cancer, the catalytic activity and substrate specificity for the majority of KMT/KDM enzymes is poorly described.

Lysine methylation signaling in pancreas physiology

Even though the principle and mechanisms of lysine methylation signaling involved in major cellular processes are shared among different cell types of an organism, several studies identified a specific importance of KMTs and KDMs activity in pancreas physiology. Both lysine methyltransferases EZH2 and SET Domain-Containing Protein 7 (SETD7) have been linked to insulin-secreting β cells progenitor identity. EZH2 is the catalytic subunit of the PRC2 complex, known to methylate lysines 9 and 27 of histone H3 (H3K9me and H3K27me), leading to the transcriptional repression of specific genes. EZH2 and H3K27me3 are required for embryonic stem cell identity and proper differentiation. During embryonic development, cells from the endoderm

of the foregut will differentiate in multipotent progenitors expressing Pdx1 (pancreatic and duodenal homeobox), a transcription factor essential to pancreatic development which gives rise to ductal cells, acinar cells and endocrine progenitors. These endocrine progenitors differentiate into all five hormone-expressing endocrine lineages (α , β , δ , ϵ and pancreatic polypeptide) which constitute the pancreatic or Langerhans islets. The decrease in EZH2 expression leads to a loss of the repressive H3K27 methylation at the *Pdx1* regulatory elements in endoderm, promoting the pancreas lineage [17]. Moreover, in-vitro inhibition of EZH2 in human ESC cultures induces an increase in the yield of functional β -cell progenitors [18].

Regarding SETD7, the enzyme is expressed in the pancreatic endoderm of *Xenopus* and mouse embryos prior to Pdx1 induction. SETD7 modulates methylation marks at pancreatic regulatory regions and is required and sufficient for pancreatic cell fate specification, notably β cells [19]. In addition, SETD7 regulates Pdx1 in two ways. First, active chromatin state induced by SETD7 trimethylation of lysine 4 of histone 3 (H3K4me3) promotes Pdx1 expression, and second, SETD7 directly increases Pdx1 transcriptional activity by methylating its lysine 131, as well as promoting H3K4 trimethylation at its target insulin and islet amyloid polypeptide gene promoters [20,21]. Finally, SETD7 has been shown to contribute to proinflammatory cytokines signaling in β cells [22]. A third KMT, euchromatic histone lysine methyltransferase 2 or G9a has been identified to play a role in pancreas physiology, as EHMT1/EHMT2 inhibitors selectively upregulate the cholesterol biosynthetic pathway in pancreatic but not liver cells [23].

On the contrary, not much is known about the impact of lysine demethylases in pancreas physiology. Consistent with the negative impact of EZH2 activity on H3K27 for pancreatic specification, H3K27me3 demethylases KDM6A and KDM6B play an important role in endoderm differentiation and have been shown to regulate the wntless-related integration site (WNT) signaling pathway through the activation of WNT3 and DKK1 [24]. Moreover, a recent study further demonstrates the relation of KDM6A and KDM6B with Langerhans islets integrity, observing that their pharmacologic inhibition protects insulin-producing β cells from inflammation-induced dysfunction and death [25].

Because of the emerging evidences that lysine methylation is involved in tumorigenesis pathways and the increasing number of KMT alterations identified in different cancers, it is now clear that lysine methylation signaling can play an important role in pancreatic tumorigenesis.

Cancer biology

Lysine methylation signaling in pancreatic cancer

Apart from driver oncogene mutations, cancer initiation and progression have been linked to various epigenetic mechanisms and aberrant PTM signalings. Pancreatic cancer is no exception, and identification of defects in DNA methylation of specific gene promoters as well as alterations in RNA and histone modifications provides some evidences that epigenetic mechanisms participate in PDAC development [26]. Two major observations have brought histone methylations into researchers' focus: the patterns of several histone methyl marks are altered in malignancies, and importantly, mutations in specific histone methyltransferases, demethylases and associated factors have been reported in many cancers, including pancreatic cancer.

