UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE des Sciences de la Vie et de la Santé (ED 414)

UMR 7242 - Biotechnologie et Signalisation Cellulaire

THÈSE présentée par :

Andrea ARGÜESO LLEIDA

soutenue le : **19 Septembre 2018**

pour obtenir le grade de : **Docteur de l'Université de Strasbourg**

Discipline/ Spécialité : Aspects moléculaires et cellulaires de la Biologie

Développement d'approches de modifications ciblées du méthylome dans les cellules mammifères

THÈSE dirigée par :

Dr WEBER Michaël Directeur de Recherche, UMR 7242 Biotechnologie et Signalisation Cellulaire

RAPPORTEURS :

Dr ARIMONDO Paola Directrice de Recherche, Institut Pasteur

Dr FRANCASTEL Claire Directrice de Recherche, Université Paris 7 Diderot, UMR 7216

AUTRES MEMBRES DU JURY :

Dr GAIDDON Christian Directeur de Recherche, U1113 INSERM Strasbourg

REMERCIEMENTS

Je tiens tout d'abord à remercier les docteurs **Paola Arimondo** et **Claire Francastel** pour avoir accepté de juger mon travail de thèse en tant que rapporteures de ce manuscrit. Je remercie également le Dr **Christian Gaiddon** d'avoir accepté de faire partie de mon jury de thèse. Je vous remercie par avance pour le temps que vous allez consacrer à ma thèse.

Ensuite, je veux remercier **Michaël** (le chef) de m'avoir accueillie dans son équipe durant toutes ces années. Je te serai toujours reconnaissante pour m'avoir donné ma première expérience professionnelle et m'avoir permis de faire partie de la *REIC team*. Merci de m'avoir fait confiance, cela m'a permis de gagner en assurance et d'apprendre de mes erreurs. Merci pour ton aide dans la rédaction de ce manuscrit, pour nos discussions et pour ta patience.

Je souhaite aussi remercier tous les membres de l'équipe pour leur soutien tout au long de ces années. Je pourrai toujours dire que j'ai vécu les deux versions REIC : version beta et version 2.0.

Dans un premier temps, je veux remercier Ambre, Ghislain et Elouan. Ils m'ont accueilli très chaleureusement quand je suis arrivée à Strasbourg et m'ont montré tous les secrets de la méthylation de l'ADN et de la chanson française ! **Ambre**, ma copine de paillasse, merci pour ton humour (que je ne comprends pas toujours, mais que j'apprécie quand même^^) et ton esprit de travail. Tu es un exemple pour moi, et je suis très heureuse de voir que tes projets personnels et professionnels avancent ! **Ghislain**, tu as toujours été de bon conseil et je t'en remercie ! J'admire ton courage et ton esprit entrepreneur, et je te souhaite le meilleur dans tes projets, toujours aux côtés de **Rose** ! **Elouan**, la personne la plus patiente avec moi dans ce laboratoire. Merci de m'avoir encadré pendant mon master, j'ai beaucoup ressenti ton départ vers « les experts » ! Merci pour toutes tes explications et les bons moments partagés. Je garderai de super souvenirs de cette première époque. Même si je ne comprenais pas toujours tout (bonjour, MEFs ?), je n'oublierai pas ces lab meetings entre cahiers et photos perdues…^^

Hala, tout d'abord, merci pour tous tes conseils scientifiques. Tu es arrivée au bon moment (^^). J'apprécie énormément ta patience avec moi, toutes tes explications et discussions, toujours très enrichissantes ! J'ai pu beaucoup apprendre à tes côtés. Merci également pour être une amie et prendre le rôle de psychologue quand j'en ai le plus besoin (vive les audios de 10min), tu as toujours eu des mots réconfortants. Et bien sûr merci de nous montrer le meilleur de la cuisine libanaise !

Manon, ma petite bretonne ! C'est très difficile de tomber sur des gens comme toi, tu as été un énorme support pendant cette thèse ! Merci pour ta bonne humeur, les discussions et pour toujours être disponible. On a passé des super moments pendant ces années, des bêtises, des balades (souvent accidentées^^) à vélo, PoS, des crises existentielles… Je te souhaite vraiment plein de réussite pour la fin de ta thèse, je suis sûre que tu vas conquérir le monde avec tes petites souris !

Merci à tout le reste de l'équipe pour son soutien, conseils et bons moments partagés au labo et en dehors. **Thomas**, je suis très heureuse d'avoir vu ton évolution depuis tes débuts comme petit stagiaire ! On a passé des bons moments au labo, surtout à table et l'inoubliable moment chasse-de-bestioles ! Courage pour la fin de ta thèse ;) ! **MiniMike**, merci pour toutes les analyses bioinfo. Mais surtout, merci de nous permettre de découvrir les talents cachés des REIC dans les

soirées Just Dance ! **Judith**, voisine de bureau et salvatrice du goûter ! Merci pour tous les bons moments, et surtout pour tous les capuccinos et les innombrable madeleines ! **Jitesh**, thanks for understanding my "franglais" when talking to you, I wish you all the best with your family and your future projects. **Eléa**, merci pour ta motivation et ton humeur ! Merci de m'avoir permis de faire partie de la super aventure PoS ! **Anaïs**, merci pour tous tes conseils. Je te souhaite plein de bonheur dans cette nouvelle étape avec ta famille. **Yannick**, on n'a pas eu beaucoup l'occasion de discuter, mais je te souhaite plein de réussites dans ton post-doc et dans ta vie !

Merci au reste de l'unité, ces personnes avec qui j'ai pu échanger tout au long de ces années. Toute votre aide et vos conseils ont été essentiels pour ma réussite.

Je veux spécialement mentionner les PARP ! Qu'est-ce qu'on ferait sans vous ? Merci pour votre aide et bien sûr pour les bons moments de rigolade et délire lors des repas de midi. Merci tout particulier à **Kathline** et **Leonel**, vous êtes des personnes adorables, je vous souhaite plein de bonnes choses pour la suite !

Je veux aussi remercier tous les gens d'ici ou d'ailleurs qui d'une manière ou d'une autre font partie de ma vie.

Gracias **Ana** por tu amistad y por ser un apoyo esencial aquí (y la liberación para hablar español :-)), nuestras charlas son siempre un placer! Merci à mes anciens « coolocs », les personnes avec lesquelles j'ai pu partager cet appartement si spécial pour moi ! Merci pour les bons moments, les crémaillères, les moments surréalistes, mais surtout merci pour me supporter même quand je deviens insupportable ! Mention spéciale à **Jojo** et **Morgane**. Momo, tu vas trop me manquer, je te souhaite plein de bonheur dans tes projets. Merci aux autres amis que j'ai rencontré à Strasbourg, spécialement toi **Jamie**, tu me manques.

Gracias a mis amigos en la distancia y por ello, gracias a las nuevas tecnologías. Gracias **Laura** por tu apoyo, por hacerme sentir que el tiempo no pasa y que los 28 son los nuevos 18. **Meg**, gracias por los desahogos varios (drama-queen) y por ser un ejemplo de perseverancia. **Alba**, gracias por tu apoyo durante estos años. Gracias a las personas que conocí en la Uni y que siguen y sé que seguirán ahí. **Paloma**, mi nueva funcionaria favorita, me hubiera encantado redactar esta tesis contigo en Zacut :-) te deseo lo mejor en esta nueva etapa, ¡y gracias por tu visita a Strasbourg! **Biotec team**, con vosotros empezó esta aventura, sois increíbles, me encanta ver vuestra evolución! **Luis**, gracias por tu apoyo y por tener un corazón enorme!

Por supuesto, gracias a mi familia. Gracias a "**las Lleida**" y a **Tito** por su apoyo en la distancia. Especialmente, gracias **papá** y **mamá** por apoyarme siempre en mis decisiones y por creer en mí más que yo misma. Y también por escuchar mis (demasiado frecuentes^^) quejas. Al final siempre acabo con una sonrisa. Aunque no nos veamos tanto como me gustaría, sé que siempre estáis ahí, os quiero.

Merci à toi **Pascal**. Cette thèse a le même âge que notre relation. Après m'avoir supporté tout ce temps, je crois qu'on est vraiment fait l'un pour l'autre^^. Grâce à toi ces quatre ans ont été plus faciles. On a vécu plein de choses, des expériences, des voyages… même un ouragan. Mais surtout ce que j'aime, c'est de partager avec toi mon quotidien et nos rigolades. Merci de me supporter, m'écouter même quand je parle trop :-) ou juste me faire des câlins. Je t'aime.

ABBREVIATIONS

(U)CAR T-cell therapy	(Universal) Chimeric Antigen Receptors T-cell therapy
2HG	2-hydroxyglutarate
5aza	5’-azacytidine
5azadC	5’-aza-2’-deoxycytidine
5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxmethylcytosine
5hmU	5-hydroxymethyluracil
5mC	5-methylcytosine
6mA	N6-methyladenine
AAV	Adeno-associated virus
AD	Alzheimer's disease
ADD	Protein domain : ATRX-DNMT3-DNMT3L
AID	Activation-induced cytidine deaminase (also AICDA)
ALL	Acute lymphoblastic leukaemia
AML	Acute myeloid leukaemia
APOBEC	Apolipoprotein B mRNA editing enzyme
AS	Angelman syndrome
BAH	Protein domain: Bromo-Adjacent Homology
BER	Base excision repair
bp	Base pair
BS-seq	Bisulfite sequencing
C	Cytosine
CD	Catalytic domain
cfDNA	Cell-free DNA
CFP1	CXXC Finger Protein 1
CGI	CpG island
ChIP-seq	Chromatin immunoprecipitation-sequencing
CLL	Chronic lymphocytic leukaemia
CML	Chronic myeloid leukaemia
CMML	Chronic myelomonocitic leukaemia
COBRA	Combined Bisulfite Restriction Analysis
CpG	CG dinucleotide (C-phosphate-G)
CRC	Colorectal cancer
CRISPR-Cas	Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated proteins
crRNA	crispr RNA
CSR	Class Switch Recombination
CT (CTA)	Cancer-testis (Antigens)
CTCF	CCCTC-binding Factor
CTCL	Cutaneous T-cell lymphoma
ctDNA	Circulating tumor DNA
CXXC	Protein domain rich in cysteins
DBD	DNA Binding Domain

dCas9	dead Cas9
Dox	Doxycycline
DHS	DNAse I hypersensitive sites
DMPs	Differentially methylated positions
DMR	Differentially methylated region, t- (tissue-specific), g- (germline), s- (somatic)
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
DNMTi	DNA methyltransferase inhibitors
ES cells/ESCs	Embryonic Stem cells (mES = mouse, hES = human)
ESCC	Oesophageal squamous cell carcinoma
EWAS	Epigenome-wide association study
FDA	US Food and Drug Administration
FdCyd	5-fluoro-2'-deoxycytidine
FMR1	Fragile X mental retardation 1
GFP	Green fluorescent protein
H3K27ac	Histone H3, acetyl group on lysine (K) 27
H3K27m3	Histone H3, triple methyl group on lysine (K) 27
H3K4m3	Histone H3, triple methyl group on lysine (K) 4
H3K9me2/3	Histone H3, double or triple methyl group on lysine (K) 9
HDAC	Histone deacetylase
HDR	Homology Directed DNA Repair
HMRs	Hypomethylated regions, c- (common for all cell types), t- (tissue-specific)
HPV	Human Papillomavirus
IAP	Intracisternal A-particle
ICF	Immunodeficiency, centromeric region instability and facial anomalies syndrome
ICR	Imprinting Control Region
IDH1 and 2	Isocitrate dehydrogenase 1 and 2
Ig	Immunoglobulin
IGF2	Insulin-like growth factor II
iPSCs	Induced pluripotent stem cells
KD	Knockdown
KO	Knockout
KRAB	Krüppel-associated Box
LINEs	Long interspersed nuclear elements
LMR	Low methylated region
LOI	Loss of imprinting
LTR	Long terminal repeats
mAb	Monoclonal antibodies
MBDs	Methyl-CpG binding domain
MDS	Myelodysplastic syndrome
MeCP2	Methyl-CpG binding protein 2
MEFs	Mouse Embryonic Fibroblasts
MHC	Major Histocompatibility Complex
miRNA (miR)	Micro-RNA
mRNA	Messenger RNA
ncRNA	Non-coding RNA
NHEJ	Non-Homologous End Joining DNA repair

NLS	Nuclear localisation signal
NPCs	Neuronal precursor cells
NSCLC	Non-small cell lung cancer
PAM	Protospacer Adjacent Motif
PBS	PCNA binding site
PCa	Prostate cancer
PCNA	Proliferating Cell Nuclear Antigen
PCR	Polymerase chain reaction
PGC	Primordial germs cells
PGC7	Also called Stella or Dppa3
piARN	Piwi-Interacting ARN
PMDs	Partially methylated domains
PRC	Polycomb repressive complex
PSA	Prostate-specific antigen
PTCL	Peripheral T-cell lymphoma
PWS	Prader-Willi syndrome
PWWP	Protein domain rich in prolines (P) and tryptophans (W)
RFTS	Protein domain: Replication Foci Targeting Sequence
RNA	Ribonucleic acid
RNApolII	RNA polymerase II
RRBS	Reduced Recombinated Bisulfite Sequencing
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
RVD	Repeat variable diresidues
SaCas9	*Staphylococcus aureus* Cas9
SAH	S-adenosyl-L-homocysteine
SAM	S-adenosyl-L-methionine
SERPINB5	Serine Proteinase Inhibitor, Clade B (Ovalbumin), Member 5 Also called MASPIN (Mammary Serine Protease Inhibitor)
sgRNA	Single-guide RNA
SHM	Somatic Hyper Mutation
SINEs	Short interspersed nuclear elements
SMUG1	Single-strand selective monofunctional Uracil-DNA Glycosylase 1
SpCas9	*Streptococcus pyogenes* Cas9
SRA	Protein domain: SET and RING Associated domain
ssDNA	Single-stranded DNA
T	Thymine
TALE	Transcription activator-like effectors
TALENs	TALE Nucleases
TDG	Thymine-DNA-Glycosylase
TEs	Transposable elements
TET	Ten-Eleven Translocation
TFs	Transcription factors
TR	Tandem repeat
tracrRNA	Trans-activating crRNA
TRDMT1	tRNA (cytosine-5-)-methyltransferase
TRE	Tet Response Element
tRNA	Transfer RNA
TSG	Tumor suppressor gene

TSS	Transcription Start Site
tTA/rtTA	Tetracycline-controlled transactivator / reverse tTA
UDG	Uracil-DNA-glycosylase
UHRF1	Ubiquitin-like containing PHD and RING finger domains 1, also called Np95 ou ICBP90
UNG2	Uracil-DNA-glycosylase
WGBS	Whole Genome Bisulfite Sequencing
WT	Wild Type
Xist	X-inactive specific transcript
ZFP	Zinc Finger Protein
α-KG	α-ketoglutarate

Résumé

La méthylation de l'ADN est une modification épigénétique sur les cytosines des dinucléotides CpG catalysée par les enzymes DNMT. Les cellules cancéreuses présentent des hyperméthylations aberrantes sur les promoteurs de gènes dits suppresseurs de tumeurs, ce qui contribue à leur répression transcriptionnelle et favorise la progression tumorale. De par sa nature réversible, la méthylation de l'ADN est une cible de choix pour des thérapies épigénétiques ; cependant, les inhibiteurs de DNMT ont une action de déméthylation globale du génome qui conduit à une forte toxicité. Mon travail a consisté à développer des stratégies de déméthylation ciblée sur des régions spécifiques du génome. Premièrement, j'ai validé une stratégie induisant une reprogrammation épigénétique spécifique et durable du gène suppresseur de tumeurs *SERPINB5* dans des cellules de cancer du sein. Deuxièmement, j'ai optimisé des stratégies d'édition de l'épigénome comme outil en recherche fondamentale.

Mots-clés : Méthylation de l'ADN, cancer, édition de l'épigénome, *SERPINB5*, TALE, dCas9

Abstract

DNA methylation takes place on cytosines of CpG dinucleotides in mammals and is catalysed by DNMT enzymes. Cancer cells are characterised by frequent promoter hypermethylation leading to transcriptional repression of tumor suppressor genes and favouring tumor progression. Because of its reversible nature, DNA methylation is a target of choice in epigenetic therapies. However, current DNMT inhibitors act in a global and non-specific manner, leading to side effects and toxicity in normal cells. During my thesis I have developed strategies to perform targeted demethylation in specific regions of the genome without affecting global methylation. First, I have validated a strategy inducing the specific and durable epigenetic reprogramming of the tumor suppressor gene *SERPINB5* in a breast cancer cell line, which can pave the way to further biomedical research. Second, I have optimised epigenome editing strategies as a regular tool in basic research.

Key words: DNA methylation, cancer, epigenome editing, *SERPINB5*, TALE, dCas9

RESUME DE THESE

Développement d'approches de modifications ciblées du méthylome dans les cellules mammifères

La méthylation de l'ADN est une modification épigénétique catalysée par famille des ADN méthyltransférases. Chez les mammifères il existe 3 enzymes conservées qui catalysent la méthylation de l'ADN : DNMT1, responsable du maintien des profils de méthylation au cours de la réplication, et DNMT3A et 3B, responsables de la méthylation *de novo*. Dans ce contexte, la méthylation de l'ADN a lieu sur les cytosines de dinucléotides CpG, qui se trouvent principalement dans les îlots CpG au niveau des promoteurs de gènes. La méthylation d'un îlot CpG inhibe de manière générale l'activité du promoteur.

Une fois mise en place, la méthylation de l'ADN est propagée de façon stable au cours des divisions cellulaires. Néanmoins il existe plusieurs voies, dites passives ou actives, permettant de déméthyler l'ADN. Le mécanisme de déméthylation passif est causé par l'absence du maintien de la méthylation pendant la réplication de l'ADN. Les mécanismes de déméthylation actifs impliquent quant à eux des réactions chimiques conduisant à des mécanismes de réparation d'ADN, sans réplication, induites notamment par les enzymes TET (*Ten Eleven Translocation*). Chez les mammifères, trois membres de la famille TET sont connus : TET1, TET2 et TET3. Ces protéines catalysent la conversion de la 5mC en 5-hydroxyméthylcytosine (5hmC), une molécule intermédiaire de la déméthylation.

La méthylation de l'ADN joue des rôles importants chez les mammifères. Elle participe à l'empreinte génomique parentale, au contrôle des transposons et à la régulation de l'expression des gènes pendant le développement. Les profils de méthylation du génome sont fortement reprogrammés au cours de l'embryogenèse précoce et la différentiation cellulaire et, une fois mis en place, ils sont censés rester stables tout au long de la vie.

Les profils de méthylation de l'ADN du génome doivent être contrôlés avec précision pour maintenir le fonctionnement normal des cellules. Il a été montré que des dérégulations des enzymes DNMT favorisent l'apparition de tumeurs dans des modèles murins *in vivo*. Chez l'homme, DNMT3A et TET2 sont mutés à une étape précoce dans les leucémies. Les cellules cancéreuses se caractérisent parfois par une déméthylation globale du génome, ce qui pourrait favoriser l'instabilité génétique et des réarrangements chromosomiques. Cependant la caractéristique principale des cellules

cancéreuses est qu'elles présentent une hyperméthylation aberrante des îlots CpG dans les promoteurs des gènes suppresseurs de tumeurs. Cette hyperméthylation participe à la répression transcriptionnelle de gènes codant pour des protéines impliquées dans des voies de régulation cruciales telles que l'inhibition de l'invasion cellulaire, le déroulement du cycle cellulaire ou l'apoptose.

Première partie

Dans la première partie de ma thèse je me suis concentré sur la méthylation comme cible thérapeutique. De par sa nature réversible, la méthylation de l'ADN constitue une cible de choix pour des thérapies épigénétiques du cancer. L'intérêt clinique est d'inhiber la méthylation des îlots CpG dans les promoteurs de gènes suppresseurs de tumeurs afin de restaurer les mécanismes endogènes de surveillance et d'éliminer les cellules cancéreuses. A ce jour, plusieurs inhibiteurs de DNMT ont été développés et testés en phase clinique. C'est le cas de la molécule 5-azacytidine, un analogue de cytosine qui inhibe l'action des DNMTs en les séquestrant sur la chromatine, utilisée dans le traitement des syndromes myélodysplasiques. Cependant son utilisation pour traiter d'autres types de cancer est limitée car l'inconvénient majeur de cette molécule est qu'elle inhibe la méthylation de l'ADN de manière non-spécifique, ce qui induit une déméthylation globale du génome et a pour conséquence de nombreux effets indirects et une toxicité très élevée sur les cellules normales.

Dans ce contexte, l'objectif de ma thèse a été de développer des stratégies de déméthylation ciblée des promoteurs des gènes suppresseurs de tumeurs sans affecter le profil de méthylation global des cellules. Mon travail a voulu répondre aux questions suivantes :

(1) Peut-on induire une déméthylation spécifique d'un promoteur d'un gène suppresseur de tumeurs dans les cellules cancéreuses ?

(2) Cette reprogrammation épigénétique du gène est-elle stable ?

Pour cela nous avons utilisé comme modèle la lignée de cellules de cancer du sein MDA-MB-231. Pour induire la déméthylation ciblée de gènes suppresseurs de tumeurs, nous avons utilisé des protéines chimériques contenant un domaine de ciblage nucléotidique (TALE ou dCas9) fusionnée au domaine catalytique de la protéine humaine TET1 (hTET1) qui permet l'hydroxylation des cytosines méthylées. Nous avons choisi comme modèle le gène *SERPINB5* (aussi appelé *MASPIN*), un gène qui a des propriétés de suppresseur de tumeur dans les cellules épithéliales et dont le promoteur est hyperméthylé dans une grande proportion des tumeurs du sein et des lignées de cellules cancéreuses du sein.

Dans un premier temps, j'ai validé que le promoteur du gène *SERPINB5* est méthylé dans notre lignée de cellules MDA-MB-231 et que le gène est réactivé après traitement avec la 5-aza-2'-deoxycytidine, ce qui démontre que *SERPINB5* est directement réprimé par la méthylation de l'ADN dans cette lignée cellulaire. Nous avons ensuite conçu deux protéines TALE (TALE1 et TALE2) ciblant le promoteur de *SERPINB5* dans le but de les fusionner avec le domaine catalytique de hTET1 pour induire une déméthylation ciblée du promoteur de *SERPINB5*. Pour la suite du projet, j'ai sélectionné la protéine TALE2 qui présentait la meilleure affinité pour *SERPINB5*. J'ai également comparé l'efficacité du système TALE avec le système CRISPR-dCas9 en créant une protéine chimérique dCas9-hTET1 et différents ARNs guide (ARNg) ciblant le promoteur de *SERPINB5*. Après transfection des cellules, j'ai analysé l'état de méthylation du promoteur de *SERPINB5* par COBRA, BS-seq et amplicon sequencing, ainsi que l'état d'expression par RT-qPCR. Les résultats obtenus ont permis de valider la stratégie expérimentale et montrent que TALE2-hTET1 induit efficacement la déméthylation et la réactivation du gène *SERPINB5* dans les cellules du cancer du sein MDA-MB-231. La reprogrammation épigénétique induite par TALE2-hTET1 est plus efficace que celle obtenue avec la stratégie dCas9-hTET1, c'est pourquoi nous avons poursuivi le projet avec la protéine chimérique TALE2-hTET1.

Par la suite j'ai étudié la spécificité de la déméthylation de *SERPINB5* en réalisant des cartes du méthylome à l'échelle du génome entier dans les cellules transfectées avec TALE2-hTET1. Afin d'étudier la stabilité de la reprogrammation épigénétique, nous avons construit une lignée MDA-MB-231 stable exprimant TALE2-hTET1-IRES-eGFP sous le contrôle d'un promoteur inductible à la doxycycline. J'ai réalisé des expériences d'induction avec la doxycycline pendant 2 semaines et étudié la cinétique de déméthylation et de réactivation de l'expression de *SERPINB5*, aussi bien au niveau des ARNm que de la protéine. Dans le but d'étudier la persistance de cette reprogrammation, j'ai étudié si les effets se maintiennent à long terme. Nous observons que les cellules continuent d'exprimer la protéine SERPINB5 jusqu'à au moins 30 jours après avoir enlevé la doxycycline. En résumé j'ai démontré que la stratégie expérimentale développée permet d'induire une reprogrammation épigénétique spécifique et durable du gène *SERPINB5* dans les cellules de cancer du sein, et ces résultats peuvent ouvrir la porte à des recherches vers une application thérapeutique.

Deuxième partie

La seconde partie de ma thèse a consisté à développer et valider des stratégies d'édition de l'épigénome comme outil en recherche fondamentale pour valider des hypothèses sur le rôle de la méthylation de l'ADN dans le génome.

Des données du laboratoire ont identifié des gènes qui sont fortement ré-exprimés dans des embryons de souris (E8.5) KO pour DNMT1 ou KO pour DNMT3A/3B qui présente une perte globale de méthylation de l'ADN. C'est le cas notamment de gènes gamétiques tels que *Dazl* ou *Asz1* qui sont méthylés au niveau de leur promoteur. Des outils de déméthylation ciblée du génome dans les cellules de souris permettraient de tester et valider les modèles de régulation de ces gènes par la méthylation de l'ADN.

J'ai construit des plasmides codant pour dCas9-hTET1 avec des ARN guide (ARNg) ciblant le promoteur de *Dazl* et je les ai transfectés dans des fibroblastes de souris (MEF). J'ai étudié l'effet sur la méthylation (par BS-Seq) et l'expression de *Dazl* (par RT-qPCR) en comparant l'efficacité de plusieurs ARNg. Mes résultats montrent que cette stratégie permet une réactivation de l'expression de *Dazl*, cependant l'efficacité est très variable selon les ARNg et l'amplitude de la démétylation est très faible. Par ailleurs, j'ai comparé l'efficacité de constructions dCas9-hTET1 et dCas9-mTET1 (domaine catalytique de la protéine TET1 d'origine souris) pour réactiver *Dazl* et, de manière surprenante, mes résultats montrent que mTET1 n'est pas efficace par rapport à hTET1. En conclusion, mes travaux ont validé l'utilisation de construction dCas9-hTET1 pour induire une déméthylation ciblée dans les cellules de souris, cependant des études sont encore nécessaires pour tester la spécificité et augmenter l'efficacité de la déméthylation.

La prochaine étape du projet est de tester la méthode basée sur le système dCas9 Suntag-hTET1, permettant le ciblage de plusieurs effecteurs à la même région. Il a été montré que ce système est plus efficate et spécifique. Des résultats preliminaires dans le laboratoire montrent une efficacité très elevée du système dans la déméthylation et reexpression du gène *Dazl* dans les celules MEF. Des expériences futures nous permettront de conclure si ce système est adaptable à d'autres contextes géniques.

Dans sa globalité, ce travail de thèse a permis de développer et valider des outils innovants d'ingénierie épigénétique permettant de moduler de manière spécifique l'état de méthylation des gènes. Ces résultats pourront avoir des impacts en recherche fondamentale ainsi que dans la recherche biomédicale.

Table of contents

Figures

INTRODUCTION

METHODS

RESULTS

1. Results (1) : Manuscript

2. Results (2)

DISCUSSION

Tables

INTRODUCTION

METHODS

RESULTS (1)

RESULTS (2)

DISCUSSION

INTRODUCTION

INTRODUCTION

DNA sequences encode all the information governing life, from the colour of our hair to the predisposition to develop a disease. These instructions are organised in genes. But how do the cells know when and where to read the information encoded in genes?

The differentiated state of a particular cell type is related to the activated or inactivated status of specific sets of genes. Gene expression is regulated in several ways. First, the DNA sequence itself harbours specific sequence motifs that dictate the binding of transcription factors in gene promoters or enhancers regions. This binding triggers the activation or repression of the genes. In addition to the information contained in DNA sequence, there are additional mechanisms involved in the modulation of gene expression beyond the DNA sequence, which are called 'epigenetics'.

Epigenetics consists in dynamic and heritable changes in DNA and chromatin modifications that modulate gene expression without altering the underlying DNA sequence. The most studied one and the focus of my thesis is cytosine methylation in DNA.

Given the fact that cytosine methylation is a reversible mark and taking advantage of the recent progress in genome and epigenome editing tools, we have developed strategies to regulate gene expression by specifically modulating gene promoter methylation in mammalian cells. In the first part of my thesis I have applied these technologies to reactivate a tumor suppressor gene in cancer cells and studied the stability of this methylation reprogramming over time. In the second part I have further optimised these epigenetic engineering tools to demethylate specific sequences and test the role of DNA methylation in the mouse genome.

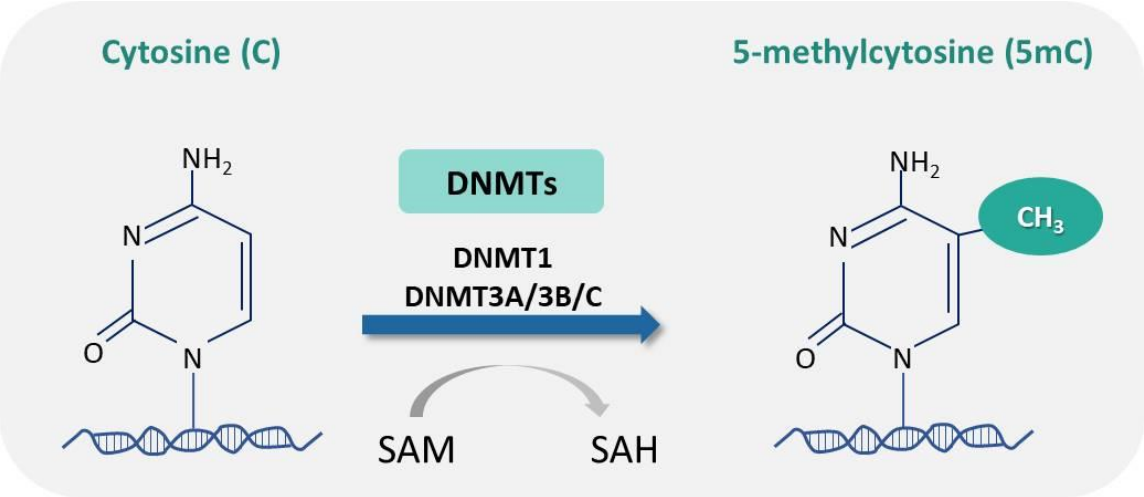
In this introduction I will first present DNA methylation, its distribution, functions and principal actors in mammalian cells. Then, I will discuss the implication of perturbed DNA methylation in cancer development and its possible role as a biomarker and as a therapeutic target. Finally, I will introduce existing epigenome editing tools, the latest advances in the field and their limitations.

1 DNA METHYLATION

1.1 Historical perspective

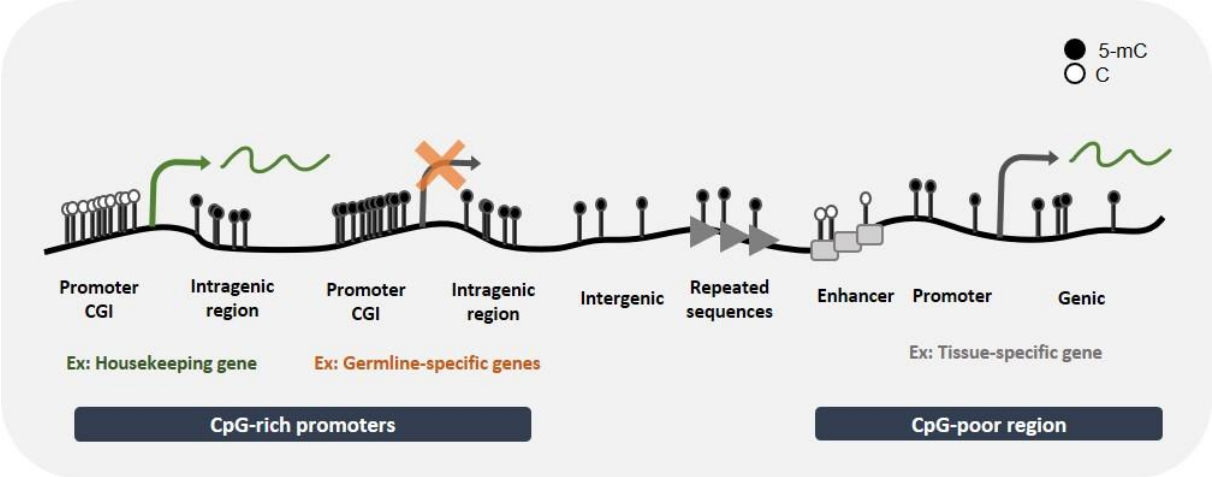
The first evidence of the existence of 5-methylcytosine (5mC) dates from 1925 in the tuberculinic acid, an unusual nucleic acid found in *Tubercle bacillus*, the bacterium responsible for tuberculosis in humans (Johnson and Coghill, 1925). This study was criticised by the scientific community because they based their observations on the optical properties of nucleic acids. It was not until 1948 that its existence was fully confirmed and its presence described in eukaryotes: it was found as a nucleotide that migrated at a higher rate than cytosine in a paper chromatography of calf thymus (Hotchkiss, 1948).

It was only years later that researchers made hypotheses on the possible biological functions of DNA methylation in eukaryotes. At this time, the implication of DNA methylation in biological processes was suggested based on observations made in bacteria because, in fact, DNA methylases had been founded by this time in bacteria but not in eukaryotes. Methylation was proposed to explain the initiation and maintenance of chromosome X inactivation (Riggs, 1975). Holliday and Pugh, 1975 proposed DNA methylation as a mechanism that controls gene expression. In the late 70's there were several papers discussing this possible role of DNA methylation. For example, McGhee and Ginder, 1979 compared the methylation status of the chicken beta-globin locus in cells where the gene was expressed and cells where the gene was not expressed with restriction enzymes able to discriminate between methylated and unmethylated cytosines. They concluded that the gene was unmethylated when expressed and vice versa. Subsequently, other studies supported the hypothesis that DNA methylation correlates with gene repression and might regulate gene activity, for example in the metallothionein-I (MT-I) gene in mouse cell lines (Compere and Palmiter, 1981). In *vitro* studies showed that methylated genes are transcriptionally inactive when transferred into *Xenopus* oocytes or in mammalian cells (Stein et al., 1982; Vardimon et al., 1982). Ultimately, the discovery and cloning of the first DNA methyltransferase (now known as DNMT1) in mouse cells (Bestor et al., 1988) stimulated many research on DNA methylation and DNA methyltransferases in mammalian cells, which led to the consensus that DNA methylation is a major epigenetic factor influencing gene activity.



Introductory Figure 1. Cytosine methylation in mammals.

DNMTs catalyse the transfer of a methyl group from the *S*-adenosyl-L-methionine to the carbon 5 of cytosine, leading to the formation of 5-methylcytosine (5mC).
SAM: *S*-adenosyl-L-methionine, SAH: *S*-adenosyl-L-homocysteine.



Introductory Figure 2. Distribution of CpG methylation in mammalian genomes.

CpGs are enriched in promoter regions (CpG islands, CGIs). CGIs in promoters are usually unmethylated in all genes, except for exceptional cases such as germline-specific genes, in which the methylation contributes to the maintenance of the gene repression. The genome is globally hypermethylated (intragenic and intergenic regions, repeated sequences). CpG-poor promoters are usually associated to tissue-specific genes and they experience a demethylation upon transcription activation. 5mCs are represented as black spheres and Cs as white spheres.

1.2 Methylation distribution in mammals

In contrast to bacteria that use DNA adenine methylation, DNA methylation in eukaryotes occurs mainly on cytosines. Cytosine methylation reaction is performed by the DNA-methyltrasferases (DNMT), which catalyse the transfer of a methyl group (-CH_3) from the cosubstrate *S*-adenosyl-L-methionine (SAM) to the carbon 5 of cytosine (**Introductory Figure 1**). *S*-adenosyl-L-homocysteine (SAH) is also formed as a product. 5mC represents around 1% of all nucleotides in mammalian cells, this percentage being relatively stable in most cell types.

In eukaryotes, DNA methylation does not occur on all cytosine bases in the genome but is targeted and restricted to some sites, obtaining different patterns of methylation. 5mC has a mosaic genomic distribution in plants, fungus and invertebrates: small regions with high methylation levels are interspersed among larger regions devoid of DNA methylation, and DNA methylation occurs mainly in transposable elements (TEs) and gene bodies but less in other genomic sequences (Bonasio et al., 2012; Feng et al., 2010; Veluchamy et al., 2013). Some invertebrates such as *D. melanogaster* and *C. elegans* do not contain detectable levels of 5mC. Nonetheless, recent studies have found methylation in these organisms at the nitrogen 6 of adenosines (N6-methyladenine, 6mA) (Greer et al., 2015; Zhang et al., 2015a) that could compensate for the absence of 5mC in processes such as silencing of repeated sequences.

In mammals (and in vertebrates in general), cytosine methylation is distributed globally throughout the genome. Cytosine methylation was found to take place preferentially in the CpG dinucleotide context in the early 60's (Doskocil and Sorm, 1962), although mammalian genomes are globally depleted in CpGs. The reason for this is evolutive. Non-methylated cytosines are spontaneously deaminated into uracil, which is recognised by the uracil-DNA-glycosylase (UDG) from the base excision repair system (BER) and corrected. However, when this spontaneous deamination takes place in methylated cytosines, the resultant product is a thymine, which is a normal DNA nucleotide and sometimes escapes the DNA repair machinery by thymine-DNA-glycosylase (TDG). This C to T transition is the most common mutation in diseases and through evolution, which leads to a global decrease in CpG content of vertebrate genomes. CpGs are mutated at a lower frequency in CpG-rich promoters, meaning that these promoters might be hypomethylated in the germ line and so CpG islands are conserved (Saxonov et al., 2006).

There are 28 million CpGs in the human genome and 22 million CpGs in the mouse genome, from which 60 to 80% are methylated in most cell types. Cytosine methylation in most cellular populations presents a classical bimodal distribution with most CpGs being either hypomethylated or hypermethylated, but few CpGs having intermediate methylation levels (**Introductory Figure 2**). DNA

methylation in non-CpG context (CHG or CHH, where H = A, C or T) is very rare in mammalian cells and is not heritable through cell division, therefore it is detected either in cells with very high *de novo* methylation activity (Embryonic Stem cells, ES cells) or in non-dividing cells like neurons or oocytes (Feng et al., 2010; Lister et al., 2009, 2013; Tomizawa et al., 2011; Varley et al., 2013; Xie et al., 2012).

1.2.1 CpG-rich regions

In vertebrates, DNA methylation levels are inversely correlated with CpG density. CpG-rich regions are associated with gene promoters and surprisingly they are mostly unmethylated independently of the gene activity (**Introductory Figure 2**). CpG dense regions are called CpG islands (CGI) and they represent 1% of the genome. 60 to 70% of gene promoters contain a CpG island around the Transcription Start Site (TSS) in the mouse or human genomes. There are several parameters defining a CGI and not a universally accepted definition, but essentially, they are regions of 0,5-2 kilobases (kb) in which the observed/expected ratio of CpGs is higher than 0.6. This ratio indicates that CGIs have not gained CpGs under positive selection during evolution, but they are sequences that have not undergone CpG loss during evolution.

Genes with promoters containing CpG islands are frequently associated with ubiquitous housekeeping genes and genes expressed in early embryos (Ponger et al., 2001; Saxonov et al., 2006). Housekeeping genes are constitutively expressed across many developmental stages and adult tissues despite the differences in the transcriptional machinery among cell types.

1.2.2 CpG-poor regions

CpG-poor regions such as exons, introns, intergenic regions and repetitive elements show a high level of methylation (Li et al., 2010; Meissner et al., 2008; Weber et al., 2007) (**Introductory Figure 2**).

Although most of the annotated gene promoters are associated with a CpG island (CGI, *see 1.2.1),* there are many genes associated with CpG-poor promoters in the genome. Those are more likely to contain a TATA box and are frequently linked to tissue-specific genes. These tissue-specific promoters often show a high level of methylation in the tissues where the genes are silent and become hypomethylated in the tissues where the genes are activated (Nagae et al., 2011; Saxonov et al., 2006). Consequently, tissue-specific differentially methylated regions (tDMRs) occur much more frequently in CpG-poor promoters compared to CpG-rich promoters (Ziller et al., 2013). An interesting feature is that a high proportion of the genes having specific hypomethylation in differentiated cells are densely methylated in both ES cells and induced Pluripotent Stem Cells (iPSCs), suggesting the possibility that the default stage of CpG-poor promoters in the embryonic stage is full methylation and that the methylation loss occurs during terminal differentiation (Nagae

et al., 2011). This suggests that tissue-specific promoter hypomethylation and selective binding of transcription factors (TFs) are coordinated to induce spatio-temporal promoter expression during cell differentiation.

CpG-poor distal regulatory regions such as enhancers also show a frequent correlation between activity and DNA hypomethylation. In fact the most frequent tDMRs between tissues and cell types occur in distal CpG poor enhancers (Hon et al., 2013; Ziller et al., 2013). Further studies showed that the loss of methylation in active enhancers occurs downstream of TF binding. For example, the binding of TFs like CTCF or REST to DNA is sufficient and necessary to create a low methylated region (LMR) (Feldmann et al., 2013; Stadler et al., 2011). Therefore, methylation of CpG-poor regulatory regions is not incompatible with gene activation, and it is suggested that demethylation is not instructive to gene expression but is a consequence of the TF binding.

The role of DNA methylation in gene bodies remains unclear. Gene body methylation has been observed in *Arabidopsis thaliana* and associated with active genes (Zhang et al., 2006), but it has also been suggested that this gene body methylation limits transcription elongation (Zilberman et al., 2007). Some studies suggest a similar repressive role of intragenic methylation in mammalian cells arguing that it triggers the formation of a chromatin structure impairing RNA polymerase II (RNApolII) elongation (Lorincz et al., 2004). However, gene body methylation is more frequent in the active human X chromosome than in the inactive one (Hellman and Chess, 2007) and is associated with higher transcription levels of genes in mammalian cells (Rauch et al., 2009), arguing against a negative role of gene body methylation on transcription elongation. What can be the biological reasons to target repressive marks to actively transcribed regions? One proposed model is that intragenic DNA methylation limits the activation of intragenic cryptic promoters within active transcribed regions (Baubec et al., 2015; Neri et al., 2017), or to facilitate a quick restoration of a repressed state of the genes if they are not required anymore.

1.2.3 CGI methylation protection

What protects CGI sequences from becoming methylated? The exact causes are still unknown, but we can think about different models that exclude the DNMTs from the promoters or promote an active demethylation of CGIs.

The underlying DNA sequence is important to determine the hypomethylated states of CGIs. Promoter fragments of 1-kb long in average recapitulate the endogenous hypomethylation pattern of CpG rich sequences when inserted away in the mouse ES (mES) cells genome (Krebs et al., 2014; Lienert et al., 2011). In addition, human DNA inserted in the mouse genome recapitulates the protection of CGI promoters against DNA methylation, further supporting that the DNA sequence

itself plays a central role in specifying the hypomethylated state (Long et al., 2016b). The presence of transcription factor binding sites in CpG rich sequences seems to have an important role. In support of this model, the binding of TFs such as SP1 has been shown to have a protective effect against DNA methylation in transfection experiments (Brandeis et al., 1994; Macleod et al., 1994). The SP1 binding site (G/T)GGGCGG(G/A)(G/A)(C/T) is one of the most commonly found in CGIs. However how the transcription factors binding could exclude DNMTs remains unknown.

Proteins containing a CXXC (cys – X – X – cys) domain could also play an important role in maintaining the hypomethylated state of CGIs. The CXXC domain mediates binding to DNA containing clustered unmethylated CpGs. This domain is usually found in epigenetic regulators. For example, the protein CXXC finger protein 1 (CFP1) binds to unmethylated CpGs and recruits the complex SETD1A that methylates the lysine 4 of histone 3 (H3K4me3). H3K4me3 is a histone mark present on CGIs independently of their transcriptional status (Guenther et al., 2007). H3K4me3 is a DNA methylation antagonist, however *Cfp1* knockout (KO) ES cells do not experience changes in DNA methylation levels nor gene expression, suggesting that other factors must be implicated in the mechanism of protection (Clouaire et al., 2012). One such factor is another CXXC domain containing protein, FBXL10 (also known as KDM2B). FBXL10 occupies all CGI sequences in mESC, and inactivation of FBXL10 leads to robust *de novo* methylation of a large fraction of CpG islands in *Fblx10-/-* ES cells (Boulard et al., 2015).

TET proteins can also prevent CGIs methylation (*See 1.4.2 TET enzymes and 5-hydroxymethylcytosine*).