Even though EZH2 and SETD7 play an important role in pancreatic islets integrity, there is no evidence of their implication in the emergence of pancreatic neuroendocrine tumors (PanNETs), a rare type of tumor derived from the diffuse pancreatic neuroendocrine system [27]. However, two other KMTs, KMT2A and KMT2B, seem to prevent cell proliferation observed in mouse PanNET. Their inactivation accelerates Langerhans islets tumorigenesis in Men1 knockout mice, a model recapitulating the early stage of human PanNET development [28]. Regarding PDAC, EZH2 seems to be the major culprit for now, but its oncogenic mechanism is not clear. EZH2 is upregulated following cerulein-induced pancreatic injury and promotes the regenerative proliferation of progenitor cells, whereas loss of EZH2 results in an impaired pancreatic regeneration and accelerates Kras (G12D)-driven neoplasia [29]. Overexpression of EZH2 is also correlated with a poor survival in various cancers, as well as tumor progression and metastasis [30]. This suggests a finely tuned role of EZH2 to balance between chronic pancreatitis regeneration and PDAC development, and the need of a highly regulated level of EZH2, as both loss and overexpression of EZH2 could have negative impact depending on the cellular environment. Moreover, it is highly probable that EZH2 oncogenic activity in PDAC is also related to nonhistone methylated substrates. For example, EZH2 trimethylates K180 of STAT3, promoting STAT3 transcription factor activity in glioblastoma stem-like cells [31]. Thus, a better characterization of EZH2 oncogenic activity in PDAC is required.

Lysine methylation of nonhistone substrates is an emerging field, with clear implications in human tumorigenesis [32]. Indeed, we recently identified two oncogenic lysine methylation signalings in PDAC, through the characterization of the lysines methyltransferases SMYD2 and SMYD3 (SET and

MYND domain containing 2/3). We found that SMYD2 is highly expressed in PDAC and that SMYD2 loss in mice impedes PDAC growth [33]. One of the major non-histone substrates identified for SMYD2 is p53, however we observed that SMYD2 oncogenic activity is p53-independent. We discovered that SMYD2 regulates the stress response kinase MAPKAPK3 through monomethylation of its lysine 355, leading to the production of inflammatory cytokines such as interleukin 6 (IL-6) and tumor necrosis factor α (TNF α) [34]. Chronic inflammation is a key trigger for pancreas tumorigenesis, and both SMYD2 and MAPKAPK3 were previously shown to be involved in IL-6 and TNF α production [35]. Concerning SMYD3, a related KMT with a clear overexpression in many cancers including PDAC, we observed that its catalytic activity is required for proper progression of Ras-induced cancers such as PDAC. By trimethylating the kinase MAP3K2 on its lysine 260, SMYD3 potentiates the Ras/Raf/MEK/ERK signaling pathway in cancer cells. Indeed, this methylation event blocks the interaction of MAP3K2 with the phosphatase complex PP2A, a key negative regulator of the kinase, leading to the constitutive activity of the methylated population of MAP3K2 [14,36].

It is highly probable that other lysine methylation signalings are involved in PDAC progression. Using public PDAC transcriptomic datasets, we analyzed the expression level of all the known SET-domain KMTs as well as other methyltransferases with potential lysine methylation activity. Concordant with previous data, we identified SMYD3 and EZH2 enzymes as one of the most overexpressed KMTs in PDAC compared to normal pancreas (Fig. 2). Interestingly, other methyltransferases emerged from this analysis, such as PRDM1 highly overexpressed in PDAC, while it is considered as a tumor-suppressor and is usually repressed or mutated in other cancers. On the other end of the scope, several KMTs are clearly repressed in PDAC, such as KMT2B and SETD5. Although the low expression of KMT2B is concordant with its capacity to negatively regulate cell proliferation in PanNET [28], no correlation between SETD5 and cancer has been demonstrated to date. Only few KMTs have been characterized so far, and more comprehensive studies are required to attest their potential tumor suppressor or oncogenic activity.