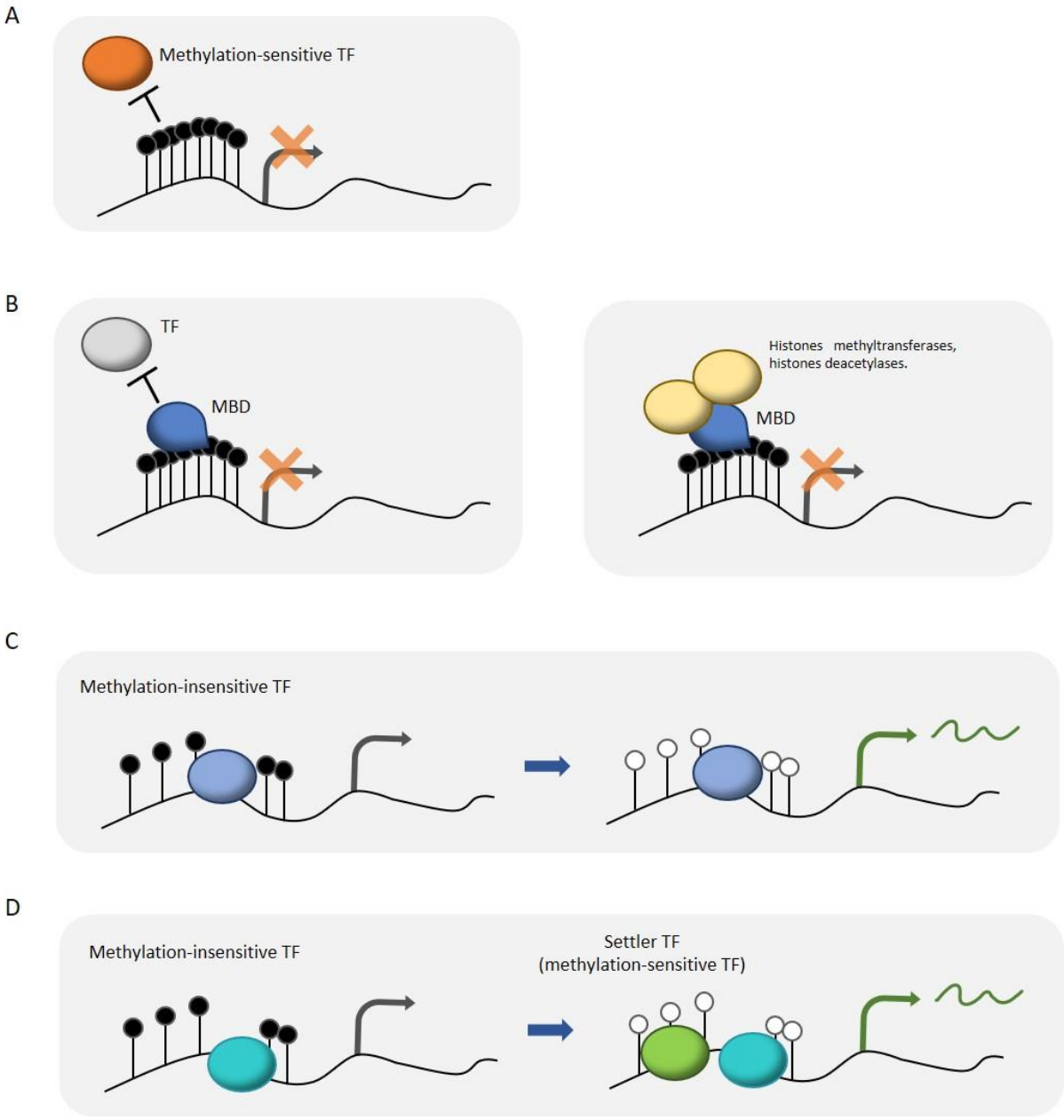
There are exceptions to this protection of CGIs against DNA methylation. First, CGIs in imprinted gene promoters (*See 1.5.3 Parental imprinting*). Second, germline-specific genes are frequently methylated during development, and this CGI methylation is directly required to maintain these genes in a repressed state in somatic cells (Auclair et al., 2014; Weber et al., 2007). Third, CGIs in bodies of genes are much more frequently methylated than promoter CGIs. The methylation of intragenic CGIs is dependent on the transcriptional status of the gene and is stimulated by transcription running across the CGI (Jeziorska et al., 2017).

Besides CGIs in annotated promoters, another class of CGIs, called orphan CGIs, has been described (Illingworth et al., 2010). These CGIs are not associated with annotated gene promoters and represent approximately half of all CGIs. They present a dynamic methylation status, and unlike CGIs at known promoters, they often become methylated during development. Orphan CGIs present promoter-like characteristics: around 40% correlate with the presence of H3K4me3 in mouse and human, and 21% of the human ones are associated to RNApolII. They might code potentially for

novel genes, non-coding RNAs or alternative promoters. However, in colorectal cancer (CRC) cells *de novo* methylation affects both known CGIs and orphan CGIs, meaning that the developmental programming is not recapitulated in cancer development (Illingworth et al., 2010).

1.2.4 DNA methylation and transcription factor binding

Cytosine methylation can impact the binding of TFs and gene expression in several ways (**Introductory Figure 3**). In general, TF binding is thought to be blocked by DNA methylation because DNA methylation modifies the affinity of TFs towards their binding sites (**Introductory Figure 3A**). Alternatively, DNA methylation can recruit proteins such as Methyl-CpG binding domain (MBD) proteins (Klose and Bird, 2006) or the zinc finger-domain containing protein Kaiso (Prokhortchouk, 2001) (**Introductory Figure 3B**). Other TFs, like CTCF and REST, are insensitive to DNA methylation and in turn, their binding to regulatory regions is sufficient to trigger active DNA demethylation (Feldmann et al., 2013; Stadler et al., 2011) (**Introductory Figure 3C**). Within the last years, several studies have tried to identify TFs whose binding to DNA is dependent on the DNA methylation status. To this aim, combination of molecular techniques such as DNAse I hypersensitive sites (DHS) mapping, Chromatin Immunoprecipitation sequencing (ChIP-seq), methylation-sensitive SELEX (Systematic evolution of ligands by exponential enrichment) or Mass-spectrometry in methylated and unmethylated contexts have been used. One study tried to identify methylation-sensitive factors by mapping new DNAse I hypersensitive sites that appear in mES cells devoid of DNA methylation (triple knockout for *Dnmt1*, *Dnmt3a* and *Dnmt3b*) (Domcke et al., 2015). This identified NRF1 as a major methylation sensisitive TF. NRF1 can bind to its binding motif when unmethylated, but, when it is methylated, it needs the assistance of an upstream TF to trigger hypomethylation at its binding site (Domcke et al., 2015) (**Introductory Figure 3D**). Interestingly, a recent report shows that some factors have a preference for methylated motifs. This is the case for several factors with roles in embryonic and organism development such as homeodomains. Crystal structure experiments showed that interactions between hydrophobic residues and the methyl group of 5mC directed this preference (Yin et al., 2017). Deciphering the proteins binding to 5mC oxidative derivatives is important to understand their possible role in biological functions, but its study is complicated because these marks are present in low frequency. However, it has been shown that these derivatives recruit transcription regulators and DNA repair proteins, supporting their role in active demethylation (Spruijt et al., 2013). UHRF2, whose possible role in DNA methylation remains unclear, also appeared as a specific reader for 5-hydroxymethylcytosine (5hmC) (Spruijt et al., 2013).



Introductory Figure 3. Cytosine methylation and TFs binding in regulatory regions.

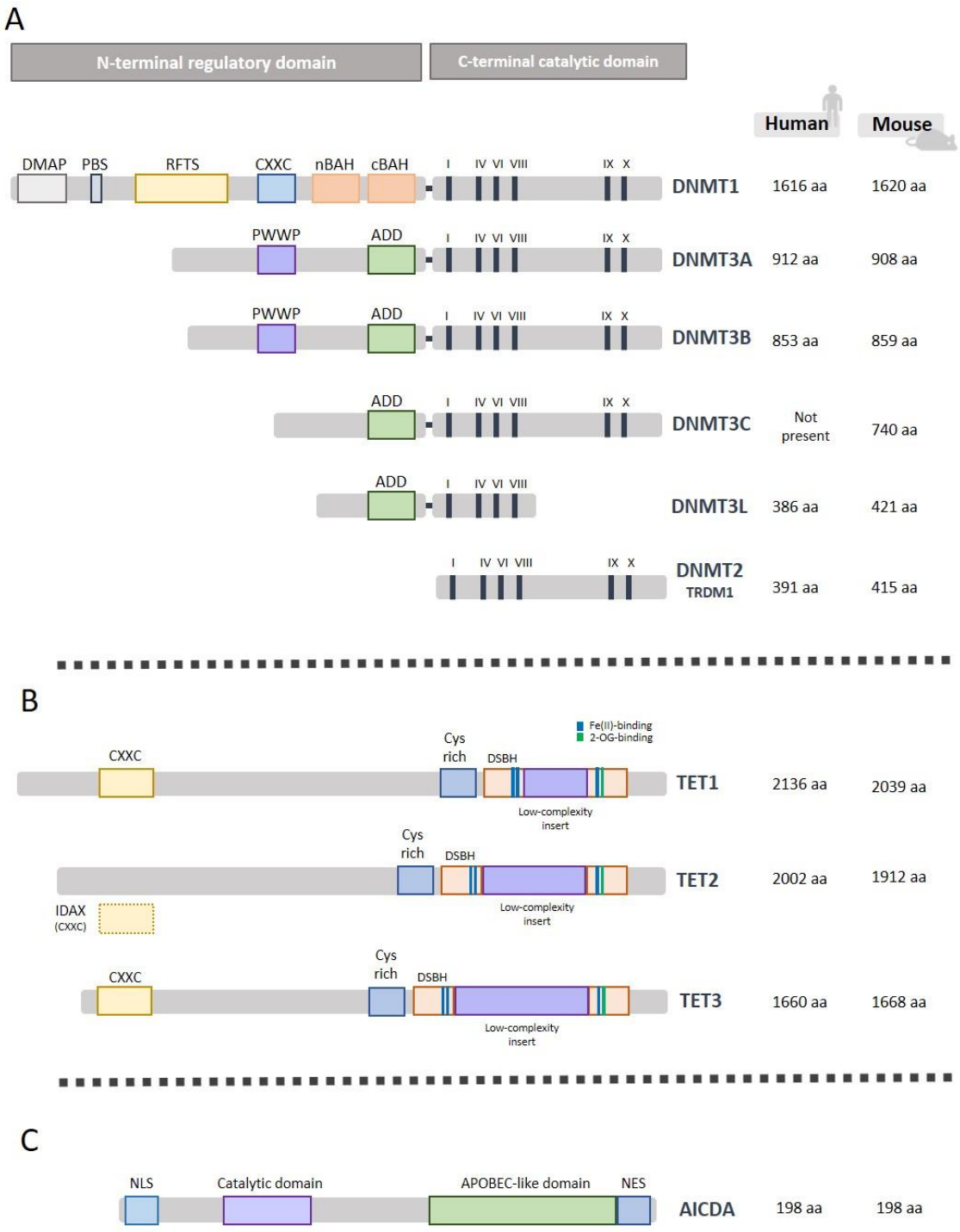
A. Methylation-sensitive TFs are blocked by DNA methylation.
B. Methyl-CpG binding domain proteins (MBDs) recognise methylated promoters. Its binding represses expression by preventing TFs binding or recruiting repressive factors
C. Methylation-insensitive TFs can bind to methylated promoters or enhancers, and their binding leads to reduced methylation.
D. A methylation-insensitive TF leads to demethylation, allowing the binding of another settler TF that triggers gene expression.

1.3 Mammalian DNA methyltransferases

Methylation is catalysed by DNA methyltransferases (DNMTs). DNMTs activities were first purified in *E. coli* (Gold and Hurwitz, 1964). There are 6 members of the DNMT family in mammals: DNMT1, DNMT3A and 3B, DNMT2, DNMT3L and the most recently discovered DNMT3C in the mouse. These enzymes allow the establishment of DNA methylation pattern *de novo* as well as its maintenance through cell division. All DNMTs with the exception of DNMT3L have a conserved catalytic domain. In the next sections I will discuss mammalian DNMTs.

1.3.1 DNMT1

DNMT1 was the first DNMT identified and purified in eukaryotic organisms (Bestor and Ingram, 1983; Bestor et al., 1988). This enzyme is involved in the maintenance of the methylation through cell division. As CpG is a palindromic site, DNA methylation can be copied from the parent strand to the newly synthesised DNA strand and inherited through generations. DNMT1 recognises preferentially hemimethylated CpGs over unmethylated DNA (Vilkaitis et al., 2005). This enzyme has been shown to be constitutively expressed in cells undergoing division, and it is especially abundant during the S-phase (Kishikawa et al., 2003). DNMT1 has two functional parts: the N-terminal domain containing regulatory functions and the C-terminal domain harbouring the catalytic activity (**Introductory Figure 4A**). The catalytic domain of DNMT1 is conserved within the other mammalian DNMTs and in plants, and interestingly, eight of the ten motifs necessary for the catalytic activity are conserved in prokaryotic methyltransferases (Finnegan and Dennis, 1993; Lauster et al., 1989). The N-terminal domain contains an interaction domain with DMAP1 (DNA methyltransferase associated protein 1) (Rountree et al., 2000), a Proliferating Cell Nuclear Antigen (PCNA) binding domain site (PBS) and the replication foci targeting sequence (RFTS) that targets the protein to the replication fork during the S-phase of the cell cycle (Leonhardt et al., 1992). There are also a CXXC DNA binding domain that contributes to regulating the catalytic activity by interacting with unmethylated CpGs in DNA (Pradhan et al., 2008; Song et al., 2011), and two bromo-adjacent homology domains (BAH). The binding of the DNMT1 CXXC domain to unmethylated CpGs creates a loop within the BAH domain that prevents the catalytic domain to access DNA and perform *de novo* methylation (Song et al., 2011). However, several groups have highlighted the possible implication of DNMT in *de novo* methylation. Arand et al., 2012 showed that DNMT1 contributes to *de novo* methylation of repeated sequences in mES cells, while others have suggested a cooperation between DNMT1 and *de novo* DNMTs (DNMT3A and 3B) (Fatemi et al., 2002; Liang et al., 2002). Moreover, depletion of DNMT1 in oocytes resulted in a slight decrease of methylation. As these are non-dividing cells, methylation performed by DNMT1 should be considered as *de novo* (Shirane et al., 2013).



Introductory Figure 4. Schematic representation of DNMT and TET proteins.

A. DNMTs have a N-terminal regulatory domain and a C-terminal catalytic domain. DNMT3L has a truncated catalytic domain and consequently does not have catalytic activity.
The DMAP, PBS and RFTS domains in DNMT1 allow binding to cofactors and the recruitment of the enzyme to the replication fork. The CXXC domain targets unmethylated sequences.
DNMT3A, 3B, 3C and 3L have a PWWP domain and a ADD domain that regulate the interaction with chromatin.
B. TET proteins share a conserved C-terminal part containing a double sheet beta helix (DSBH) domain. This part is the catalytic domain.
C. AID has a catalytic domain conserved among members of the APOBEC family. The N-terminal domain contains the Nuclear Localisation Signal (NLS) and the C-terminal the Nuclear Exportation Signal (NES).

How is DNMT1 recruited to the replication fork? DNMT1 is part of a complex. It interacts with the PCNA; however, this interaction is not the only mechanism required for faithful maintenance methylation by DNMT1 during replication (Spada et al., 2007). Another important cofactor is Ubiquitin-like containing PHD and RING finger domains 1 (UHFR1), which plays multiple roles in regulating DNMT1 activity. UHRF1 specifically binds to hemimethylated CpG sites thanks to its SRA (SET and RING associated) domain, which helps recruiting DNMT1 to hemimethylated DNA at the replication fork (Bostick et al., 2007; Sharif et al., 2007). The UHRF1 C-terminal RING domain also functions as an E3 ubiquitin ligase that promotes histone H3 ubiquitylation, which is then subsequently recognised by DNMT1 through its RTFS domain (Nishiyama et al., 2013). Finally, numerous studies show that UHRF1 also directly interacts with DNMT1 and stimulates its catalytic activity in multiple ways (Bashtrykov et al., 2014; Li et al., 2018).

While mouse ES are surprisingly viable despite DNMT1 depletion, this enzyme plays an essential role in development as shown by the embryonic lethality of *Dnmt1* KO mice at around embryonic day E9.5. These embryos present a hypomethylated genome compared to wild-type embryos (Lei et al., 1996; Li et al., 1992).

1.3.2 DNMT3A and DNMT3B

De novo methylation of the genome takes place after the waves of demethylation in the pre-implantation embryo and the germ cells during embryonic development (Kafri et al., 1992; MONK et al., 1987). In addition, *Dnmt1* KO mES cells still retained remarkable methylation levels and had the ability to methylate foreign sequences (Lei et al., 1996). These facts suggested that other enzymes involved in the *de novo* methylation of DNA remained to be discovered.

DNMT3A and 3B were identified by homology search compared to the sequence of a bacterial type II cytosine-5 methyltransferase (Okano et al., 1998a). Recombinant DNMT3A and 3B showed no preference for hemimethylated DNA, suggesting that they could be the long-sought mammalian *de novo* methyltransferases. This idea was subsequently demonstrated by knockout studies in ES cells and embryos (Okano et al., 1999). Both proteins contain a PWWP (P = proline, W = tryptophan) domain and an ATRX-DNMT3-DNMT3L (ADD) PHD-related domain in the N-terminal region that regulate interaction with chromatin (**Introductory Figure 4A**). The PWWP domain has been proposed to target DNMT3A or 3B to transcribed gene bodies previously marked with H3K36me3 (Baubec et al., 2015; Dhayalan et al., 2010). The ADD domain mediates the binding of DNMT3A and 3B to histone H3 unmethylated on the lysine 4 (Ooi et al., 2007; Otani et al., 2009), and this interaction stimulates the catalytic activity of the enzymes (Guo et al., 2015; Li et al., 2011a).

DNMT3B shows a preference for methylating non-nucleosome DNA *in vitro* (Morselli et al., 2015). Human DNMT3A and 3B have been shown to have intrinsic sequence preference for the nucleotides flanking the CG: DNMT3A methylation hotspots present frequently a T at position -2 and a C at position +2, whereas for DNMT3B the nucleotides are a T at position -1 and a G at position +1 (Wienholz et al., 2010). These results are in concordance with previous reports indicating a preference of DNMT3A for CpGs flanked by pyrimidines *in vitro* (Lin et al., 2002).

DNMT3A and DNMT3B have both redundant and specific functions in development. DNMT3B is strongly expressed during early stages of embryogenesis, being the principal enzyme implicated in *de novo* methylation during embryonic implantation (Auclair et al., 2014; Borgel et al., 2010). However DNMT3A and 3B act synergistically during development to methylate the bulk genome because DNA methylation levels are only partially reduced in single *Dnmt3a* or *Dnmt3b* KO embryos (Auclair et al., 2014; Okano et al., 1999). On the other hand, DNMT3B also has specific functions in development: it is the only enzyme required for the methylation of a subset of CGIs on autosomes and all CGIs on the inactive X chromosome (Auclair et al., 2014; Gendrel et al., 2012). In contrast to DNMT3B, DNMT3A expression is not restricted to early embryonic stages but DNMT3A is also expressed at late embryonic stages and in differentiated cells. For example, DNMT3A is the main enzyme required to establish DNA methylation in differentiating gametes. Conditional inactivation of *Dnmt3a* in gametes, but not of *Dnmt3b*, abolishes the acquisition of methylation at imprinting control regions and genome-wide (Kaneda et al., 2004; Shirane et al., 2013; Smallwood et al., 2011). Consequently, female mice depleted for DNMT3A in germ cells are sterile as their offspring die at mid gestation due to improper setting of DNA methylation at maternal imprinting genes (Kaneda et al., 2004).

Both enzymes are essential in development. *Dnmt3a* KO mice follow a normal embryonic development until birth but they become runted and die around 4 weeks after birth (Okano et al., 1999). In contrast, *Dnmt3b* KO embryos develop normally until embryonic day E9.5 but then show multiple developmental defects and die before embryonic day E15.5. Embryos double KO for *Dnmt3a* and *Dnmt3b* show an even stronger phenotype: they fail to complete gastrulation and die around embryonic day E11.5 (Okano et al., 1999). These phenotypes are in agreement with the expression patterns of the enzymes and their implication in DNA methylation at different developmental stages.

The possible role of DNMT3A and 3B in the maintenance of methylation has been widely discussed, showing evidence that DNMT1, 3A and 3B cooperate to maintain the methylation at repetitive elements, single genes or imprinted regions (Arand et al., 2012; Chen et al., 2003; Liang et al., 2002).

1.3.3 DNMT3L

DNMT3-like protein (DNMT3L) shows an homologous sequence with DNMT3A and DNMT3B, but it is the only DNMT that does not contain a functional catalytic domain (it is truncated) (Aapola et al., 2000) (**Introductory Figure 4A**). It has been shown that DNMT3L interacts with the other DNMT3 (3A and 3B) and acts as a cofactor stimulating their activities. It interacts with unmethylated H3K4 and with the catalytic domain of DNMT3A (Jia et al., 2007; Ooi et al., 2007). This is consistent with data showing that methylation of H3K4 is not compatible with DNA methylation.

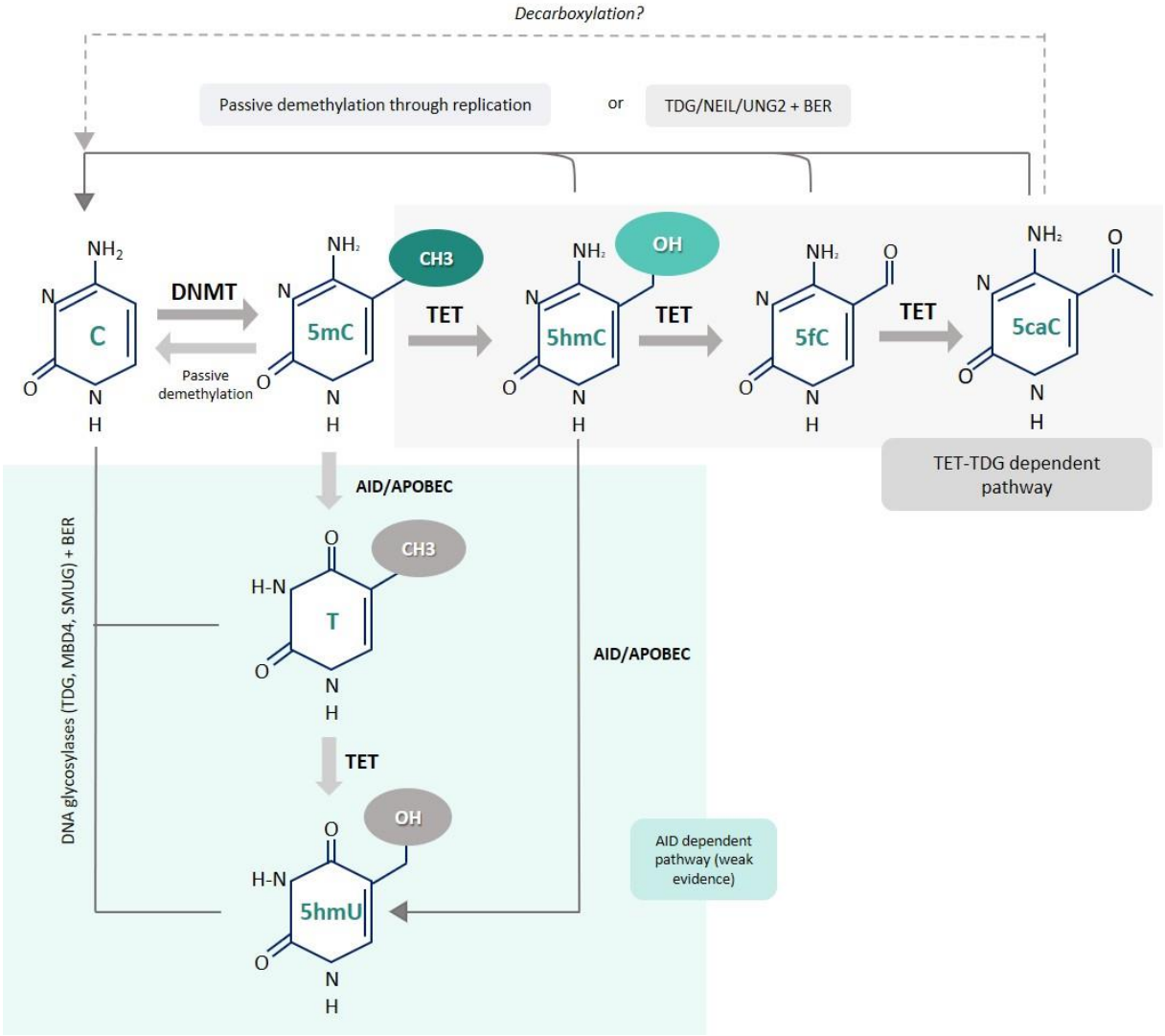
DNMT3L expression is mostly restricted to germ cells, where it is required to establish correct methylation at imprinted genes, TEs and genome-wide (Bourc'his, 2001; Hata et al., 2002; Shirane et al., 2013; Smallwood et al., 2011; Zamudio et al., 2015). Conditional inactivation of *Dnmt3L* in germ cells showed that mice are viable but not capable of establishing the *de novo* methylation in the germline of both sexes, leading to male infertility and death of the offspring from KO females (Bourc'his, 2001; Bourc'his and Bestor, 2004).

1.3.4 DNMT3C

DNMT3C is the most recent discovered DNMT and is found only in mice but not in humans. It evolved through duplication of the *Dnmt3b* gene around 50 million years ago and had been considered a pseudogene until its discovery (Barau et al., 2016) (**Introductory Figure 4A**). This enzyme is expressed in male germ cells, where it methylates and silences young transposons during spermatogenesis (transposons that have appeared relatively recently in evolution) (Barau et al., 2016; Jain et al., 2017).

1.3.5 TRDMT1 (DNMT2)

TRDMT1/DNMT2 (tRNA (cytosine-5-)-methyltransferase) lacks the regulatory N-terminal domain present in the other members of the family (**Introductory Figure 4A**). It is highly conserved among species, even in organisms lacking detectable levels of DNA methylation such as *D. melanogaster* or *S. pombe*, which suggests that it might play other functions than methylating DNA. Even if DNMT2 forms a complex with DNA, no methyltransferase activity is observed on DNA (Dong, 2001; Okano et al., 1998b). DNMT2-deficient mES cells are able to perform *de novo* and maintenance methylation of endogenous or newly integrated DNA (Okano et al., 1998b), indicating that DNMT2 is not involved in DNA methylation. Instead, Goll, 2006 showed that DNMT2 methylates transfer RNAs (tRNA), in particular the cytosine 38 that lies in the anticodon loop of aspartic acid tRNA. More recently, it has been shown that DNMT2 methylates additional tRNA residues in drosophila and mammals, promoting tRNA stability (Schaefer et al., 2010; Tuorto et al., 2012).



Introductory Figure 5. Pathways for cytosine demethylation.

5mC is established and maintained by the DNMTs. 5mC can be erased in a replication-dependent process by lack of methylation maintenance (passive demethylation).
Active demethylation is achieved through the TET-TDG dependent pathway. TET enzymes perform iterative oxidations leading to the formation of 5hmC, 5fC and 5caC. Those intermediates can be lost in a replication-dependent manner or can be actively excised via DNA glycosylases (TDG, NEIL) followed by BER. 5caC could also be decarboxylated to return to the unmethylated C, but a responsible enzyme has not yet been discovered.
5mC and 5hmC could be deaminated by cytosine deaminases like those from the AID/APOBEC system, and the products (T and 5hmU) excised by DNA glycosylases (TDG, MBD4, SMUG1), followed by BER. However, there is little experimental evidence for this AID dependent pathway.

1.4 DNA demethylation pathways

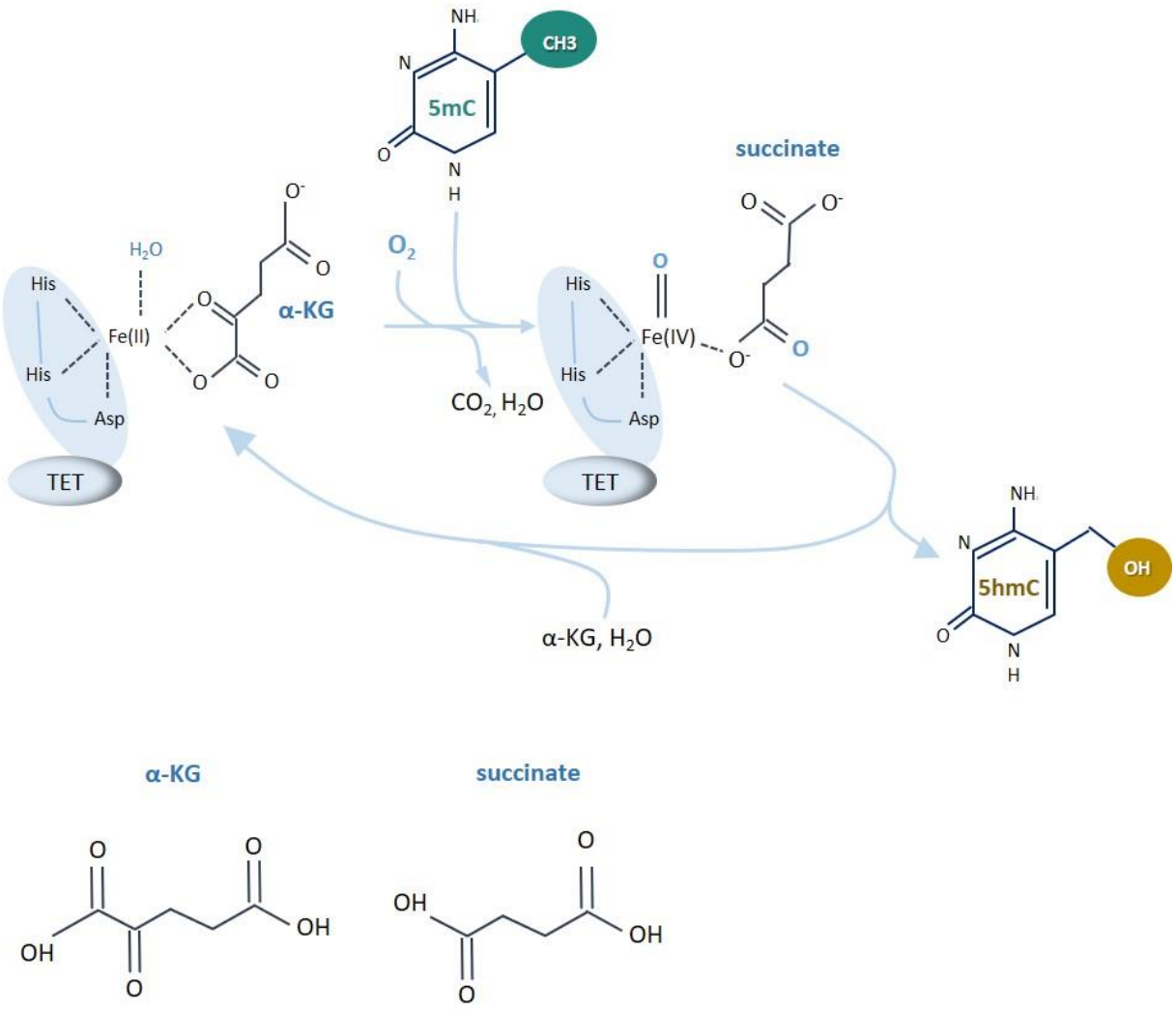
Cytosine methylation is a reversible modification. Critical phases in development require DNA demethylation such as methylation reprogramming during early embryogenesis and germ cell development (*See 1.5.1 DNA methylation reprogramming during development*), or demethylation of specific gene promoters during cell differentiation. Plants possess an active DNA demethylation system, catalysed by specific DNA glycosylases (ROS1, DME or DML2/3) capable of directly removing 5mC from DNA and replacing it by an unmethylated cytosine through BER pathways (Zhu, 2009). DNA glycosylases also exist in mammals (TDG, MBD4) but there is no evidence that these enzymes possess robust excision activity against 5mC, suggesting that direct BER of 5mC is not a major pathway for demethylation in mammals. Other possible mechanisms for DNA demethylation in mammals are discussed in this chapter.

1.4.1 Passive demethylation

Passive demethylation is the simplest model to explain loss of methylation marks in DNA: it occurs by replication-dependent dilution of 5mC in the absence of a functional DNA methylation maintenance system through cell divisions. In this way, DNA methylation is diluted by 50% after each round of cell division because methylation is not copied on the new DNA strand during the replication process (Razin and Riggs, 1980) (**Introductory Figure 5**). This can be achieved by downregulating the maintenance machinery (DNMT1, UHRF1) or inhibiting its access to chromatin. Recent work suggests that passive demethylation is a major contributor to epigenetic reprogramming in early embryos and germ cells (Messerschmidt et al., 2014).

1.4.2 TET enzymes and 5-hydroxymethylcytosine

The family of Ten eleven translocation (TET) enzymes is the only one for which we have strong evidence of its role on active DNA demethylation. Active demethylation pathways independent of replication have already been evoked during the 80's (Sullivan and Grainger, 1987; Wilks et al., 1984). Paroush et al., 1990 suggested that the activation of the *α-actin* gene in myoblasts required active demethylation steps in the absence of DNA replication, but they did not explain the molecular mechanisms. It was not until 2009 that TET1 was discovered. Trypanosomes have a DNA base called base J (β-D-glucosyl hydroxymethyluracil). The first steps for the synthesis of base J involve the hydroxylation of the methyl group of thymine by the proteins JBP1 and JBP2. By searching for homologues of these proteins in humans, Iyer et al., 2009 found 3 proteins: TET1/2/3. TET1 is the founding member of the family and a fusion partner of MLL gene in acute myeloid leukaemia (AML) (Lorsbach et al., 2003).



Introductory Figure 6. Catalytic reaction performed by TET enzymes.

Schematic representation of the reaction generating 5hmC by TET enzymes. Conserved residues in TET protein (His-His-Asp) bind Fe(II) and coordinate water and α-ketoglutarate (α-KG). Oxidation of α-KG leads to the formation of CO_2 and a high-valent Fe(IV)-oxo intermediate, still bound to the TET enzyme. This intermediate performs an oxidative transfer in the 5mC, leading to the formation of 5hmC and the regeneration of the Fe(II) species. Adapted from Kohli and Zhang, 2013.

i. TET proteins and oxidised derivatives of 5mC

TET proteins are 2-oxoglutarate-Fe(II)-dependent dioxygenases that catalyse the conversion of 5mC into 5hmC using α-ketoglutarate (α-KG) as a substrate (Ito et al., 2010; Tahiliani et al., 2009) (**Introductory Figure 6**). This reaction was first shown to take place in cultured cells and *in vitro* (Tahiliani et al., 2009). The same year, relatively high amounts of 5hmC (0,6% of total nucleotide) were reported in Purkinje neuronal cells (Kriaucionis and Heintz, 2009), suggesting a physiological role of this nucleotide in epigenetic control of gene expression. TET proteins perform also iterative oxidations leading to the formation of 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (**Introductory Figure 5**). 5fC and 5caC were found as products of TET active demethylation in mES cells and mouse organs (Ito et al., 2011; Pfaffeneder et al., 2011).

5hmC, 5fC and 5caC are considered intermediates of demethylation. There are several pathways that can lead to the replacement of these oxidised derivatives with unmodified cytosine. 5fC and 5caC can be efficiently excised by TDG and then repaired by the BER system via substitution with a C (He et al., 2011; Maiti and Drohat, 2011) (**Introductory Figure 5**). Multiple studies have biochemically validated this pathway. In addition, *Tdg* knockdown (KD) mES cells showed higher levels of 5fC and 5caC at different regions in the genome (Raiber et al., 2012; Shen et al., 2013). Recently, new candidates have been proposed to be involved in the excision procedure. UNG2, a uracil-DNA-glycosylase, was proposed to participate in the TET-mediated active demethylation pathway (Xue et al., 2016). Moreover, NEIL1 and NEIL2 cooperate with TDG in removing the oxidised methylcytosines (Schomacher et al., 2016). Alternatively, other studies show that oxidised products 5hmC, 5fC and 5caC may facilitate replication-dependent passive loss of 5mC during cell division as DNMT1 does not recognise them (Inoue et al., 2011). Indeed, it was established that the presence of 5hmC significantly decreases (by 60 fold) the catalytic activity of DNMT1 *in vitro* (Hashimoto et al., 2012). The decarboxylation of 5caC has also been proposed as a mechanism to return to the unmethylated cytosines (Schiesser et al., 2012), but the responsible enzyme has not been identified so far. One study suggests that the DNMT enzymes can directly remove the carboxyl group *in vitro* (Liutkevičiūtė et al., 2014). As both 5fC and 5caC are detected in mES cells, early embryonic development and neurons, they could be of physiological importance and help to regulate gene expression.

5mC levels are relatively constant between cell types, but the levels of 5hmC and the other intermediates are more variable. For example, 5hmC levels are higher in the nervous system compared to other cell types (Globisch et al., 2010; Ruzov et al., 2011), which could be related to the fact that most adult neurons are non-dividing cells. Moreover, there is a different distribution of 5mC, 5hmC and 5fC over the entire genome, meaning that TET proteins have different affinities for specific regions (Raiber et al., 2012). In ES cells, 5hmC occurs almost exclusively at CpGs and this modification is enriched at CpG-poor distal regulatory regions like enhancers (Yu et al., 2012). This correlates with the fact that the TFs binding leads to the demethylation in regulatory regions (Feldmann et al., 2013), as we discussed in *2.2 CpG-poor regions*. However, it does not correlate with TET1 binding sites, mostly found at CGIs in promoters. Consistent with that, CpG-rich promoters contain high levels of 5hmC and 5fC in CGIs surrounding the TSS, which could be involved in keeping them in a hypomethylated state (Raiber et al., 2012).

TET1 and TET2 proteins, but not TET3, are expressed in mouse ES cells. Recently it has been shown that miR-29 family regulates TET1 expression by directly repressing it during early differentiation in mES cells (Cui et al., 2016). TET2 is also abundantly expressed in the haematopoietic system, playing an essential role. Indeed, TET2 loss of function leads to various forms of haematological malignancies (*See 2.1.4 DNA methylation actors are frequently mutated in cancer - TETs*). TET3 is expressed in oocytes and has been proposed to participate in the global demethylation of the maternal and paternal genomes after fertilisation (Guo et al., 2014; Shen et al., 2014; Wang et al., 2014a). Even if each TET protein might play specific roles, some studies suggest that they also have redundant functions. For example, the combined deficiency of TET1 and TET2 shows that they have overlapping roles in embryonic development (Dawlaty et al., 2013; Huang et al., 2014).

ii. Structure of TET proteins

TET proteins contain a CXXC domain (with the exception of TET2), a cysteine-rich domain and a conserved catalytic domain (**Introductory Figure 4B**).

CXXC domain

The CXXC domain is found in several chromatin binding proteins such as DNMT1, MLL and CFP1, and it mediates binding to DNA containing unmethylated CpGs. TET1 and TET3 proteins (but not TET2) have a CXXC domain. The CXXC domain of TET1 potentially binds DNA; this function has been widely discussed but it remains controversial. Frauer et al., 2011 found that the CXXC domain of TET1 has no DNA binding activity and is not essential for the TET1 catalytic activity. On the contrary, there is evidence that the TET1 CXXC domain strongly binds CpG-rich DNA either unmodified or containing 5mC or 5hmC, and that this binding is impaired after mutation of the CXXC domain (Xu et

al., 2011; Zhang et al., 2010). In agreement with the binding of the CXXC domain to CpG-rich DNA, ChIP experiments revealed that TET1 binds mostly to CGIs in promoters and in intragenic regions. As discussed earlier, the binding of TET proteins to CpG islands prevents the accumulation of cytosine methylation in promoters, helping to maintain them in a hypomethylated state. Indeed, TET1 inhibition leads to increased levels of DNA methylation at many TET1 binding sites in promoters (Ficz et al., 2011; Williams et al., 2011; Wu et al., 2011a; Xu et al., 2011).

In mES cells, TET1 is also involved in facilitating the recruitment of Polycomb Repressive Complex 2 (PRC2) to promoters that are consequently marked with the H3K27me3 repressive histone modification (Wu et al., 2011a). This suggests that TET1 can also regulate gene expression independently of its catalytic activity. This idea is also supported by another study that demonstrates the role of TET1 in transcriptional repression by recruiting the SIN3A co-repressor complex at some loci (Williams et al., 2011). Taken together, these studies suggest that TET1 has a dual function in the regulation of gene expression by opposing DNA methylation at CpG-rich sequences and participating in the repression of some Polycomb-target genes.

It is important to notice that an isoform of TET1 coming from an alternative promoter has been discovered recently. $TET1^{ALT}$ lacks the CXXC domain. It is not expressed in ES cells, but it becomes activated in embryonic and adult tissues (Good et al., 2017). This isoform is aberrantly expressed in cancer, and it does not prevent CGI methylation, which correlates with the pattern of aberrant hypermethylation observed in tumor suppressor genes.

Interestingly, the *TET2* gene lost its CXXC domain during evolution through a chromosomal inversion and only codes for the catalytic domain. The CXXC domain initially encoded within the ancestral *TET2* gene is now encoded in the *IDAX* gene. The IDAX CXXC protein binds unmethylated cytosines *in vitro* and it is found at CGIs in promoters. IDAX interacts with TET2 and promotes the recruitment of TET2 to its target regions in DNA (Ko et al., 2013).

The TET3 CXXC domain can also bind to unmethylated CG-rich DNA, and is essential for the biological function of TET3, at least in *Xenopus* (Xu et al., 2012). Surprisingly, the CXXC domain of TET3 has a higher affinity towards 5caC. It suggests that it could be involved in the 5caC excision by the BER system, specifically in neuronal cells where it seems to localise at specific promoters of neuronal development (Jin et al., 2016).

Catalytic domain

The catalytic domain of TET proteins is characteristic of Fe2+ and α-KG-dependent dioxygenases (Loenarz and Schofield, 2008). It consists in a double-stranded beta helix localised in

the C-terminal end of the protein. A conserved histidine coordinates the Fe2+. TET proteins catalyse the oxidative decarboxylation of α-KG by using molecular oxygen as a substrate. As a result, the TET enzyme-bound high-valent Fe(IV)-oxo intermediate converts the 5mC into 5hmC by an oxidative transfer (Kohli and Zhang, 2013). In the human and mouse TET1 protein, the catalytic domain alone is sufficient to perform a similar catalytic activity than the full-length protein (Ito et al., 2010; Tahiliani et al., 2009) (**Introductory Figure 6**).

Vitamin C (ascorbic acid) acts as a cofactor of TET proteins enhancing the oxidation of 5mC into 5hmC in mES cells or Mouse Embryonic Fibroblasts (MEFs) (Blaschke et al., 2013; Minor et al., 2013; Yin et al., 2013). Vitamin C was shown to promote iPSC reprogramming (Esteban et al., 2010). Further studies showed that vitamin C regulates somatic cell reprogramming by modulating the activity of TET1 (Chen et al., 2013b).

iii. Physiological functions of Tet genes

Tet1 -/- and *Tet2* -/- mice are viable and fertile (Dawlaty et al., 2011; Quivoron et al., 2011). *Tet1* mutant mice are grossly normal and develop normally, although some mutant mice have a smaller size at birth (Dawlaty et al., 2011). The inactivation of TET2 is also compatible with normal development, but *Tet2* mutant mice develop haematological malignancies, highlighting the crucial role of TET2 in maintaining homeostasis of the haematopoietic system (Li et al., 2011b; Quivoron et al., 2011). Double inactivation of *Tet1* and *Tet2* is compatible with development, although some *Tet1* -/- *Tet2* -/- double mutant mice die around birth. Some of them overcome the loss of these enzymes by increasing the *Tet3* expression, suggesting that TET3 can compensate in order to maintain 5hmC levels in development. Therefore, the three *Tet* genes could have redundant functions during development (Dawlaty et al., 2013). Indeed, *Tet1*/*Tet2*/*Tet3* triple knockout (TKO) embryos show developmental defects after the gastrulation step (E7.5), indicating that TET proteins are required at this moment, possibly by regulating the *Nodal* signal pathway that is essential in mesoderm development (Dai et al., 2016). Single *Tet3*-/- mice die perinatally for unknown reasons (Gu et al., 2011b). Consistent with the high level of *Tet3* expression in oocytes, conditional KO of *Tet3* in the oocyte impairs genome demethylation after fertilisation, which leads to impaired development and a reduction in fertility (Gu et al., 2011b).