As lysine methylation signaling is a dynamic and reversible process, the characterization of lysine demethylases' role in PDAC is also essential. Lysine-specific demethylase 1A (LSD1) is one of the best characterized KDMs and demethylates active and repressive histone marks (H3K4me1/2 and H3K9me1/2, respectively) as well as nonhistone

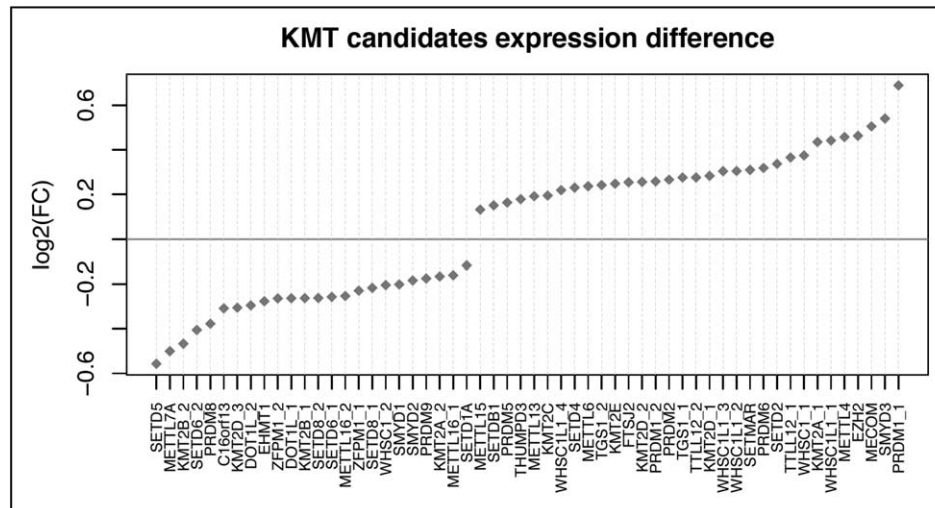


FIGURE 2. KMTs expression fold change between PDAC and normal pancreas. Distribution of log₂ (fold change) values for KMT candidates with Benjamini–Hochberg false discovery rate. *P*-value less than 5.10⁻² in PDAC vs. normal pancreas contrast. KMT candidates are sorted from lowest log₂ (fold change) to highest. Data obtained from GSE 22780, GSE 15471 and GSE 32676 manually selected for their clinical relevance, totalizing 72 pancreas cancer patients and 54 controls and analysis of 262 probes corresponding to transcripts of established and candidate KMTs. Thirteen different KMTs present multiple specific probes identified with significant expression fold changes. As we chose to avoid possible bias in biological significance in case of splicing event, we labeled the multiple probes with enumeration after KMT name.

substrates prone to regulate tumorigenesis such as p53, E2F1 and DNMT1 [37–39]. In PDAC, overexpression of LSD1 is frequent and correlates with poor prognosis, and its loss represses proliferation of pancreatic cancer cells [40]. Because of its large number of substrates deregulated in cancer, LSD1 is probably involved in multiple oncogenic mechanisms. For instance, Qin *et al.* [40] suggest a synergistic effect of LSD1 and hypoxia-induced factor1 α , promoting the glycolytic process which will fuel pancreatic cancer proliferation. Other interesting demethylases are described as KDM2B, known to demethylate active chromatin marks such as H3K4me3 and H3K36me2, whose level of expression correlates with PDAC progression [41]; KDM4B, which participates in the epithelial-to-mesenchymal transition (EMT) in PDAC [42]; and KDM6B, involved in H3K27 demethylation and whose expression decreases as the malignant grade of PDAC progresses [43].

These different observations underline the importance of the regulation of lysine methylation signaling in PDAC initiation and progression, and identify potential new relevant biomarkers and new therapeutic targets.

New therapeutic approaches in pancreatic ductal adenocarcinoma based on lysine methylation signaling

Current efforts are engaged to target aberrant DNA methylation or histone acetylation signaling [43].