1.4.3 AID/APOBEC enzymes

The AID/APOBEC family of enzymes are able to deaminate cytidine in DNA and/or RNA (Conticello et al., 2005). AID (Activation-induced cytidine deaminase) is a protein encoded by the *Aicda* gene. It has a high sequence homology with apoB, an mRNA editing enzyme, so it was proposed that AID could have a role in mRNA editing (Muramatsu et al., 1999). However, AID was shown to be an important factor implicated in the process of antibody diversification in B lymphocytes through Class Switch Recombination (CSR) or Somatic Hyper Mutation (SHM) (Muramatsu et al., 2000; Revy et al., 2000). Mice mutants for the *Aicda* gene show defects in CSR and SHM (Muramatsu et al., 2000), while mutations of the gene in humans cause the hyper IgM syndrome type 2 (Revy et al., 2000). The mechanism of action of the enzyme is as follows: AID deaminates cytosine in ssDNA to form uracil as a product, which is then recognised as a mismatch U:G and excised by the uracil-N glycosylase. The abasic site is then repaired by the BER system. This repair might involve an error-prone polymerase as we find an unusual high error rate (Longerich et al., 2006; Noia and Neuberger, 2007). It has been proposed that a similar mechanism could be implicated in 5mC demethylation: 5mC would be deaminated by AID giving T as a product, and the T:G mismatch could be recognised by a glycosylase such as TDG or MBD4 and repaired by an error-free polymerase that would fill the abasic site with an unmethylated cytosine (**Introductory Figure 5**). In support of this, the TDG enzyme has been found to be essential for embryonic development (Cortellino et al., 2011). Moreover, *Tdg* mutant mice show aberrant patterns of DNA methylation and TDG forms a complex with AID and the DNA damage-associated protein GADD45 (Cortellino et al., 2011). However, *in vitro*, AID is less active on 5mCs than unmethylated cytosines, suggesting that AID might not directly deaminate 5mCs (Bransteitter et al., 2003; Larijani et al., 2005). Alternatively, AID could also deaminate the oxidation derivatives of 5mC such as 5hmC, creating 5-hydroxymethyluracil (5hmU) that could be excised by the 5hmU glycosylase SMUG1 and repaired with unmethylated cytosine (**Introductory Figure 5**). In support of this model, overexpression of AID stimulates 5hmC demethylation of transfected DNA in human HEK293 cells. The authors propose a model of active demethylation in which TET proteins first oxidise 5mC to 5hmC, which is then deaminated by the AID/APOBEC deaminases into 5hmU, excised by 5hmU glycosylases and repaired by the BER pathway with unmethylated cytosines (Guo et al., 2011) (**Introductory Figure 5**).

However, the role of AID in active demethylation remains unclear and debated. The first evidence of AID implication in 5mC demethylation was proposed by Morgan et al., 2004. They found that AID was able to mutate 5mC in *E. coli*, and that AID and APOBEC1 produced T after 5mC deamination *in vitro*. They also suggested a role of AID in epigenetic reprogramming during development as they found it expressed in oocytes, primordial germ cells (PGCs) and ES cells.

However, *Aid-/-* mice develop normally and do not present developmental failures, which argues against a role of AID in epigenetic reprogramming. Evidence of a role of AID *in vivo* was shown by the demethylation of exogenous DNA in zebrafish embryos, coupling deamination of 5mC into T by AID (or APOBEC2), followed by MBD4 glycosylase excision of the T (Rai et al., 2008). Other studies showed also efficient AID deamination-dependent demethylation of 5mC in zebrafish (Abdouni et al., 2013). Nevertheless, the results obtained in zebrafish were not reproducible by others (Shimoda et al., 2014).

In mammals, further studies showed that the deaminase activity of purified AID on 5mC or 5hmC could not be demonstrated (Nabel et al., 2012), but that another member of the family, APOBEC3A, has strong deamination activity towards 5mC but not towards TET-oxidised derivatives (Schutsky et al., 2017; Wijesinghe and Bhagwat, 2012). By using AID null fertilised mouse oocytes, the authors showed that AID was implicated in the global genome demethylation taking place in the mouse zygote (Santos et al., 2013). Moreover, AID deficient PGCs were used to study the potential role of AID in gametic methylation reprogramming. These cells presented a mild hypermethylation in introns, intergenic regions and transposons compared to WT cells, suggesting a role for AID in epigenetic reprogramming (Popp et al., 2010). Around the same time, AID was shown to be required to demethylate pluripotency-related gene promoters such as *Oct4* and *Nanog* in the process of generating iPSCs (Bhutani et al., 2010). A more recent study also showed that AID promotes demethylation and activation of the pluripotency genes *OCT4* and *NANOG* in somatic cell nuclear transfer (SCNT) experiments in bovine preimplantation embryos (Ao et al., 2016). However, here again, the role of AID in epigenetic reprogramming is still highly debated, as illustrated by other reports, suggesting that other factors than AID are required for epigenetic reprogramming in the mouse germline. Indeed, no *aid* expression can be detected in PGCs between E8.5 and E11.5, at the time of methylation reprogramming (Hajkova et al., 2010; Kagiwada et al., 2013).

In summary, numerous published articles indicate that the role of AID in cytosine demethylation is still controversial. It seems that AID may participate in locus-specific demethylation but not genome-wide methylation reprogramming in development. Further studies are necessary to fully understand its role and the interplay between AID and other active demethylation pathways.

1.5 Roles of DNA methylation in development

DNA methylation is linked to the regulation of gene expression during development, but also to the maintenance of genomic stability through the silencing of mobile elements. In the next paragraphs, I will explain briefly several points.

1.5.1 DNA methylation reprogramming during development

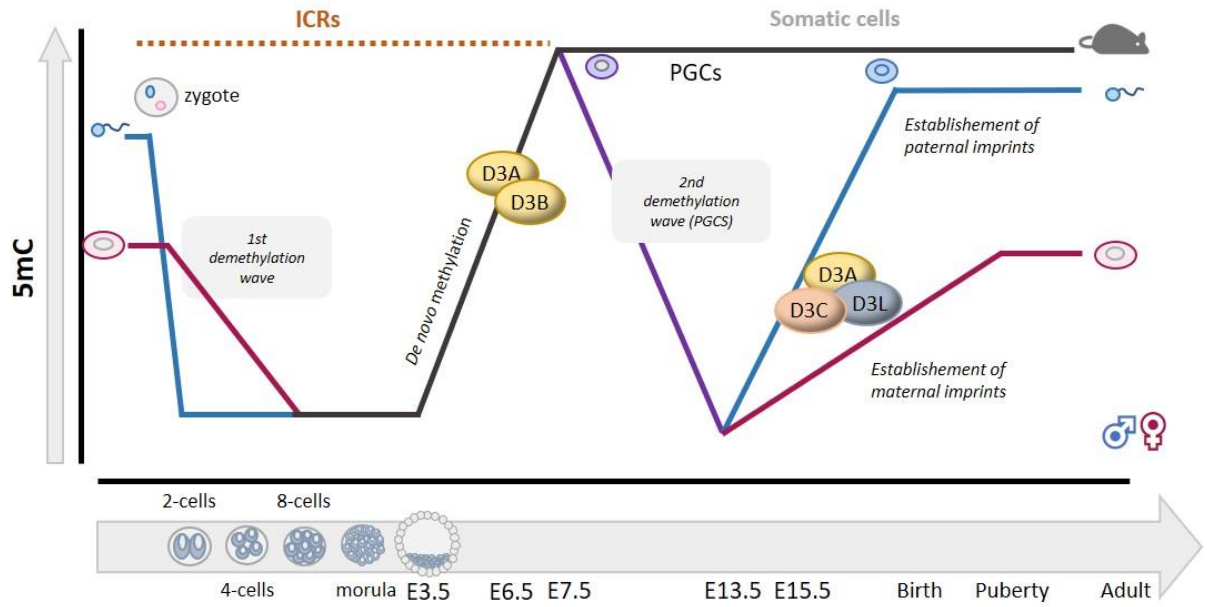
Epigenetic reprogramming, including DNA demethylation, takes place at two moments during mammalian development: a first wave of demethylation in the early embryo and a second wave of demethylation in PGCs (**Introductory Figure 7**). These waves are important to control the pluripotent state and to establish genomic imprinting.

i. Reprogramming during preimplantation development

DNA methylation is erased after fertilisation (MONK et al., 1987; Santos et al., 2002), reaching its lowest level at the blastocyst stage (embryonic day E3.5 in the mouse) (Smith et al., 2012). The parental genome undergoes rapid demethylation 4-5 hours after fertilisation. The mechanisms remain unclear, but it might be due to a combination of passive demethylation and active demethylation, probably involving the BER pathway (Hajkova et al., 2010; Santos et al., 2013). The TET3 hydroxylase is involved in this mechanism. Indeed, a TET3-dependent increase of 5hmC can be observed in the paternal pronucleus, associated with an accumulation of this enzyme in the male pronucleus (Gu et al., 2011b; Wossidlo et al., 2011). The maternal genome undergoes a demethylation that until recently was thought to be mostly replication-dependent (Saitou et al., 2012). Some reports suggested that the maternal genome is protected from TET3-mediated conversion to 5hmC by the PGC7 protein (also known as DPPA3 or STELLA) (Nakamura et al., 2007, 2012). However, more recent studies in the mouse using optimised bisulfite sequencing methods for low amounts of cells have shown that both paternal and maternal genomes contain oxidative derivatives of 5mC and undergo a TET3-dependent passive and active demethylation (Guo et al., 2014; Shen et al., 2014; Wang et al., 2014a). However, the active demethylation facilitated by TET3 is independent from TDG (Guo et al., 2014).

Some sequences in the genome escape this wave of demethylation: the imprinted loci and the repetitive elements (*See 1.5.3 Parental imprinting and 1.5.4 Maintenance of genome stability*) (**Introductory Figure 7**).

After embryo implantation, global DNA methylation is re-established by DNMT3A and 3B genome-wide, and the genome remains globally methylated in all somatic cell types of the body (**Introductory Figure 7**). CGIs are protected from this *de novo* methylation, except those in the promoters of developmental and gametic genes that acquire methylation preferentially by DNMT3B and are silenced (Auclair et al., 2014).



Introductory Figure 7. DNA Methylation reprogramming during mouse development.

After the fertilisation, a first wave of demethylation takes place and both parental genomes are demethylated in the preimplantation embryo until the blastocyst stage. Some transposon sequences and the imprinting control regions (ICR) are resistant to this demethylation. Methylation is established *de novo* by DNMT3A and DNMT3B after implantation. The genome will remain methylated in all somatic cells in adult tissues. Only the PGCs (Primordial Germ Cells) will undergo a second wave of demethylation between E8.5 and E13.5. Germ cells will be *de novo* methylated by DNMT3A, 3L and 3C. Methylation acquisition is different between sexes: male gametes are completely methylated before birth whereas female gametes acquire methylation after birth.

ii. Reprogramming in gametes

Germ cell progenitors, called PGCs (Primordial germ cells), arise in the epiblast at E6.5 in the mouse. As these cells are derived from somatic cells that already acquired a global somatic methylation, it is necessary that they undergo a global demethylation to restore their pluripotent status (Hackett et al., 2012; Saitou et al., 2012). They undergo a global genome demethylation between E9.5 and E12.5, in parallel to their division and migration to the genital crests. At E13.5 these cells present a globally demethylated genome. Kinetic studies during PGC development showed that imprinted regions, CGIs on the X chromosome and methylated CGIs of germline-specific genes become demethylated with a slower kinetics than the rest of the genome, suggesting a two-phase demethylation process in PGCs (Guibert et al., 2012; Seisenberger et al., 2012). Interestingly, some genomic regions resist this global demethylation and retain high levels of methylation in PGCs. These regions are mostly associated with repeated elements, and most of them are also resistant to demethylation in early embryos (Guibert et al., 2012; Seisenberger et al., 2012) (**Introductory Figure 7**).

Even if the global mechanism of demethylation is still debated, it is known that proteins implicated in the *de novo* and maintenance methylation (DNMT3A, DNMT3B, UHRF1) are downregulated in PGCs, and proteins involved in active demethylation, such as TET and BER, are expressed (Kagiwada et al., 2013; Seisenberger et al., 2012). The actual general consensus is that demethylation in PGCs occurs largely through a replication-dependent process (Kagiwada et al., 2013; Ohno et al., 2013), whereas TET1 is not required to initiate global demethylation but has a contribution in ensuring hypomethylation of defined sequences in the genome (Hill et al., 2018; Yamaguchi et al., 2012).

After global demethylation, germ cells acquire *de novo* methylation when they differentiate into gametes and undergo the meiosis process. This process is different in male and female gametes. In male gametes, DNMTA, 3L and 3C cooperate to establish *de novo* methylation of the genome after the E14.5 embryonic stage (Barau et al., 2016; Bourc'his and Bestor, 2004; Kaneda et al., 2004). On the contrary, female gametes are arrested in a hypomethylated state at the metaphase stage before birth, and oocytes are then remethylated essentially by DNMT3A and 3L during the phase of oocyte growth after birth (Smallwood et al., 2011) (**Introductory Figure 7**).

1.5.2 X chromosome inactivation

Because females have two X chromosomes whereas males have just one, a mechanism is required to compensate the gene dosage on the X chromosome. In mammals, this occurs by repressing one of the female X chromosomes (Lyon, 1961). A non-coding RNA (ncRNA) called *Xist* (X-

inactive-specific transcript) is essential to initiate the process of X-chromosome inactivation (Penny et al., 1996). It is expressed only from the X-inactivated chromosome and its product does not encode a protein (Brockdorff et al., 1992; Brown et al., 1991). *Xist* coats the inactive X chromosome (Brown et al., 1992) and leads to a series of epigenetic events triggering the X-inactivation *in cis*, which includes exclusion of the transcription machinery, recruitment of Polycomb proteins, histone modifications like H3K27me3 and variants like H2A, and DNA methylation of promoters CGIs (Barakat and Gribnau, 2012; Chadwick and Willard, 2001; Plath et al., 2003; Zhao et al., 2008).

X-chromosome inactivation starts randomly in either the paternal or the maternal chromosome during embryonic development (between E3.5 and E8.5) and is maintained in somatic cells. DNA methylation has been shown to take place only after the genes have been inactivated (Prissette et al., 2001). *De novo* methylation is not essential for the initiation of X chromosome inactivation (Sado, 2004), meaning that it contributes to the stabilisation of a long-term repression (Csankovszki et al., 2001; Sado et al., 2000). X-linked methylation is dependent of DNMT3B: mice deficient for *Dnmt3b* (but not *Dnmt3a*) show a total loss of methylation in X chromosome (Auclair et al., 2014). Surprisingly, some X-linked genes escape the process of inactivation, and this correlates with a lack of DNA methylation and H3K27me3 on their promoters (Auclair et al., 2014; Yang et al., 2010).

1.5.3 Parental imprinting

Imprinted genes are expressed exclusively from one allele depending on their parental origin: either the paternal allele or the maternal allele. There are approximately 150 imprinted genes in mice and most of them are also imprinted in humans. These genes are often organised in clusters, where their allele-specific expression is regulated by epigenetic modifications on CpG-rich regions called imprinted control regions (ICRs). ICRs harbour differential DNA methylation status between both alleles, which is acquired in the parental gametes and makes them germline DMRs (gDMRs). These ICRs can be in the promoter of genes or in intergenic regions. Methylated ICRs can directly repress the expression of an imprinted gene if they are in the gene promoter, or alternatively they can regulate several distant genes that are part of large imprinted regions. In this case, ICRs can function as insulators or promoters of long non-coding RNAs (ncRNA) that repress several distant genes in *cis* (Arnaud, 2010). Imprinting is a good example of how DNA methylation is not always linked to transcriptional repression: it can also activate gene expression by the exclusion of repressor proteins. For example, in the *Igf2/H19* locus, DNA methylation inhibits CTCF binding to allow *Igf2* activation through distant enhancers (Bell and Felsenfeld, 2000; Hark et al., 2000). Imprinting is established during gametogenesis. How is DNA methylation regulated in imprinted loci? There are three phases: *de novo* differential DNA methylation of the gametic DMRs (gDMR) is first established

in the parental cells during gametogenesis depending on the sex of the parent, then it is stably maintained after the fertilisation and during embryogenesis, and finally methylation imprints are erased in PGCs and reset according to the sex of the embryo.

The mechanism of the ICRs demethylation in the PGCs is not clear but, as discussed before, it could involve a combination of mechanisms including active demethylation by AID (*See 1.4.3 AID/APOBEC enzymes*) (Popp et al., 2010), TET proteins expressed in PGCs (Yamaguchi et al., 2013), and passive dilution (Kagiwada et al., 2013). The methylation imprints are then re-established in PGCs during the phase of genome remethylation that occurs at different developmental stages depending on the sex: paternal imprints are established after E14.5 and completed before birth (Li et al., 2004), whereas, on the contrary, maternal imprints are established after birth during the phase of oocyte growth and maturation (Lucifero et al., 2004). Conditional invalidation of DNMTs in PGCs have shown that the enzymes responsible for the *de novo* methylation of imprinted regions are DNMT3A and DNMT3L (Bourc'his, 2001; Bourc'his and Bestor, 2004; Kaneda et al., 2004; Kato et al., 2007).

Differential DNA methylation at imprinted loci is maintained during embryogenesis: the methylated allele escapes the first wave of DNA demethylation in the early embryo development while the non-methylated allele does not acquire *de novo* methylation. The principal mechanisms protecting imprinted alleles from demethylation in embryogenesis involve maintenance methylation by DNMT1, essential to maintain imprinted methylation patterns during early cleavage stages and embryogenesis (Hirasawa et al., 2008; Li et al., 1993), as well as additional factors MBD3, Stella/PGC7 and ZFP57 (Li et al., 1993, 2008; Nakamura et al., 2007; Reese et al., 2007). The protection from *de novo* aberrant methylation could be carried out by histone marks that exclude the DNMTs (Stewart et al., 2015).

Interestingly, in addition to germline DMRs that are stably propagated in development, it has been shown that there are also many secondary somatic DMRs (sDMRs) established in imprinted loci after implantation in a DNMT3B-dependent manner (Auclair et al., 2014; Bhogal et al., 2004; Sato et al., 2011).

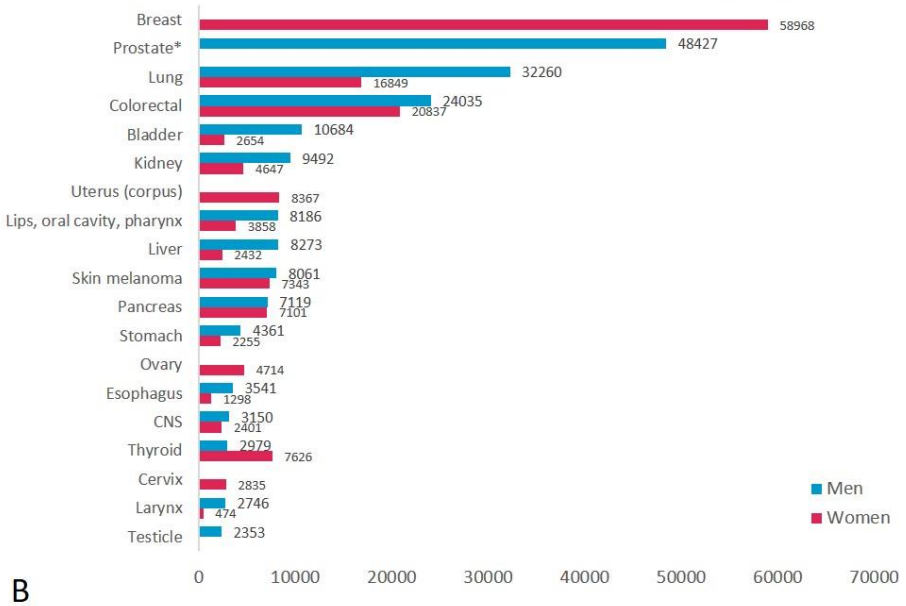
1.5.4 Maintenance of genome stability

A big proportion of mammalian introns and intergenic sequences contains non-coding DNA or repetitive DNA. There are several types of repeated regions in the genome: tandem repeats (major and minor satellites mostly in the pericentromeric regions) and transposable elements (TEs) such as long interspersed nuclear elements (LINE), short interspersed nuclear elements (SINEs) and long terminal repeats (LTR). TEs constitute almost 50% of the human genome. The expression of

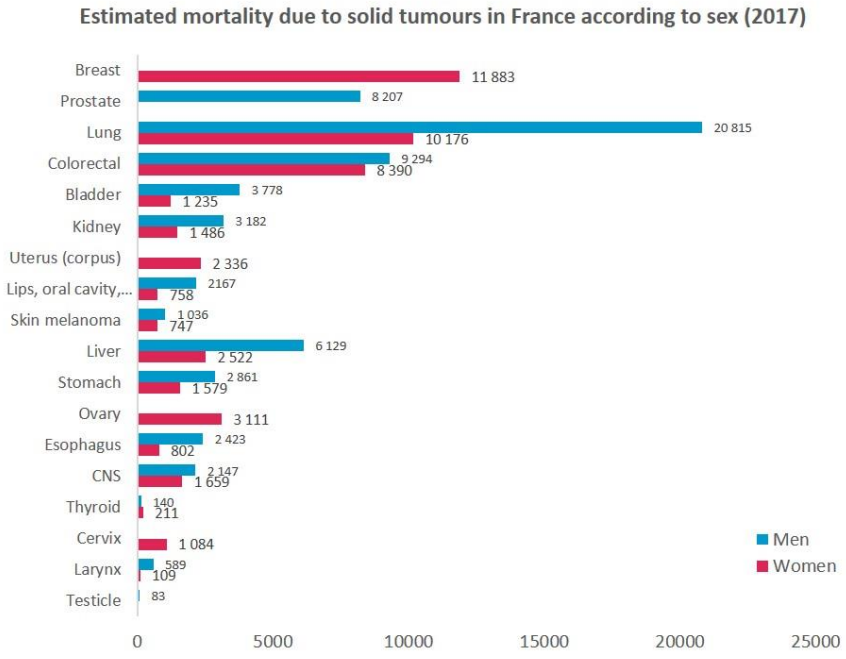
transposons can potentially impair genome integrity and stability because they are mutagenic, causing insertion in or around protein-coding genes and chromosomal rearrangements.

One of the main defence mechanisms involved in repressing the expression and mobility of transposons in mammalian genomes is DNA methylation. In mice, DNMT1 and UHRF1 are essential to maintain TEs repression in the genome: absence of DNA methylation maintenance in *Dnmt1-/-* or *Uhrf1-/-* mice leads to a reactivation of TEs like intracisternal A-particles (IAPs), LINE-1 and SINE-1 (Sharif et al., 2007; Walsh et al., 1998). In addition, there are molecular pathways required to specifically recruit DNA methylation at TEs in development, in particular TRIM28 (also known as KAP1) and the histone methyltransferase SETDB1 (also known as ESET). These proteins are recruited to TEs by KRAB zinc-finger protein (ZFP) in order to catalyse H3K9 methylation and facilitate the recruitment of DNA methylation at these regions (Leung et al., 2014; Rowe et al., 2013). The epigenetic control of TEs is also essential to preserve genome integrity in germ cells. DNMT3L and DNMT3C are essential to establish DNA methylation of TEs in mouse gametes (*see1.3.3 DNMT3L and 1.3.4 DNMT3C*). Failure to methylate TEs in germ cells of *Dnmt3l-/-* or *Dnmt3c-/-* mice leads to robust expression of evolutionarily young retrotransposons during spermatogenesis at the onset of meiosis, which leads to meiotic catastrophes and a lack of production of functional spermatozoa (Barau et al., 2016; Bourc'his and Bestor, 2004; Zamudio et al., 2015). A mechanism implicating PIWI-interacting RNAs (piRNAs) has evolved to specifically induce TEs silencing and initiate the *de novo* DNA methylation of TEs in the male germline (Aravin et al., 2008; Kuramochi-Miyagawa et al., 2008).

A



B



Introductory Figure 8. Projection of cancer incidence and mortality in France in 2017.

A. Classification of solid tumors by estimated incidence in metropolitan France in 2017.
B. Classification of solid tumors by estimated mortality in metropolitan France in 2017.
*Data for 2017 projection in prostate cancer are not available; data from 2013 have been used in this graphic.
Source: Partenariat Francim/HCL/ Santé Publique France /INCa

2 DEREGULATION OF DNA METHYLATION IN CANCER

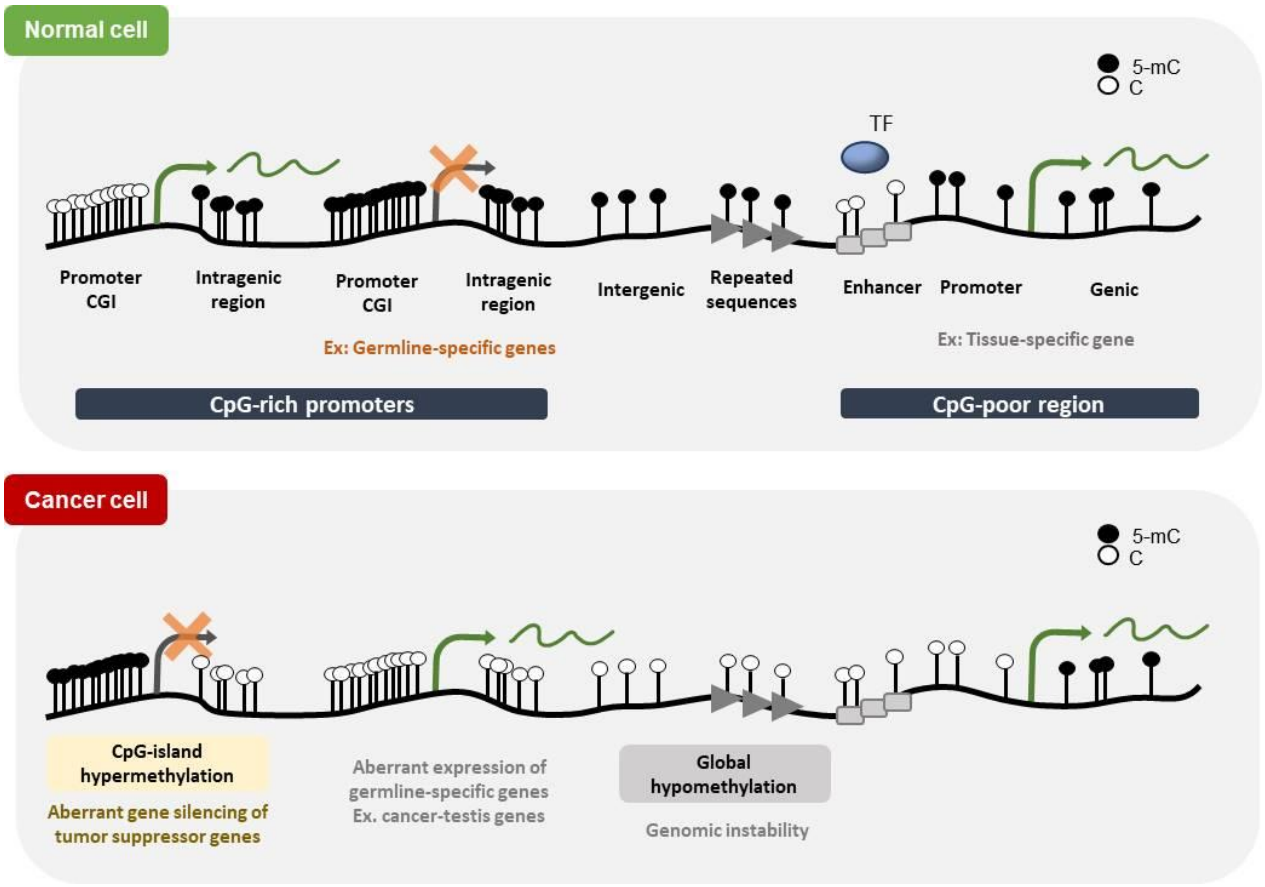
Cancer is one of the leading causes of death in the world. Data from World Health Organisation (WHO) estimate that almost 9 million people died in 2016 because of this disease. Upper-middle income and high-income countries present higher rates of deaths due to non-communicable diseases than low-income countries, and several types of cancer are found among these diseases (lung, liver, stomach, kidney and breast).

New cancer cases are increasing each year. In France, around 400.000 people were diagnosed in 2017 and around 150.000 unfortunately died because of cancer. In France, the most common cancer in men is the one affecting the prostate, whereas the most common cancer in women is breast cancer. Concerning the mortality, lung cancer is the most lethal in men and breast cancer in women (**Introductory Figure 8**) (Jéhannin-Ligier et al., 2017).

Cancer can affect any part of the body. We distinguish "solid tumors" from cancers affecting the lymphatic or haematopoietic system. The main characteristic of all types of cancer is that cells grow in an uncontrolled way and are disseminated through the body. Malignant tumors are constituted from cancer cells and from non-cancer tissues like blood vessels or immune cells that help to create a favourable microenvironment for tumor growth. Cancer cells invade surrounding tissues and they can also be detached from the parental tumor and reach distant parts of the body in a process called metastasis.

Even if current advances allow to cure cancer in almost half of the patients, prognosis is variable from one patient to another. It depends on different factors like the tissue affected, the state of the tumor at the moment of detection, the presence of metastasis or genetic factors.

Genetic and epigenetic events are involved and cooperate in tumorigenesis and tumor development. Since 30-50% of cancer cases could be prevented, detection tools have become essential. In this chapter I will discuss DNA methylation implication in cancer and how it can be used as a biomarker. In fact, DNA methylation-based methods can help to detect cancer at early stages and to choose the correct treatment of the disease. Besides, because of its reversible nature, it is also a potential therapeutic target for anti-cancer therapies.



Introductory Figure 9. Altered DNA methylation profiles in cancer cells.

Cancer cells harbour a global hypomethylated genome. This hypomethylation is related to the aberrant activation of genes with oncogenic potential such as cancer-testis genes and genomic instability. Recent whole genome bisulfite sequencing (WGBS) studies have confirmed that almost half of the genome in cancer cells is hypomethylated. Focal aberrant hypermethylation of CGIs is another hallmark of cancer cells. This hypermethylation is common in the promoter of tumor suppressor genes and is linked with the repression of genes regulating key functions in cells.

2.1 DNA methylation perturbations in cancer and other diseases

Somatic mutations have been widely studied in cancer initiation and progression. Key genes in cell cycle regulation or DNA repair are frequently deleted or mutated in tumors. But these mutations are not the only drivers of tumorigenesis: a large number of genes are also epigenetically silenced through aberrant promoter DNA methylation in cancer cells. Both genetic and epigenetic alterations cooperate and act synergistically during tumor evolution to drive both gene loss or gain of function, which is translated in aberrant gene expression. Despite aberrant promoter CGIs methylation, cancer cells present a globally hypomethylated genome (**Introductory Figure 9**). Studying the connections between genetic and epigenetic alterations constitute a major objective in cancer research, with potential clinical applications. We are going to discuss in the next paragraphs the roles of DNA methylation in cancer.

2.1.1 Genome-wide methylation studies in cancer

Substantial genome hypomethylation in cancer cells was already observed in the 80's in different kind of tumors (Feinberg and Vogelstein, 1983; Gama-Sosa et al., 1983). On the other hand, one common feature of cancer cells is the aberrant hypermethylation of CGIs in gene promoters. These genes are usually tumor suppressor genes implicated in crucial biological functions for the cell. Altogether, hyper or hypomethylation events suggest that DNA methylation balance is necessary to ensure the correct functioning of the cells.

Until recently it was difficult to precisely compare global DNA methylation pattern in tumors and healthy tissues because of technical restrictions. In recent years, technical advances have overtaken these barriers with the development of methods like Whole-Genome Bisulfite Sequencing (WGBS), allowing to perform genome-wide studies that give novel insights into the methylome of cancer cells.

Recent methylome studies have highlighted several points that repurpose cancer DNA methylation hallmarks:

(i) Global loss of DNA methylation in cancer cells takes place in megabase-scale partially methylated domains (PMDs) or hypomethylated blocks that can make up almost half of the genome. PMDs were first described in cultured fibroblasts (Lister et al., 2009), and reduction of methylation in PMDs is widely observed in several cancer types: colon (Berman et al., 2011; Vidal et al., 2017), breast (Hon et al., 2012; Legendre et al., 2015) and other solid tumors. Megabase-scale PMD hypomethylation is now viewed as a universal epigenetic marker of most cancer types and PMDs are less hypomethylated in most of the normal tissues (Timp et al., 2014; Zhou et al., 2018). However,

large megabase regions of hypomethylation are not observed in haematological cancer (Wahlberg et al., 2016).

This global loss of methylation in PMDs occurs generally in gene-poor, low CG density, late replicating regions with weak regulatory activity (Berman et al., 2011; Vidal et al., 2017). But, when it occurs in a promoter or enhancer region, the loss of methylation is associated with increased expression of the related gene (Berman et al., 2011). Other reports correlate global hypomethylation with gene silencing associated with gain of repressive chromatin marks like H3K27me3 or H3K9me3 (Hon et al., 2012). It remains to be clarified if PMD hypomethylation is a driving force of cancer or a side effect of cancer cells proliferation.

(ii) The decrease in global methylation is progressive through tumor progression, being more intense in metastasis samples than primary tumors (Vidal et al., 2017)

(iii) Cancer samples can be clustered according to their tissue of origin (Vidal et al., 2017).

(iv) Global genome hypomethylation is accompanied by focal hypermethylation of specific DNA regions in cancer cells. This gain of aberrant methylation takes place frequently at CGIs in gene promoters and is associated with silencing of genes like tumor suppressor genes (TSGs). These genes are silenced either unselectively in all cancer types or in a cancer type-specific manner (Vidal et al., 2017). Interestingly, several reports show that genes associated with aberrant gain of methylation in CGIs are frequently not expressed in normal cells (Wahlberg et al., 2016), meaning that aberrant CGIs methylation is not necessarily driving changes in gene expression.

(v) Methylation-prone sequences are rich in repressive histone marks like H3K27me3 (Berman et al., 2011). On the contrary, methylation resistant sequences show an enrichment for enhancer marks such as p300 or H3K27ac (Berman et al., 2011; Heyn et al., 2016). Hypomethylated enhancers present an enrichment at TFs binding sites known to protect DNA from methylation (Sp1, Nrf1…) (Berman et al., 2011) or to promote oncogenic activity (Heyn et al., 2016).

(vi) DNA hypomethylation in cancer cells is frequent in repetitive elements, which significantly contributes to overexpression of repeats in cancer cells (Hon et al., 2012).

In the next paragraphs, two of these cancer methylome studies are described in more detail.

Berman et al., 2011 was the first study comparing the methylome of human colorectal tumors and surrounding normal tissues by WGBS. They showed the existence of long-range regions (>100 kb) hypomethylated in the tumor that represent almost half of the total genome. These regions correlated with nuclear lamina association domains, suggesting a role of DNA hypomethylation in chromatin organisation. The loss of methylation in these regions exposed (with

low frequency) enhancers and transcription factor binding sites and was associated with genes upregulation. At the same time, focal hypermethylation was observed within the hypomethylated regions. 95% of methylation prone elements correlated with CGIs and were known to be enriched with Polycomb repressive marks in hESC. From those CGIs, 29% correlated with known promoters and methylation was associated with gene repression. Interestingly, methylation resistant elements were enriched with TFs binding sites as Sp1 or Nrf1. Those TFs had already been described for their roles in DNA methylation protection in cancer cells (Gebhard et al., 2010).

A more recent study carried out by the group of Manel Esteller reported WGBS methylomes in 8 normal tissues and 13 associated cancer samples (Vidal et al., 2017). Not surprisingly, global DNA methylation was reduced in primary cancer samples compared to normal tissues, with a significant diminution in repetitive elements. In addition, they observed hypomethylated regions (HMRs) related to transcription control sequences in normal tissues that they divided in c-HMRs (common for all cell types) or t-HMRs (tissue-specific). 12% of c-HMRs acquired methylation in cancer. Interestingly, tumors could be clustered depending on their c-HMRs methylation status. A remarkable point of the article is the comparison among three samples coming from the same individual: normal tissue, primary colorectal cancer and liver metastasis. They found that the progression from the healthy tissue to the primary tumor, and in turn the metastases is coupled to increased genome hypomethylation during tumorigenesis: hypomethylated DMRs passed from 87663 in the primary tumor to 205459 in the liver metastases. Higher 5mC depletion in advanced or metastatic tumors has already been associated with bad prognosis (Li et al., 2014). At the same time, tumors gained methylation at specific loci, with the highest levels observed in the metastatic samples, but the number of hypermethylated DMRs augmented only slightly in metastases compared to the primary tumor (12364 in primary tumor, 15373 in metastasis). These data confirmed that large-scale loss coupled to focal gain of DNA methylation is a hallmark of human cancer progression.

It is not clear what mechanisms drive genome hypomethylation in cancer cells, and whether hypomethylation is a driver or a consequence in tumor progression. It has been found recently that PRMT6, an arginine methyltransferase catalysing asymmetric dimethylation of histone H3 arginine 2 (H3R2me2a), negatively regulates DNA methylation by inhibiting the association of UHRF1 to chromatin, and its overexpression is correlated with global hypomethylation in cancer (Veland et al., 2017). As this protein is overexpressed in several types of cancer, its upregulation could be one mechanism leading to the hypomethylated pattern observed in cancer cells (Yang and Bedford, 2013). Alternatively, PMD hypomethylation could simply reflect the accumulation of cell divisions because late replicating regions are more prone to incomplete maintenance methylation in each

round of cell division (Zhou et al., 2018). Further studies should be carried out to decipher the events involved in cancer-specific global DNA hypomethylation.

2.1.2 Consequences of global hypomethylation in cancer cells

Several consequences in cancer cells have been attributed to the global loss of methylation, such as activation of TEs, genomic instability, activation of normally silenced genes and loss of imprinting. These changes in genome stability and gene expression caused by the hypomethylated status of cancer cells all contribute to create a selective advantage for the cancer cells, promoting tumor progression.

i. Repetitive sequences

As we saw in chapter I (*see 1.5.4 Maintenance of genome stability*), transposon sequences in the genome are silenced by DNA methylation, contributing to genome stability. Hypomethylation in cancer cells can lead to the expression and transposition of repeated sequences like the mobile elements LINEs, SINEs and LTR transposons. LINE-1 elements represent approximatively 17% of human genome, so their methylation degree is considered as a mark of global methylation status (Cordaux and Batzer, 2009). In the context of cancer, LINE-1 hypomethylation and reactivation have been documented in many cancers and are associated with poor prognosis in several types of cancer like colorectal, oesophageal, lung or gastric cancers (Iwagami et al., 2013; Kawano et al., 2014; Saito et al., 2010; Shigaki et al., 2013; Swets et al., 2016).

The consequences of transposon overexpression in cancer cells are genomic instability and aberrant gene expression. Most TEs in the human genome are not capable of moving, but a small proportion remains active. Activation of these TEs can lead to mutagenic retrotransposition events that disrupt gene function (by mutating the open reading frame or the *cis*-regulatory elements, or by influencing mRNA splicing and stability) and promote cancer progression. This was first observed in the *APC* gene in colorectal cancers (Miki et al., 1992). More recently, this event has been observed in the tumor suppressor genes *MCC* and *ST18* (Shukla et al., 2013). High-coverage whole-genome sequencing of cancer samples revealed hundreds of somatic TEs insertion events, mostly LINE-1 insertions, in epithelial tumors (Lee et al., 2012a). Importantly, many of these insertions occur near genes with known tumor suppressor functions, suggesting a potential contribution of LINE-1 insertions to cancer development (Lee et al., 2012a). Global genome hypomethylation also induces genomic instability (Gaudet, 2003), and one possible explanation is that the insertion of TEs like LINE-1 promotes genome instability by inducing chromosomal breakages (Liu et al., 1997).

ii. Gene activation

Global hypomethylation can lead to gene activation in several ways. Importantly, not all reactivated genes in cancer are oncogenic. For an activated gene to be considered as oncogenic, it should be implicated in functions like cell proliferation or angiogenesis.

Even if they cannot jump within the genome, hypomethylated expressed transposon sequences can also impact cancer development by regulating the expression of nearby genes. As LINE-1 are distributed all along the genome, gene expression can be up or downregulated depending on where they are located. For example, proto-oncogenes *MET*, *RAB3IP* and *CHRM3* are activated as a consequence of the demethylation of LINE-1 elements that are located within their enhancers or promoters in CRC metastasis development (Hur et al., 2014). Another example is the hypomethylation of a LINE-1 element in the intron 18 of the *ALK* gene correlating with the activation of novel cancer-specific oncogenic *ALK* transcript in melanomas (Wiesner et al., 2015). LINE-1 promoter hypomethylation can also activate an alternative transcript of the *MET* oncogene in bladder cancer (Wolff et al., 2010). However, if LINE-1 sequences are present in the intragenic body, their hypomethylation can lead to transcriptional repression (Aporntewan et al., 2011).

The loss of methylation in gene bodies can also expose cryptic enhancers, alternative promoters or alternative splicing sites. For example, characterisation of DNA methylomes of patients with chronic lymphocytic leukaemia (CLL) revealed gene-body hypomethylation in enhancer regions, potentially affecting the binding of transcription factors and playing a functional role in leukaemogenesis (Kulis et al., 2012).

Another consequence of global hypomethylation is the aberrant activation of “cancer-testis” (CT) genes leading to the formation of cancer-testis antigens (CTA), which have an immunogenic potential. These genes are expressed exclusively in germ cells and selectively silenced in somatic tissues through processes involving promoter CGI methylation (Weber et al., 2007). Thus, DNA methylation is the main mechanism required to restrict CT genes expression in somatic cells. CT genes can be testis specific or testis predominant, meaning that low doses of expression have been observed in somatic cells (da Silva et al., 2017; Van Tongelen et al., 2017). The biological functions of CT genes are very diverse during gametogenesis, meaning that they could contribute to different functions in tumor development (Whitehurst, 2014). Gametogenesis and tumorigenesis share some phenotypical characteristics like hypomethylation, activation of silenced genes and angiogenesis (Simpson et al., 2005), and the aberrant re-expression of CT genes in tumors suggests a potential oncogenic role and involvement in tumorigenesis (Wang et al., 2016). In a recent review from the De Smet group, they classify methylation-dependent CT genes in relation to cancer initiation and

progression events. For example, there are CT genes associated with cell proliferation (*MAEL*, *DDX3*), whereas others are associated with immortality (*BORIS*/*CTCFL*) or metastasis (*CT-GAVRA3*) (Van Tongelen et al., 2017). Different genome-wide studies have been developed to identify potential CT genes. For example a recent study used transcriptomic data to identify CT-genes predominantly expressed in testis and tumors, and they found that some CT genes are expressed in different types of tumors (e.g. *PBK*, *SPATA22*, *IL4I1*, *HIST1H1A*) whereas others are mostly tumor type-specific such as *C17ORF104* (melanoma), *FUT5* (colon), *PAGE1* (ovary) and *CSNK1A1L* (prostate) (da Silva et al., 2017). CT genes can be associated with good or bad prognosis. Interestingly, there is evidence that the CTA related with good prognosis produce a signal for CD8+ entry in the tumor, leading to the induction of apoptosis (da Silva et al., 2017). As CTA are not expressed in normal somatic cells, they represent an interesting target for the development of targeted therapies in cancer, including immunotherapy (Simpson et al., 2005).

iii. Loss of imprinting (LOI)

Loss of imprinting (LOI) is another molecular characteristic of tumor cells that is linked to the loss of methylation. LOI implies loss of monoallelic gene regulation, activating the biallelic expression of a proto-oncogene or repressing the normally active copy of a gene with tumor suppressor characteristics. To test the function of LOI in tumorigenesis, Holm and colleagues performed transient conditional inactivation of *Dnmt1* in mESCs (Holm et al., 2005). This generated ESCs that lost all imprinting methylation marks and in which the reactivation of DNMT1 restored global DNA methylation but failed to re-establish methylation patterns in ICRs. MEF cells derived from these ESCs presented characteristics of transformed cells like a high growth rate and immortality, and injection of these cells in immunodeficient mice lead to tumor formation. These results suggested a potential causal role for LOI in cancer.

Several imprinted genes have been associated with tumor development. One of the most common LOI marks in cancer is the activation of the silenced maternal copy of *IGF2* (insulin-like growth factor II), a growth promoting gene. The biallelic aberrant expression was first reported in Wilms tumors of the kidney, the most common tumor in children (Ogawa et al., 1993; Rainier et al., 1993). Since then, LOI in *IGF2* has been reported in colorectal, liver, lung and ovary cancers (Cui, 2007). In chronic myeloid leukaemia (CML), it appears to be linked with the progression of the disease (Randhawa et al., 1998).