Inhibitors of histone deacetylases (HDACs) and bromodomain and extra-terminal motif proteins already showed promising preliminary results in several hematological malignancies and solid tumors [44,45]. On the contrary, lysine methylation signaling as a therapeutic target is at a less advanced stage, mostly because of its higher specificity of mechanism. Current attempts to regulate DNA methylation or histone acetylation in cancer aim at global epigenomic switch that could be more prone to side-effects or mitigated results, because of the complexity and imbalance between subsets of genes to reactivate or repress in order to tackle a specific cancer. However, targeting lysine methylation signaling could prove to be more efficient, or at least more specific than other current epigenetic therapies. For example, a study carried in pancreatic endocrine cells to study global transcriptomic changes upon chromatin-targeting small molecules demonstrated that while HDAC inhibitors regulate several hundreds of transcripts irrespective of the cell type, compounds targeting histone methyltransferases modulate the expression of restricted gene sets in distinct cell types [23].

Global inhibitors affecting similarly DNA, RNA and protein methyltransferases exist, such as 3-deazaneplanocin A (DZNep), an inhibitor of the S-adenosyl-l-homocysteine hydrolase and sinefungin, a SAM analogue, whereas another one, AMI-1, is a protein arginine N-methyltransferase inhibitor. However, as each KMT and KDM seems to have a

Cancer biology

restricted number of highly specific substrates and can play contradictory functions depending on cellular environment, the clinical relevance of these global inhibitors is questionable. Thus, a particular effort has recently been made to develop specific KMT and KDM inhibitors with a hope for future epigenetic therapy [46,47]. Many KMTs and KDMs have inhibitors usable for research, but few are suitable for a clinical purpose yet. Current specific KMT inhibitors include compounds targeting the histone methyltransferase DOT1L (disruptor of telomeric silencing 1-like), EZH2, EHMT1 and 2, SETD7 and SMYD2 as presented in the review by McGrath and Trojer [46]. Few novelties were developed since, such as inhibitors of SMYD3 and SETD8, as well as less specific compounds preferentially targeting SUV39H1 and SETD2 [48]. Finally, a different approach to inhibit KMTs has been recently developed, by blocking the interaction of the enzyme within the core complex required for its activity. For example, MM-102 compound inhibits the H3K4 methyltransferase KMT2a also known as MLL1 by blocking its interaction with WDR5 (WD repeat-containing protein 5), a subunit essential for its methylation of H3K4 [49], and MAK683 compound inactivates EZH2 through the inhibition of the embryonic ectoderm development protein EED, an essential subunit of the polycomb repressive complex 2 [50].

Because of direct evidences of DOT1L and EZH2 implication in specific types of cancer, pharmaceutical compounds targeting these KMTs have been developed and are currently being studied in clinical trials. Thus, several EZH2 inhibitors such as EPZ-6438 (Tazemetostat), MAK683, CPI-1205 and GSK2816126 are evaluated in 12 clinical studies in advanced solid tumors and lymphomas. Also, DOT1L inhibitor EPZ-5676 (Pinometostat) has been evaluated in two phase-1 studies in leukemia, completed in 2016 with only preliminary results available to date, suggesting an acceptable safety with a confirmation of DOT1L inhibition in tumors [51].

Regarding the development of lysine demethylase inhibitors, fewer compounds have been designed: GSK-J1 compound targets H3K27me3 demethylase KDM6B [52] and EPT103182 inhibits H3K4me3 demethylases KDM5a and KDM5b. A particular effort has been shown to target LSD1, with the development of several inhibitors such as GSK2879552, ORY-1001, IMG-7289, 2-PCPA (tranylcypromine) and 4SC-202 which are evaluated in several phase I and I/II clinical trials in various cancers (notably malignant melanoma, acute myeloid leukemia and small cell lung cancer).