2.1.3 Consequences of the focal hypermethylation of CGIs

Hypermethylation of CGIs is a hallmark of cancer cells. It was first observed in the promoter of the *RB* gene implicated in retinoblastoma tumors (Greger et al., 1989). Besides CGI hypermethylation of genes with known tumor suppressor functions, aberrant CGI methylation can also take place in the promoters of upstream TFs and non-coding RNAs (ncRNAs) that modulate expression of genes with tumor suppressor potential.

i. Tumor suppressor genes

Even if all TSG are not methylated in all kind of tumors, TSG CpG island-promoter hypermethylation is a hallmark of all cancers, affecting 5-10% of CGIs that are unmethylated in normal cells (Baylin and Jones, 2011). We can talk about a "DNA methylation signature" that defines tumor subtypes and that could be used as a biomarker for early detection and prognosis. Epimutations are as important as genetic mutations in cancer development, and there is a remarkable correlation between both of them. Indeed, genetic and epigenetic inactivation in cancer frequently target the same genes, and combining genetic and epigenetic studies to identify genes that are subject to both mutations and hypermethylation help to identify key tumor suppressor genes that strongly correlate with poor outcome in breast and colorectal cancers (Chan et al., 2008; Yi et al., 2011). Genes that are commonly silenced by DNA hypermethylation participate in processes such as DNA repair (*BRCA*, *MLH1*), immortality (*CASP8*) or metastasis (*CADM1*) and angiogenesis (*THBS-1*). In the **Introductory Table 1** there is a non-exhaustive compilation of coding and non-coding genes silenced by CGIs hypermethylation in human cancer.

However, the functionality of DNA-methylation dependent silencing of tumor suppressor genes and its role in cancer progression is debatable. First, some TSGs are methylated in some type of tumors but not in others. Second, some genes are overexpressed in some tumors but downregulated in others. That means that there are genes that can have both proto-oncogenic or tumor suppressor functions (Chen et al., 2018; Deng, 2009; Shen et al., 2018). And finally and most importantly, there are many cases where aberrant CGIs hypermethylation takes place in genes that were already repressed in non-cancerous tissues, suggesting that their hypermethylation does not participate in the silencing and thus in the tumoral progression (Sproul et al., 2012; Wahlberg et al., 2016). There is evidence that TSGs repressed by aberrant DNA hypermethylation are marked with the repressive histone mark H3K27me3 prior to their silencing. That would mean that genes are tagged for further repression via methylation by Polycom complex in a similar manner that occurs during embryonic development (Gao et al., 2014a; Schlesinger et al., 2007).

Involved in..	*Gene*	Gene name	Consequence	Tumor-type	References
DNA repair	*BRCA1*	Breast cancer 1	Genomic instability	Breast, ovary	Esteller et al., 2000a; Pathania et al., 2011
	MLH1	MutL Homolog 1	Genomic instability	Colon, endometrium, stomach	Kidambi et al., 2016; Simpkins et al., 1999
	MGMT	O6-methylguanine-DNA methyltransferase	Genomic instability	Multiple types	Wang et al., 2008
	WRN	Werner syndrome ATP-dependent helicase	Genomic instability	Cervical cancer	MASUDA et al., 2012
Cell cycle control	*p15INK4b*	Cyclin Dependent Kinase Inhibitor 2B	Aberrant division	Leukaemia	Humeniuk et al., 2013; Krajnović et al., 2013
	p16INK4a	Cyclin Dependent Kinase Inhibitor 2A	Aberrant division	Multiple types	Krajnović et al., 2013
	RASSF1A	Ras association domain family member 1	Aberrant division	Multiple (breast, ovarian, lung, prostate, colon)	Hesson et al., 2007
	RB1	Retinoblastoma	Aberrant division	Retinoblastoma, glioblastoma	Nakamura et al., 2001; Ohtani-Fujita et al., 1997
	SEPT9.	Septin-9	Aberrant division	Colorectal	Nian et al., 2017
	WNT5A	Wnt Family Member 5A	Aberrant division	Colorectal	Galamb et al., 2016
	APC	Adenomatous polyposis coli	Aberrant division	Colorectal, breast	Liang et al., 2017; Zhou et al., 2016
Cell adhesion	*CADM1*	Cell adhesion molecule 1	Metastasis	Multiple types (CRC, cervical, liver)	Baars et al., 2016; Chen et al., 2011; Zhang et al., 2011b
	CDH1	Cadherin-1	Metastasis	Breast, stomach, leukaemia, prostate	Caldeira et al., 2006; Keil et al., 2014; Machado et al., 2001
	CDH11	Cadherin-11	Metastasis	Colon, breast, oesophagus, gastric, bladder	Li et al., 2012; LIN et al., 2015
	DACH1	Dachshund Family Transcription Factor 1	Metastasis	Mutiple types (stomach, breast, prostate)	Chen et al., 2015a; Yan et al., 2014; Zhao et al., 2015
	EXT1	Exostosin-1	Metastasis	Leukaemia, skin	Ropero, 2004
	TIMP3	Metalloproteinase inhibitor 3	Metastasis	Stomach, HNSCC, breast	Guan et al., 2013; Lui et al., 2005; Rettori et al., 2013
	LAMA3	Laminin subunit alpha 3	Metastasis	Breast, prostate, bladder	Sathyanarayana et al., 2003a, 2003b, 2004
Apoptosis	*CASP8*	Caspase-8	Immortality (inhibition of apoptose)	Glioblastome	Skiriute et al., 2012
	DAPK	Death-associated protein kinase 1	Immortality (inhibition of apoptose)	Gastrointestinal, cervical, lung, head and neck	Agodi et al., 2015; Sanchez-Cespedes et al., 2000; Tang et al., 2000; Yuan et al., 2017
	TMS1	Target of methylation-induced silencing 1	Immortality (inhibition of apoptose)	Breast, prostate, ovary	Das et al., 2006; GORDIAN et al., 2009; Levine et al., 2003; Terasawa et al., 2004
	MASPIN	Mammary serine protease inhibitor	Immortality (inhibition of apoptose)	Breast, prostate, colon, lung	Berardi et al., 2013; Maass et al., 2002
Angiogenesis	*THBS-1*	Thrombospondin-1	Anti-angiogenic	Neuroblastoma	Yang et al., 2003
Transcription factors	*GATA4* (activator)	GATA Binding Protein 4	Aberrant silencing of target genes	Colon, stomach	Akiyama et al., 2003; Hellebrekers et al., 2009
	GATA5 (activator)	GATA Binding Protein 5	Aberrant silencing of target genes	Colon, stomach	Akiyama et al., 2003; Hellebrekers et al., 2009
	HIC1 (repressor)	Hypermethylated in Cancer 1	Aberrant activation of target genes	Prostate, NSCLC, breast, gastric, haematological malignancies	Zheng et al., 2012
ncRNA	*MIR-127*	Target gene: BCL6	Aberrant division	Prostate, bladder, colon	Saito et al., 2006
	MIR-124a	Target genes: CDK6, Rb	Aberrant division	Mutiple (colon, breast, lung, leukaemias, lymphomas)	Lujambio et al., 2007
	MIR-145	Target genes: MUC1, FSCN1, NEDD9 and SOX9	Metastasis	Multiple (CRC, lung, breast, bladder, prostate)	Cui et al., 2014
	MIR-375	Target gene: IGF1R	Aberrant division	Oesophageal squamous cell carcinoma, breast	Kong et al., 2012; Liu et al., 2016

Introductory Table 1. Selection of Tumor Suppressor Genes silenced by DNA promoter hypermethylation in human cancers.

Adapted from Esteller, 2007; Heyn and Esteller, 2012; Llinàs-Arias and Esteller, 2017.

The role of promoter DNA hypermethylation in cancer initiation is evidenced by a report showing that methylation of two well established tumor suppressor genes (*HIC1* and *RASSF1A*) is sufficient to initiate cellular transformation and tumorigenic phenotype (Teng et al., 2011). Targeted epigenome editing tools have also become valuable tools to test the causal role of DNA methylation. For example, targeted promoter methylation of *MASPIN* enhanced breast cancer cell colony production, suggesting indeed that DNA hypermethylation of key genes can drive cancer progression (Rivenbark et al., 2012). Targeted epigenome editing will be discussed in more details in Chapter 3.

ii. Transcription factors (TFs)

Hypermethylation can also take place in the CGIs of promoters of upstream TFs, which in turn leads to downregulation of the expression of their target genes. This is the case for the GATA4 and GATA5 transcription factors that are silenced by hypermethylation in colorectal and gastric cancer, together with some of their downstream target genes with known antitumor functions (Akiyama et al., 2003). Expression of exogenous GATA5 was able to restore expression of some of the target genes (Akiyama et al., 2003). Another example is HIC1 (Hypermethylated in Cancer 1), a transcriptional repressor frequently mutated in a variety of solid tumors and haematological malignancies (Zheng et al., 2012) (**Introductory Table 1**).

iii. Non-coding RNAs (ncRNAs)

Finally, aberrant hypermethylation can also contribute to the silencing of ncRNA. Micro RNAs (miRNAs) with a potential tumor suppressor function are frequently methylated in cancer cells, leading to an aberrant upregulation of their target genes. In the report of Saito et al, the authors showed that 17 out of 313 human miRNAs were upregulated after treatment with DNA methylation and histone acetylation inhibitor factors in bladder cancer cells. Especially, miR-127 induction leads to the downregulation of the proto-oncogene *BCL6* in several cancer types (Saito et al., 2006). Other examples of miRNAs downregulated by DNA methylation in cancer can be found in **Introductory Table 1**.

In the future, new advances in DNA methylation and transcriptome analysis at genome scale will help to identify all the genes and ncRNAs that are repressed by DNA methylation in various cancer types. This will allow to better understand the differences among tumor subtypes and to create tools to predict, diagnose and select the most convenient therapies for each patient.

Function	Gene	Alteration and cancer type
DNA methyltransferases	*DNMT1*	* Mutational inactivation CRC (Kanai et al., 2003) * Overexpression: CRC (Marzo et al., 1999), prostate (Patra et al., 2001), hepatocarcinome (Saito et al., 2003), breast (Agoston et al., 2005; Al-Kharashi et al., 2017), ovary (Wu et al., 2007), pancreas (Peng et al., 2006), oesophagus (Zhao et al., 2011)
	DNMT3A	*Mutation: AML (Desai et al., 2018; Ley et al., 2010; Ribeiro et al., 2012; Yamashita et al., 2010), MDS (Walter et al., 2011), ALL (Neumann et al., 2013), CMML (Jankowska et al., 2011), T-cell lymphomas (Couronné et al., 2012) * Overexpression: liver (Zhao et al., 2010)
	DNMT3B	* Overexpression: breast (Butcher and Rodenhiser, 2007), colon (Ibrahim et al., 2011; Nosho et al., 2009), prostate (Kobayashi et al., 2011) * Deletion: MLL mouse model (Zheng et al., 2016)
5mC hydroxylases	*TET1*	* Translocation in AML (Lorsbach et al., 2003) * Mutations : Rare in AML (Dolnik et al., 2012) and solid tumors such as CRC (Seshagiri et al., 2012), rare mutation in solid tumors, reviewed in (Scourzic et al., 2015) * Downregulated by methylation in multiple types of cancer (Li et al., 2016) * Overexpression in MLL-rearranged leukaemia (Huang et al., 2013)
	TET2	* Mutation or deletion in haematological malignancies : AML, MDS (Delhommeau et al., 2009), CMML (Kosmider et al., 2009), mastocytosis (Tefferi et al., 2009), T-cell lymphome (Quivoron et al., 2011), B-cell lymphomas (Asmar et al., 2013) * Rare mutation in solid tumors, reviewed in (Scourzic et al., 2015)
	TET3	* Mutation rare in solid tumors such as CRC (Seshagiri et al., 2012)
Isocitrate dehydrogenases	*IDH1/2*	* Mutation: chondrosarcomas (Amary et al., 2011), AML (Figueroa et al., 2010; Green and Beer, 2010; Mardis et al., 2009), glioma (Lu et al., 2012; Turcan et al., 2012; Yan et al., 2009), MDS (Kosmider et al., 2010) Prostate (Ghiam et al., 2012), thyroid carcinomas (Hemerly et al., 2010), brain (Balss et al., 2008)
Methyl-binding-protein 1 and 2	*MBD*	Sansom et al., 2007

Introductory Table 2. Proteins implicated in DNA-methylation that are altered in cancer.

2.1.4 DNA methylation actors are frequently mutated in cancer

The important role of DNA methylation in tumor development is reinforced by the fact that several actors of the methylation / demethylation pathways are often deregulated or mutated in cancer (**Introductory Table 2**).

i. DNMTs

DNMT1 is upregulated or downregulated in several cancers. An overexpression of DNMTs is related to an augmentation in the aberrant methylation of TSG in a variety of tumors. An upregulation of DNMT1 is observed in colorectal cancers, prostate cancers and hepatocellular carcinomas (Marzo et al., 1999; Patra et al., 2001; Saito et al., 2003). Knockdown of *DNMT1* in oesophageal squamous cell carcinoma (ESCC) or cervix cancer cells resulted in TSG demethylation and re-expression, consequently inhibiting cell proliferation and inducing apoptosis (Zhang et al., 2011c; Zhao et al., 2011). With the exception of one study that reported *DNMT1* mutations in rare cases of colorectal cancer (Kanai et al., 2003), no recurrent mutations in the *DNMT1* gene were found in cancer.

In contrast, *DNMT3A* mutations are frequent in haematological malignancies. *DNMT3A* is mutated in a big proportion (around 25%) of adult cases of AML (Ley et al., 2010). *DNMT3A* mutations are generally acquired in the early development of the tumor, suggesting a causal role for tumoral progression (Shlush et al., 2014). In addition, another study showed that *DNMT3A* mutations, among other mutations, can be detected in blood DNA years before the diagnosis of AML, suggesting again a causal role on the progression of the disease (Desai et al., 2018). DNMT3A presents a mutation hotspot at the Arg882 (R882H), which does not abolish catalytic activity but seems to have dominant negative activity and leads to focal methylation alterations (Emperle et al., 2018; Lu et al., 2016; Spencer et al., 2017). *DNMT3A* mutation is a poor-prognosis marker in AML (Ribeiro et al., 2012). In addition to AML, *DNMT3A* mutations are also frequent (up to 20 to 30% of cases) in other haematological malignancies such as myelodysplastic syndromes (MDS) (Walter et al., 2011), chronic myelomonocitic leukaemias (CMML) (Jankowska et al., 2011), acute lymphoblastic leukaemias (ALL) (Neumann et al., 2013) or T-cell lymphomas (Couronné et al., 2012). In contrast to *DNMT3A*, *DNMT3B* is rarely mutated in haematological cancers, however, its depletion promoted MLL-AF9 leukaemia progression in a mouse model (Zheng et al., 2016). Other reports have shown that all DNMTs are overexpressed in AML and CML, which correlates with aberrant hypermethylation of TSG (Mizuno, 2001). In solid tumors, overexpression of DNMT3A and 3B is more frequent than their inactivation by mutations. DNMT3A upregulation has been reported in liver tumors (Zhao et al., 2010) and upregulation of DNMT3B in breast (Butcher and Rodenhiser, 2007) or colorectal cancer

(Ibrahim et al., 2011; Nosho et al., 2009). Mouse models also support a functional role of DNMT overexpression in cancer progression. For example, the overexpression of *Dnmt3b* enhanced the formation of colon tumors in the model of Apc Min/+ mice (Linhart et al., 2007). An interesting finding from this study is that DNMT3B can methylate the same set of genes in tumoral and non-tumoral tissues, suggesting that aberrant methylation caused by DNMT3B overexpression could initiate gene silencing in normal cells (Linhart et al., 2007).

ii. TETs

Somatic mutations of *TET2* are frequent in haematological malignancies such as AML, MDS (Delhommeau et al., 2009), systemic mastocytosis (Tefferi et al., 2009) and CMML (Kosmider et al., 2009). They are present in 10-25% of cases of AML or MDS, and up to 50% of CMML cases. Mutations can be due to chromosomal arrangements but also to point mutations (Weissmann et al., 2012). There are several evidences that designate *TET2* mutation as an early event in leukaemogenesis: *TET2* mutations are present in pre-leukaemic stem cells (Chan and Majeti, 2013; Itzykson et al., 2013) and their presence can predispose to the development of malignancies correlating with age (Busque et al., 2012; Genovese et al., 2014). This predisposition is supported by mouse models with conditional or total loss of *Tet2* (Moran-Crusio et al., 2011; Quivoron et al., 2011). As *TET2* mutation alone is not always sufficient to drive cancer development, additional mutations are required to trigger the malignant transformation. Thus, *TET2* alteration provides clonal advantage to cells, and these malignant precursors can acquire oncogenic mutations that lead to the tumorigenic process. Mutations in *TET2* lead to alterations in 5hmC patterns (Ko et al., 2010). Somatic mutations of *TET1* or *TET3* in human haematological cancers are rare compared to *TET2*. However, TET1 has been implicated in the initiation of B-cell lymphoma in mouse models (Cimmino et al., 2015) or in combination with *TET2* deletion (Zhao et al., 2015), suggesting a tumor suppressor potential. Other reports point that *TET1* has an oncogenic potential in MLL-rearranged leukaemia (Huang et al., 2013).

Less is known about TET mutations in solid tumors. 5hmC levels are dramatically decreased in breast, liver, lung, pancreas or prostate cancers compared to surrounding tissues. This reduction of 5hmC is associated to the reduction in the expression of three types of TET proteins (Yang et al., 2013). Data show that cells with high rate of proliferation correlates with low levels of 5hmC (Bachman et al., 2014), which is consistent with what is observed in cancer cells.

iii. Other mutations

Other proteins related to the metabolism of DNA methylation are frequently mutated in cancer. For example, Isocitrate Dehydrogenase 1 and 2 (IDH1 and 2), enzymes responsible of the

conversion of isocitrate into α-KG, are frequently mutated in gliomas (Yan et al., 2009), haematological diseases (Figueroa et al., 2010; Green and Beer, 2010; Kosmider et al., 2010; Mardis et al., 2009) and other cancers (Waitkus et al., 2018). The mutations preferentially occur at the arginine 132 of IDH1 and the arginine 140 or 172 of IDH2. These point mutations confer new functions to the enzymes: IDHs are then able to convert α-KG into 2-hydroxyglutarate (2HG), which is a competitive inhibitor of α-KG. The decrease in α-KG impacts on the activity of TET enzymes, impairing the demethylation process and leading to a hypermethylation signature (Figueroa et al., 2010; Losman and Kaelin, 2013). Moreover, Methyl-binding-protein 1 and 2 have been reported to be mutated and increase the risk of lung and breast cancer (Sansom et al., 2007) (**Introductory Table 2**).

2.1.5 Other pathologies associated to DNA methylation alterations

DNA methylation anomalies are present in many diseases other than cancer. In the next paragraphs some examples will be presented.

i. Imprinting related disorders

As explained before, DNA methylation regulates the allele-specific expression of imprinted genes. A deregulation in the gene dosage of imprinted genes can trigger the development of several disorders. This deregulation can be caused by methylation alterations that lead to the loss of imprinting in some ICRs. For example, the Prader-Willi syndrome (PWS) and the Angelman syndrome (AS) are neurogenetic disorders caused by loss of imprinting in 15q11-q13, which can be attributed to deletions in the ICR or, in a significant proportion of patients, to abnormal DNA methylation of the ICR (Buiting et al., 2003). The Beckwith-Wiedmann (Reik et al., 1995) and the Silver-Russel syndromes (Abu-Amero et al., 2010) affect the intrauterine growth and are caused by epimutations at the ICR on 11p15.5 in the *H19/IGF2* locus for a significant proportion of patients. Interestingly, gain of methylation in patients is associated with transmission of deletions or single nucleotide variants at CTCF and OCT4 binding sites in the ICR resulting in hypermethylation (Abi Habib et al., 2014; Sparago et al., 2004), which illustrates the complex interplay between genetics and epigenetics in the disease. Another disease caused by imprinting defects is transient neonatal diabetes mellitus, presenting genetic defects in the imprinted *ZAC/HYMAI* locus in 6q24 or methylation aberrations at the CGI overlapping exon 1 of *ZAC/HYMAI* (Temple and Shield, 2002).

ii. Neurological disorders

Among neurological disorders, the Fragile X syndrome is caused by the aberrant expansion and methylation of the CGG repeat on the 5' UTR and adjacent promoter region of the *FMR1* gene. The

gene is then silenced in patients with this mental retardation disorder (Oberlé et al., 1991; Oostra and Willemsen, 2009). Methyl-CpG binding protein 2 (MeCP2) regulates neuronal gene expression by binding to methylated DNA sequences (CpG or CH context, H being A, T or C). Mutations in this protein are frequents in patients with Rett syndrome (Amir et al., 1999; Chen et al., 2015b).

DNMT1 was found to be mutated in patients suffering from hereditary sensory and autonomic neuropathies with dementia and hearing loss (HSANIE) and autosomal dominant cerebellar ataxia with deafness and narcolepsy (ADCA-DN) (Klein et al., 2011; Winkelmann et al., 2012). Both syndromes have heterozygous mutations in the RFTS domain responsible for the recruitment of DNMT1 to chromatin after DNA replication. These mutations affect the recruitment of DNMT1 to chromatin but also the interaction with UHRF1 (Smets et al., 2017), leading to local changes in DNA methylation in cells of patients (Sun et al., 2014).