Finally, another way to target lysine methylation signaling is to block the specific recognition of

methyl-lysine motifs by proteins harboring a MBD, leading to the loss of signal transduction into appropriate biological responses. This mechanism is thoroughly under evaluation for acetylation signaling with bromodomain inhibitors, but only one specific compound has been developed for methylation. UNC1215 is a potent, selective antagonist of the triple malignant brain tumor domain of L3MBTL3, a domain with broad binding capacity over mono and dimethylated lysines [53]. Even though none of these exciting interfering compounds have been tested for pancreatic cancer in clinical trials, several preclinical studies point to the high interest that these targeted therapies may represent for PDAC patients.

Because of the multiple escape pathways that cancers develop to resist under anticancer treatments, combination of cytotoxic chemotherapies with one or several targeted therapies is often required. Even if no therapeutic association showed a benefit for PDAC patients to date, new combinations are being tested to improve PDAC treatment. For instance, regulation of histone methylation by DZNep affects the expression of important miRNA involved in EMT and metastasis in pancreatic cancer. This compound has been shown to modulate EZH2 activity and H3K27me3 while it synergistically enhances the antiproliferative activity of gemcitabine [54,55]. However, further characterization of DZNep has shown that it is not only an EZH2 inhibitor but rather a global histone methylation inhibitor, affecting other methylation marks such as H4K20me3. Nevertheless, another study confirmed that EZH2 depletion by siRNA sensitizes pancreatic cancer cells to cytotoxic agents such as doxorubicin and gemcitabine and leads to a significant induction of apoptosis [56]. These works suggest that the combination of EZH2 inhibitors with standard cytotoxic chemotherapies could offer new therapeutic solutions for patients with pancreatic cancer.

Our laboratory observed in mouse xenografts of human PDAC that pharmacological inhibition of SMYD2 by BAY-598 significantly reduces cancer growth when combined with standard chemotherapy [33^{*}]. We also demonstrated that combined inhibition of SMYD2 and its substrate MAPKAPK3 strongly impeded the growth of PDAC xenografts in mice. Likewise, we observed that deletion of SMYD3 also impairs PDAC development in a mouse model and potentiates the antitumor activity of Trametinib, an MEK1/2 kinase inhibitor already used to treat metastatic melanoma [14]. Furthermore, deletion of both SMYD2 and SMYD3 in mice has no visible phenotype, suggesting that specific inhibitors of these enzymes would have only minimal collateral toxicity. Both studies suggest a possible combinatorial strategy that

might prevent treatment toxicity by lowering dosages and reduce drug resistance by simultaneously targeting several oncogenic actors.

Finally, targeting methylation may also directly prevent drug resistance of cancers, as KMTs, such as G9a, seem to be involved in this mechanism. Indeed, PDAC cell lines resistant to gemcitabine overexpress G9a, and this resistance is mediated by G9a-dependant IL-8 stimulation. Both IL-8 and G9a inhibition seem to override gemcitabine resistance in pancreatic cancer cells [57].

CONCLUSION

So far, PDAC has resisted most preventive, diagnostic and therapeutic challenges. Despite intensive research on biomarkers and new therapeutic approaches, clinical results are still modest. There is today an unprecedented and diversified offer of new therapeutic options targeting the epigenome and PTMs such as lysine methylation. Remarkably, some lysine methyltransferase inhibitors have shown exciting preclinical effects in PDAC animal models, but there is still a long way from these models to clinical demonstration. However, the very fact that inhibition of lysine methylation signaling is already under clinical trial in other cancer types is a major milestone toward the development of successful new therapies for PDAC.

Acknowledgements

None.

Financial support and sponsorship

This work was supported in part by grants from the Agence National pour la Recherche (ANR JCJC), La Ligue contre le Cancer (Comité Isère Ligue) and Initiatives de Recherches Stratégiques (UGA IDEX).

G.R. is a recipient of a 'poste d'accueil INSERM' doctoral fellowship. A.C. is a recipient of Fondation pour la Recherche Médicale doctoral fellowship (FRM grant number ECO20170637491 to A.C.). N.L. is a recipient of a postdoctoral fellowship from the French National Research Agency in the framework of the 'Investissements d'avenir' program (ANR-15-IDEX-02).

Conflicts of interest

The authors claim no conflicts of interest.

REFERENCES AND RECOMMENDED READING

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Hidalgo M. New insights into pancreatic cancer biology. *Ann Oncol* 2012; 23(Suppl 10):x135–x138.

2. Rahib L, Smith BD, Aizenberg R, *et al.* Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res* 2014; 74:2913–2921.
3. Waddell N, Pajic M, Patch A-M, *et al.* Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature* 2015; 518:495–501.
4. Ryan DP, Hong TS, Bardeesy N. Pancreatic adenocarcinoma. *N Engl J Med* 2014; 371:2140–2141.
5. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Res* 2011; 21:381–395.
6. Sims RJ, Reinberg D. Is there a code embedded in proteins that is based on posttranslational modifications? *Nat Rev Mol Cell Biol* 2008; 9:815–820.
7. Xhemalce B. From histones to RNA: role of methylation in cancer. *Brief Funct Genomics* 2013; 12:244–253.
8. Frye M, Jaffrey SR, Pan T, *et al.* RNA modifications: what have we learned and where are we headed? *Nat Rev Genet* 2016; 17:365–372.
9. Dillon SC, Zhang X, Trievel RC, Cheng X. The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol* 2005; 6:227.
10. Taverna SD, Li H, Ruthenburg AJ, *et al.* How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* 2007; 14:1025–1040.
11. Huang J, Perez-Burgos L, Placek BJ, *et al.* Repression of p53 activity by Smyd2-mediated methylation. *Nature* 2006; 444:629–632.
12. Saddic LA, West LE, Aslanian A, *et al.* Methylation of the retinoblastoma tumor suppressor by SMYD2. *J Biol Chem* 2010; 285:37733–37740.
13. Levy D, Kuo AJ, Chang Y, *et al.* Lysine methylation of the NF- κ B subunit RelA by SETD6 couples activity of the histone methyltransferase GLP at chromatin to tonic repression of NF- κ B signaling. *Nat Immunol* 2011; 12:29–36.
14. Mazur PK, Reynoird N, Khatri P, *et al.* SMYD3 links lysine methylation of MAP3K2 to Ras-driven cancer. *Nature* 2014; 510:283–287.
15. Schneider R, Bannister AJ, Kouzarides T. Unsafe SETs: histone lysine methyltransferases and cancer. *Trends Biochem Sci* 2002; 27:396–402.
16. Morin RD, Mendez-Lago M, Mungall AJ, *et al.* Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* 2011; 476:298–303.
17. Xu C-R, Cole PA, Meyers DJ, *et al.* Chromatin 'prepattern' and histone modifiers in a fate choice for liver and pancreas. *Science* 2011; 332:963–966.
18. Xu C-R, Li L-C, Donahue G, *et al.* Dynamics of genomic H3K27me3 domains and role of EZH2 during pancreatic endocrine specification. *EMBO J* 2014; 33:2157–2170.
19. Kofent J, Zhang J, Spagnoli FM. The histone methyltransferase Setd7 promotes pancreatic progenitor identity. *Development* 2016; 143:3573–3581.
20. Maganti AV, Maier B, Tersey SA, *et al.* Transcriptional activity of the islet β cell factor Pdx1 is augmented by lysine methylation catalyzed by the methyltransferase Set7/9. *J Biol Chem* 2015; 290:9812–9822.
21. Ogiwara T, Vanderford NL, Maier B, *et al.* Expression and function of Set7/9 in pancreatic islets. *Islets* 2009; 1:269–272.
22. Fujimaki K, Ogiwara T, Morris DL, *et al.* SET7/9 enzyme regulates cytokine-induced expression of inducible nitric-oxide synthase through methylation of lysine 4 at histone 3 in the islet β cell. *J Biol Chem* 2015; 290:16607–16618.
23. Kubicek S, Gilbert JC, Fomina-Yadlin D, *et al.* Chromatin-targeting small molecules cause class-specific transcriptional changes in pancreatic endocrine cells. *Proc Natl Acad Sci U S A* 2012; 109:5364–5369.
24. Jiang W, Wang J, Zhang Y. Histone H3K27me3 demethylases KDM6A and KDM6B modulate definitive endoderm differentiation from human ESCs by regulating WNT signaling pathway. *Cell Res* 2013; 23:122–130.
25. Backe MB, Andersson JL, Bacos K, *et al.* Lysine demethylase inhibition protects pancreatic β cells from apoptosis and improves β -cell function. *Mol Cell Endocrinol* 2017; In press.
26. McCleary-Wheeler AL, Lomber GA, Weiss FU, *et al.* Insights into the epigenetic mechanisms controlling pancreatic carcinogenesis. *Cancer Lett* 2013; 328:212–221.
27. Modlin IM, Oberg K, Chung DC, *et al.* Gastroenteropancreatic neuroendocrine tumours. *Lancet Oncol* 2008; 9:61–72.
28. Lin W, Francis JM, Li H, *et al.* Kmt2a cooperates with menin to suppress tumorigenesis in mouse pancreatic islets. *Cancer Biol Ther* 2016; 17:1274–1281.
29. Mallen-St. Clair J, Soydaner-Azeloglu R, Lee KE, *et al.* EZH2 couples pancreatic regeneration to neoplastic progression. *Genes Dev* 2012; 26:439–444.
30. Völkel P, Dupret B, Le Bourhis X, Angrand P-O. Diverse involvement of EZH2 in cancer epigenetics. *Am J Transl Res* 2015; 7:175–193.
31. Kim E, Kim M, Woo D-H, *et al.* Phosphorylation of EZH2 activates STAT3 signaling via STAT3 methylation and promotes tumorigenicity of glioblastoma stem-like cells. *Cancer Cell* 2013; 23:839–852.
32. Hamamoto R, Saloura V, Nakamura Y. Critical roles of nonhistone protein lysine methylation in human tumorigenesis. *Nat Rev Cancer* 2015; 15:110–124.
33. Reynoird N, Mazur PK, Stellfeld T, *et al.* Coordination of stress signals by the lysine methyltransferase SMYD2 promotes pancreatic cancer. *Genes Dev* 2016; 30:772–785.

The authors demonstrate the implication of the lysine methyltransferase SMYD2 in PDAC progression and provide preclinical evidence of the therapeutic value to target its signaling.

Cancer biology

34. Wysk M, Yang DD, Lu H-T, *et al.* Requirement of mitogen-activated protein kinase kinase 3 (MKK3) for tumor necrosis factor-induced cytokine expression. *Proc Natl Acad Sci U S A* 1999; 96:3763–3768.
35. Xu G, Liu G, Xiong S, *et al.* The histone methyltransferase Smyd2 is a negative regulator of macrophage activation by suppressing interleukin 6 (IL-6) and tumor necrosis factor α (TNF- α) production. *J Biol Chem* 2015; 290:5414–5423.
36. Mazur PK, Gozani O, Sage J, Reynold N. Novel insights into the oncogenic function of the SMYD3 lysine methyltransferase. *Transl Cancer Res* 2016; 5:330–333.
37. Xie Q, Bai Y, Wu J, *et al.* Methylation-mediated regulation of E2F1 in DNA damage-induced cell death. *J Recept Signal Transduct* 2011; 31:139–146.
38. Huang J, Sengupta R, Espejo AB, *et al.* p53 is regulated by the lysine demethylase LSD1. *Nature* 2007; 449:105–108.
39. Wang J, Hevi S, Kurash JK, *et al.* The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nat Genet* 2009; 41:125–129.
40. Qin Y, Zhu W, Xu W, *et al.* LSD1 sustains pancreatic cancer growth via maintaining HIF1 α -dependent glycolytic process. *Cancer Lett* 2014; 347: 225–232.
41. Tzatsos A, Paskaleva P, Ferrari F, *et al.* KDM2B promotes pancreatic cancer via Polycomb-dependent and -independent transcriptional programs. *J Clin Invest* 2013; 123:727–739.
42. Li S, Wu L, Wang Q, *et al.* KDM4B promotes epithelial-mesenchymal transition through up-regulation of ZEB1 in pancreatic cancer. *Acta Biochim Biophys Sin* 2015; 47:997–1004; gmv107.
43. Yamamoto K, Tateishi K, Kudo Y, *et al.* Loss of histone demethylase KDM6B enhances aggressiveness of pancreatic cancer through downregulation of C/EBP α . *Carcinogenesis* 2014; 35:2404–2414.
44. Hessmann E, Johnsen SA, Siveke JT, Ellenrieder V. Epigenetic treatment of pancreatic cancer: is there a therapeutic perspective on the horizon? *Gut* 2017; 66:168–179.
45. Mazur PK, Herner A, Mello SS, *et al.* Combined inhibition of BET family proteins and histone deacetylases as a potential epigenetics-based therapy for pancreatic ductal adenocarcinoma. *Nat Med* 2015; 21:1163–1171.
46. McGrath J, Trojer P. Targeting histone lysine methylation in cancer. *Pharmacol Ther* 2015; 150:1–22.
47. Morera L, Lübbert M, Jung M. Targeting histone methyltransferases and demethylases in clinical trials for cancer therapy. *Clin Epigenetics [Internet]* 2016; 8:57; Dec [cited 2017 Sep 6];8(1). Available from: <http://clinicaepigeneticsjournal.biomedcentral.com/articles/10.1186/s13148-016-0223-4>.
48. Kaniskan HÜ, Konze KD, Jin J. Selective inhibitors of protein methyltransferases. *J Med Chem* 2015; 58:1596–1629.
49. Karatas H, Townsend EC, Cao F, *et al.* High-affinity, small-molecule peptidomimetic inhibitors of MLL1/WDR5 protein–protein interaction. *J Am Chem Soc* 2013; 135:669–682.
50. Kim W, Bird GH, Neff T, *et al.* Targeted disruption of the EZH2–EED complex inhibits EZH2-dependent cancer. *Nat Chem Biol* 2013; 9:643–650.
51. Shukla N, Wetmore C, O'Brien MM, *et al.* Final report of phase 1 study of the DOT1L inhibitor, pinometostat (EPZ-5676), in children with relapsed or refractory MLL-r acute leukemia. *Blood* 2016; 128:2780.
52. Kruidenier L, Chung C, Cheng Z, *et al.* A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. *Nature* 2012; 488:404–408.
53. James LI, Barsyte-Lovejoy D, Zhong N, *et al.* Discovery of a chemical probe for the L3MBTL3 methyllysine reader domain. *Nat Chem Biol* 2013; 9:184–191.
54. Avan A, Crea F, Paolicchi E, *et al.* Molecular mechanisms involved in the synergistic interaction of the EZH2 inhibitor 3-deazaneplanocin a with gemcitabine in pancreatic cancer cells. *Mol Cancer Ther* 2012; 11:1735–1746.
55. Mody HR, Hung SW, AlSaggar M, *et al.* Inhibition of S-adenosylmethionine-dependent methyltransferase attenuates TGF 1-induced EMT and metastasis in pancreatic cancer: putative roles of miR-663a and miR-4787-5p. *Mol Cancer Res* 2016; 14:1124–1135.
56. Ougolkov AV, Bilim VN, Billadeau DD. Regulation of pancreatic tumor cell proliferation and chemoresistance by the histone methyltransferase enhancer of zeste homologue 2. *Clin Cancer Res* 2008; 14:6790–6796.
57. Pan M-R, Hsu M-C, Luo C-W, *et al.* The histone methyltransferase G9a as a therapeutic target to override gemcitabine resistance in pancreatic cancer. *Oncotarget* 2016; 7:61136–61151.