Changes in DNA methylation might also play a role, but not necessarily a causal one, in other neurological disorders such as Alzheimer's disease (AD), the most common form of dementia. Immunoassays showed that AD patients present reduced methylation levels in the temporal neocortex neuronal nuclei (Mastroeni et al., 2009) and hippocampus (Chouliaras et al., 2013). However, other studies based on bisulfite sequencing do not identify dramatic changes in DNA methylation in patients (Bakulski et al., 2012; Rao et al., 2012). A recent study by the group of Manel Esteller compared methylomes from patients with different neurodegenerative diseases: AD, dementia with Lewy bodies, Parkinson's disease and Alzheimer-like neurodegenerative profile associated with Down's syndrome. Using two different approaches, they found 709 or 1545 DMRs that were common among the 4 diseases and thus represent a common set of aberrant CpG methylation in neurodegenerative diseases. Moreover, around 16% of the genes containing a DMR presented changes in expression (Sanchez-Mut et al., 2016). The shared aberrant methylation pattern could be useful to identify biomarkers for detection or therapy.

iii. Immune disorders

The *DNMT3B* gene is homo or heterozygously mutated in 50% of patients with Immunodeficiency, centromere instability, facial anomalies (ICF) syndrome (Ehrlich, 2003; Hansen et al., 1999; Xu et al., 1999). ICF is an autosomal recessive disease, characterised by highly decreased levels of immunoglobulins in serum and DNA rearrangements in centrosomal pericentromeric regions of chromosomes 1, 9 and 16 associated with DNA hypomethylation. DNMT3B deficiency impairs lymphocyte maturation or activation (Ehrlich, 2003). Patients suffer from bacterial and viral infections and die frequently at a young age. The deregulation of gene expression in ICF patients is due mostly to changes on the Polycomb-dependent histone landscape, suggesting that DNMT3B is

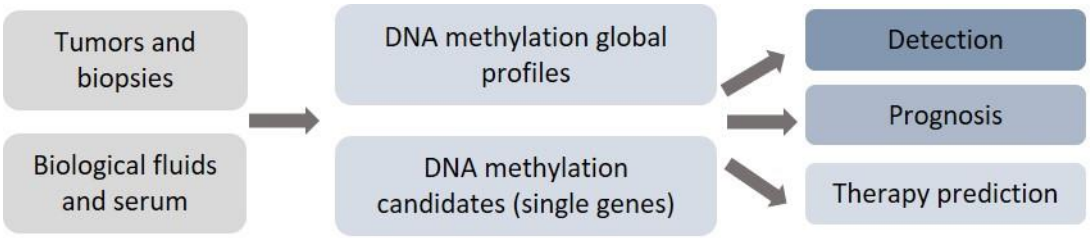
required for the deposition of histone repressive marks like H3K27me3 (Jin et al., 2008). Genome-wide studies showed that hypomethylation in DNMT3B mutant ICF patients occurs mainly in CpG islands of germline genes, imprinted genes and X-linked genes in females (Velasco et al., 2018). Hypomethylation also occurs in intragenic regions and it is linked to altered transcription regulation patterns and altered splicing events, affecting genes with biological functions important to the development of the disease (Gatto et al., 2017). Other types of ICF are linked with mutations of *ZBTB24* (de Greef et al., 2011), *CDCA7* or *HELLS* (Thijssen et al., 2015). In all cases, hypomethylation of the juxtacentrometic chromatin is a hallmark of the disease.

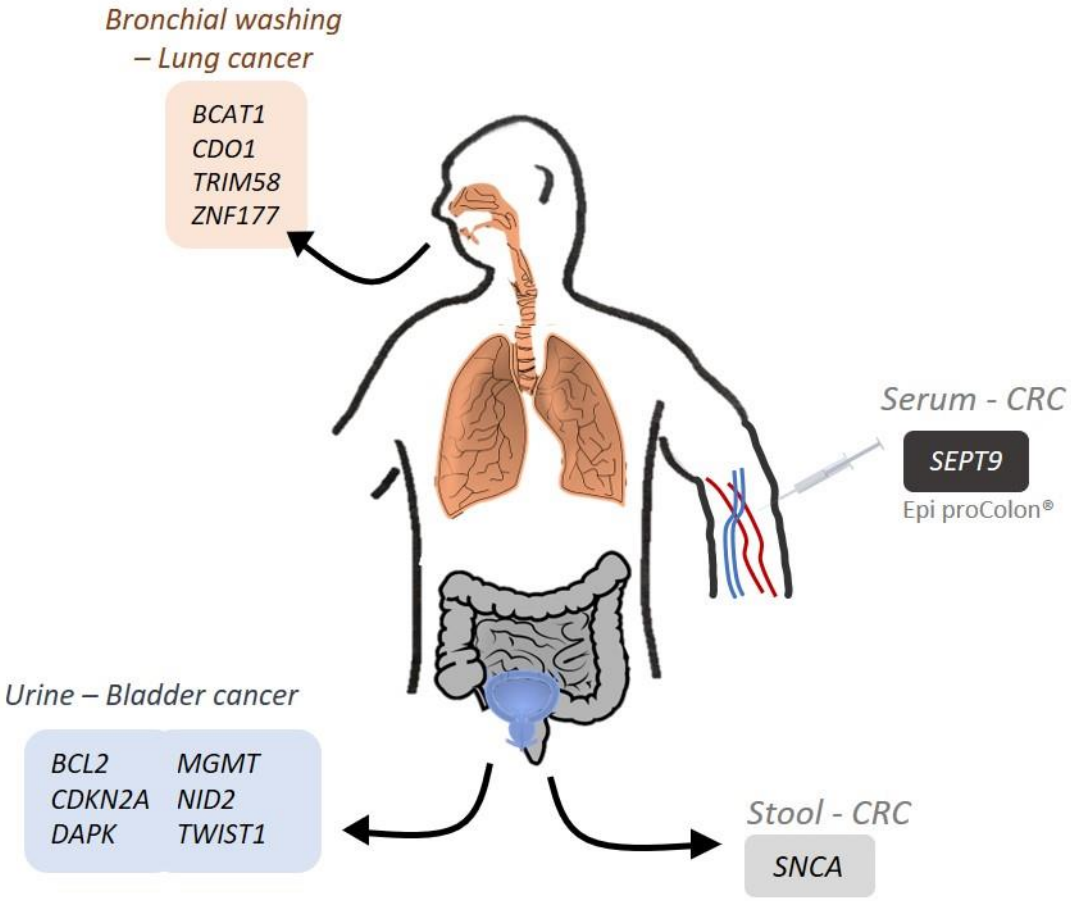
Differential methylation among the MHC (Major Histocompatibility Complex) region potentially increases the genetic risk to develop rheumatoid arthritis (RA) (Liu et al., 2013). In an epigenome-wide association study (EWAS), DNA methylation levels in ten genes were found to be significantly different in B cells from patients compared to healthy individuals (Julià et al., 2017). Interestingly, differential methylation in nine out of those ten genes was also found in systemic lupus erythematosus (SLE) (Julià et al., 2017). These results highlight some genes and shared DNA methylation patterns could be a common driver for the development of these diseases.

Recently, it has been shown that expression levels of the principal actors of DNA methylation (DNMTs and TETs proteins) are deregulated in the hippocampus of multiple sclerosis (MS) patients correlating with demyelination (Chomyk et al., 2017). The authors also found differentially methylated positions (DMPs). If these DMPs were present near a gene TSS, their methylation status inversely correlated with expression. These gene candidates are expressed mostly in neurons and astrocytes (Chomyk et al., 2017), suggesting DNA methylation as a mechanism contributing to altered gene expression in MS patients.

2.2 DNA methylation as a biomarker in cancer clinics

Screening of cancer biomarkers is important to identify high risk populations and to detect cancer at early stages. Late detections of cancers are frequent and imply that the tumors have progressed through advanced stages without the possibility of performing surgical operations or accurate therapies. This late diagnosis is usually linked to poor survival rate. Moreover, current detection methods for cancer present some limitations. For example, mammography cannot detect small tumors in breast and colonoscopy is an invasive and time-consuming method to detect colon cancer. Moreover, histological evaluations require biopsies and their diagnostic depends on the quality of the tissue and the interpretation of the technician. It is also important to take into consideration the heterogenicity of tumors. For example, prostate biopsies frequently underestimate the cancer aggressiveness because they might be not representatives of the actual tumor core





Introductory Figure 10. Epigenetic biomarkers in oncology based on DNA promoter methylation-status.

DNA-methylation screenings have identified DNA methylation biomarkers that can be used in the detection, prognosis and prediction to therapy response. As cancer cells shed DNA into biological fluids near the tumor, non-invasive methods are being developed. These markers are tumor-specific. For example, an increase in the methylation level of the promoter of the genes *BCAT1*, *CDO1*, *TRIM58* and *ZNF177* present in bronchial washing is linked to lung cancer. Until now, only one epi-biomarker has been approved by the FDA for its use in clinics: *SEPT9* promoter methylation in serum is related to colorectal cancer (Epi proColon®). Further research will allow new biomarkers based on methylation to reach the clinical stage.

(Jatkoe et al., 2015). In the next paragraphs, we will discuss recent advances in the area of non-invasive methods to detect cancer based on DNA methylation.

2.2.1 Non-invasive detection techniques for detection of DNA methylation

Epigenetic changes such as promoter hypermethylation of tumor suppressor genes occur early and progressively during tumor initiation and disease evolution, which makes it a good marker to analyse the progression and stage of cancer. For example, changes in methylation can predict the likelihood of metastasis and the recurrence in patients with high risk of relapse after surgery. Thus, determining the DNA methylation pattern in cancer samples provides a new biomarker for early detection, identification of the cancer type, progression and prediction of therapy response (**Introductory Figure 10**).

Whole-genome screening is often used to identify biomarkers comparing tumoral and non-tumoral tissues, whereas locus-specific assays (quantitative methylation-specific PCR, MethyLight assay, pyrosequencing…) are used to study a candidate or a panel of candidates in a specific type of cancer. The best control in the identification of novel biomarkers would be a biopsy of the same tissue that is not affected by the tumor, but sometimes the authors compare samples against healthy individuals.

Individual tumors shed cancer cells and cancer DNA into the blood or closely related body fluids like sputum, saliva, urine or stools. These biological fluids represent a minimally invasive source for biomarkers, which overcome the limitations caused by tumor heterogeneity and the invasive nature of tumor biopsies. Circulating tumor (ctDNA) or cell-free DNA (cfDNA) can be used to study biomarkers. cfDNA is composed by DNA from normal apoptotic cells but also DNA coming from tumors. ctDNA can be distinguished because of the presence of genetic or epigenetic alterations, even if it represents less than 1% of total cfDNA (Diaz and Bardelli, 2014; Diehl et al., 2008; Tanić and Beck, 2017). Urine can be used to detect methylation biomarkers of bladder and prostate cancers (Daniūnaitė et al., 2011; Jatkoe et al., 2015; Renard et al., 2010), sputum or saliva to detect methylation biomarkers of lung and oral cancers (Diaz-Lagares et al., 2016; Miglio et al., 2015; Su et al., 2016), stool to detect methylation biomarkers of colorectal or pancreatic cancers (Li et al., 2015b), and blood to detect methylation biomarkers of breast, colon and prostate cancers (Brait et al., 2017; Liu et al., 2015; Shan et al., 2016; Shirahata and Hibi, 2014) (**Introductory Figure 10**). The principal limitations of these techniques are the available quantity of ctDNA and the background noise. Some examples will be detailed in the next paragraphs and a non-exhaustive list of candidate genes can be found in **Introductory Table 3**.

Methylated gene	Detection/prognosis	DNA marker source	Reference
Prostate			
ERα	Detection	Serum	Brait et al., 2017
Erβ	Detection	Serum	Brait et al., 2017
MCAM	Detection	Serum	Brait et al., 2017
RASSF1	Detection	Urine	Daniūnaitė et al., 2011
RARB	Detection	Urine	Daniūnaitė et al., 2011
GSTP1	Detection / prognosis	Urine / plasme	Jatkoe et al., 2015; Mahon et al., 2014; Wu et al., 2011
APC	Prognosis	Urine	Jatkoe et al., 2015
PCDH10	Prognosis	Serum	Deng et al., 2016
Colorectal			
FBN1	Detection	Stool	Li et al., 2015
NGFR	Detection	Plasma	Lofton-Day et al., 2008
SNCA	Detection	Stool	Li et al., 2015
TMEFF2	Detection	Plasma	Lofton-Day et al., 2008
VIM	Detection	Serum	Shirahata and Hibi, 2014; Shirahata et al., 2010
SEPT9 - Epi proColon® *	Detection / prognosis	Serum	Lofton-Day et al., 2008; Tham et al., 2014
NDRG4, BMP3 - Cologuard® *	Detection	Stool	Pickhardt, 2016
TAC1	Prognosis	Serum	Tham et al., 2014
Bladder			
BCL2	Detection	Urine sediment	Friedrich et al., 2004
CDKN2A	Detection	Urine sediment	Hoque et al., 2006
DAPK	Detection	Urine sediment	Friedrich et al., 2004
MGMT	Detection	Urine sediment	Hoque et al., 2006
NID2	Detection	Urine	Renard et al., 2010
TWIST1	Detection	Urine	Renard et al., 2010
Breast			
FHIT	Detection	Serum	Liu et al., 2015
GSTP1	Detection	Serum	Yamamoto et al., 2012
hMLH1	Detection	Serum	Shan et al., 2016
HOXD13	Detection	Serum	Shan et al., 2016
PCDHGB7	Detection	Serum	Shan et al., 2016
P16	Detection	Serum	Shan et al., 2016
RASSF1α	Detection	Serum	Shan et al., 2016
SFN	Detection	Serum	Shan et al., 2016
Lung			
3OST2	Detection NSCLC	Sputum	Su et al., 2016
BCAT1	Detection	Bronchial washing	Diaz-Lagares et al., 2016
CDO1	Detection	Bronchial washing	Diaz-Lagares et al., 2016
MGMT	Detection	Bronchial washing and sputum	Miglio et al., 2015
MIR-31	Detection NSCLC	Sputum	Su et al., 2016
MIR-210	Detection NSCLC	Sputum	Su et al., 2016
RASSF1A	Detection NSCLC	Sputum	Su et al., 2016
TMEFF2	Detection NSCLC	Serum	Lee et al., 2012
TRIM58	Detection	Bronchial washing	Diaz-Lagares et al., 2016
ZNF177	Detection	Bronchial washing	Diaz-Lagares et al., 2016
SHOX2, PTGER4 - Epi ProLung®	Detection	Serum	Weiss et al., 2017
Head and neck			
CDKN2A	Detection	Saliva	Rosas et al., 2001
DAPK1	Detection	Saliva	Rosas et al., 2001
HOXA9	Detection	Saliva	Guerrero-Preston et al., 2011
MGMT	Detection	Saliva	Rosas et al., 2001
NID2	Detection	Saliva	Guerrero-Preston et al., 2011

Introductory Table 3. DNA methylation-based biomarkers for detection and prognosis of cancer in biofluids. (*) FDA approved.

2.2.2 DNA methylation profiling for cancer detection

One main application of DNA methylation biomarkers in body fluids is cancer detection. Serum prostate-specific antigen (PSA) is frequently upregulated in prostate cancer (PCa) but its detection sometimes leads to unneeded biopsies because of lack of sensitivity and specificity. Combination of this biomarker with promoter TSG methylation of the genes *MCAM*, *ERα* and *ERβ* present in serum increases the sensitivity and specificity of the early PCa diagnosis (Brait et al., 2017). *GSTP1* is frequently methylated in prostate tumors and could serve as a biomarker for early detection (Van Neste et al., 2012). In body fluids, it could be used in complement of PSA screening (Wu et al., 2011b) or other genes hypermethylation like *APC* (Jatkoe et al., 2015).

Carcinoembryonic antigen (CEA) and CA19-19 are used as biomarkers of colorectal cancer in clinics: their levels are usually elevated in patients with this pathology and they are good markers of prognosis and recurrence after treatment, but their accuracy is sometimes limited (Vukobrat-Bijedic et al., 2013). Methylation-based biomarkers have been studied to improve detection rates and several kits have been approved by the US Food and Drug Administration (FDA) for the detection of colorectal cancer. Epi proColon® is a blood-based kit studying the methylation status of Septin-9 (*SEPT9*) (Lamb and Dhillon, 2017). Cologuard® is a stool-based studying the methylation status of *NDRG4* and *BMP3* genes among other factors (Pickhardt, 2016).

Lung cancer is the leading cause of cancer-related death worldwide, in part due to its late detection. New biomarkers detectable in non-invasive techniques are required. It has been shown that a panel of 4 genes (*BCAT1*, *CDO1*, *TRIM58* and *ZNF177*) form an epigenetic signature effectively found in bronchial fluids (Diaz and Bardelli, 2014). The advantage of such a multi-target approach is to overcome the limits of sensitivity of individual gene candidates. Epi proLung® is a blood-based kit approved by the FDA for the detection of lung cancer using the methylation biomarkers *SHOX2* and *PTGER4* (Weiss et al., 2017). Similarly, a multigene panel analysing the methylation status in serum of *HOXD13*, *SFN*, *RASSF1α*, *P16*, *PCDHGB7* and *hMLH1* could be useful to detect breast cancer (Shan et al., 2016). As 15-20% of breast cancer cases are triple negative, meaning that they do not present genetic mutations in key genes (oestrogen receptor, progesterone receptor and Erb-B2 receptor tyrosine kinase 2), it is necessary to find good molecular candidates to perform non-invasive tests. Further research is needed to adapt known biomarkers to body fluid-based test that allow a quick detection and monitoring of cancer.

2.2.3 DNA methylation profiling for cancer prognosis

Beyond its power for detection and diagnosis, DNA methylation patterns can be used to predict prognosis, tumor recurrence and survival rates. Methylation patterns can help to discern

more aggressive from less aggressive tumors and then choose the adequate treatments (Hao et al., 2017).

Concerning relapse, studying patients that had experienced a Non-small cell lung cancer (NSCLC) tumor recurrence within 40 months after surgery and patients without recurrence, DNA methylation of a panel of four genes (*CDH13*, *RASSF1*, *CDKN2A* and *APC*) was linked with early tumor recurrence (Brock et al., 2008). Methylation-based biomarkers like *TAC1* and *SEPT9* are interesting candidates to monitor the recurrence of colorectal cancer post operation (Tham et al., 2014). DNA methylation of TSGs promoters has also been associated with poor-prognosis in melanoma (Bustos et al., 2018), ovary cancers (Chang et al., 2017) or breast cancers (Mathe et al., 2016; Ward et al., 2016). Moreover, methylation patterns can be used to classify tumoral subtype for example in colorectal cancers (Hinoue et al., 2012), breast cancers (Fang et al., 2011) or childhood ALL (Milani et al., 2010), where it can also predict relapse in patients.

Metastasis detection is crucial to prognosis. DNA methylation aberrations in genes linked to cell adhesion are common in metastatic stages of cancer. For example, E-cadherin or cadherin-11 promoter hypermethylation is common through metastasis progression (Carmona et al., 2012; Graff et al., 1995). DNA methylation signatures can also be helpful to differentiate metastases from primary cancers, such as colon cancer metastases to the liver or lung (Hao et al., 2017).

2.2.4 DNA methylation profiling for therapy response

Epigenetic profiles are also linked to drug response sensitivity. Some markers can predict the response to chemotherapeutic agents. For example, *MGMT* silencing by promoter hypermethylation in glioma or glioblastoma is linked to an increased sensitivity to alkylating agents-based therapies like temozolamide or carmustine. *MGMT* removes aberrant alkyl groups from DNA, so its presence in cells would reverse the effect of the alkylating agents (Esteller et al., 2000b; Hegi et al., 2005).

Among other examples, *APAF1* negative melanomas are chemoresistant and apparently not capable of undergoing apoptosis (Soengas et al., 2001), *SLFN11* methylation decreases colorectal cancer cells sensitivity to cisplatin (He et al., 2017) and *IGFBP-3* does the same in NSCLC (Caceres et al., 2010).

New research should be carried out to identify methylation-based biomarkers highly specific, sensitive and adaptable to routine clinical tests. The main objective is the establishment of reliable biomarker panels allowing the identification and detection of cancer comparable to classical tests. If these biomarkers are not sensitive enough to detect the presence of the disease independently, they could be implemented in addition to currently studied molecular signatures. Such tests based on

blood or body fluids will be cost-effective, less time-consuming, less painful and more advantageous to patients.

2.3 DNA methylation as a therapeutic target for cancer

There are different methods used in cancer therapy and they can be used in combination to improve treatment response and increase survival rate. Because of their reversible nature, epigenetic marks such as DNA methylation are also a target of choice in anti-cancer therapies. There are some epigenetic drugs currently used in clinics and in clinical trials. In the next paragraphs, I will present an overview on classical and new cancer therapies, with a focus on DNA methylation therapies.

2.3.1 Classical methods

Classical methods include surgery to remove the tumors, followed by irradiation of the tumors (radiotherapy) and/or treatment by medication (chemotherapy). In some advanced stages, surgery is not an option and it is necessary to find a correct combination of treatments.

In radiotherapy, cells within the tumor environment are targeted. The targeted cells undergo massive DNA damages that slow their growth and eventually kill them. Depending on the type of cancer, the location and the size of the tumor, different types of radiotherapy can be performed. Radiotherapy can be performed individually or in combination with drug treatments (Sharma et al., 2016).

Chemotherapy agents target quickly dividing cells and disrupt the cell-cycle. Some of the most commonly used chemotherapy drugs are alkylating agents like cyclophosphamide or temozolomide (Ralhan and Kaur, 2007). These agents induce damage in the DNA, leading to DNA strand breaks, arrest of cell division and cell death. They can act in a wide range of cancers. One of their main limitations is that cells can be resistant to therapy via proteins such as MGMT (Kaina et al., 2007, *see 2.2.4 DNA methylation profiling for therapy response*). Moreover, as they induce DNA damage, they increase the risk of developing a secondary cancer such as leukaemia (Davies, 2001). Other chemotherapy agents include antimetabolites and DNA topoisomerase inhibitors.

Both radio and chemotherapy present many side effects because of their global effect: normal cells in division are also destroyed during the treatment. Moreover, cancer cells can become resistant to treatments and increase the aggressiveness of the cancer. For these reasons, new targeted cancer therapies have been developed.

Drug	Selectivity	Year of approval	Approval indication	Reference
5'-azacytidine	DNMT1, DNMT3A, DNMT3B, TRDMT1 and RRM2	2004	MDS, AML, CMML	Kaminskas, 2005
5'-aza-2'-deoxycytidine (decitabine)	DNMT1, DNMT3A, DNMT3B	2006	MDS	Kantarjian et al., 2006
SAHA (Vorinostat)	HDAC classes I, II and IV	2006	CTCL	Mann et al., 2007
Romidepsin (FK228)	HDAC class I	2009	CTCL and PTCL	Barbarotta and Hurley, 2015; Coiffier et al., 2012
Ruxolitinib	JAK1/2	2011	Myelofibrosis	Mascarenhas and Hoffman, 2012
Belinostat (PXD101)	HDAC classes I, II and IV	2015	PTCL	Lee et al., 2015
Panobinostat (LBH589)	HDAC classes I, II and IV	2015	Relapsed/refractory multiple myeloma	Raedler, 2016

Introductory Table 4. Epigenetic drugs approved by the FDA.

AML = Acute myeloid leukaemia
CTCL = Cutaneous T-cell lymphoma
MDS = Myelodysplastic syndrome
PTCL = Peripheral T-cell lymphoma

2.3.2 Targeted cancer therapies

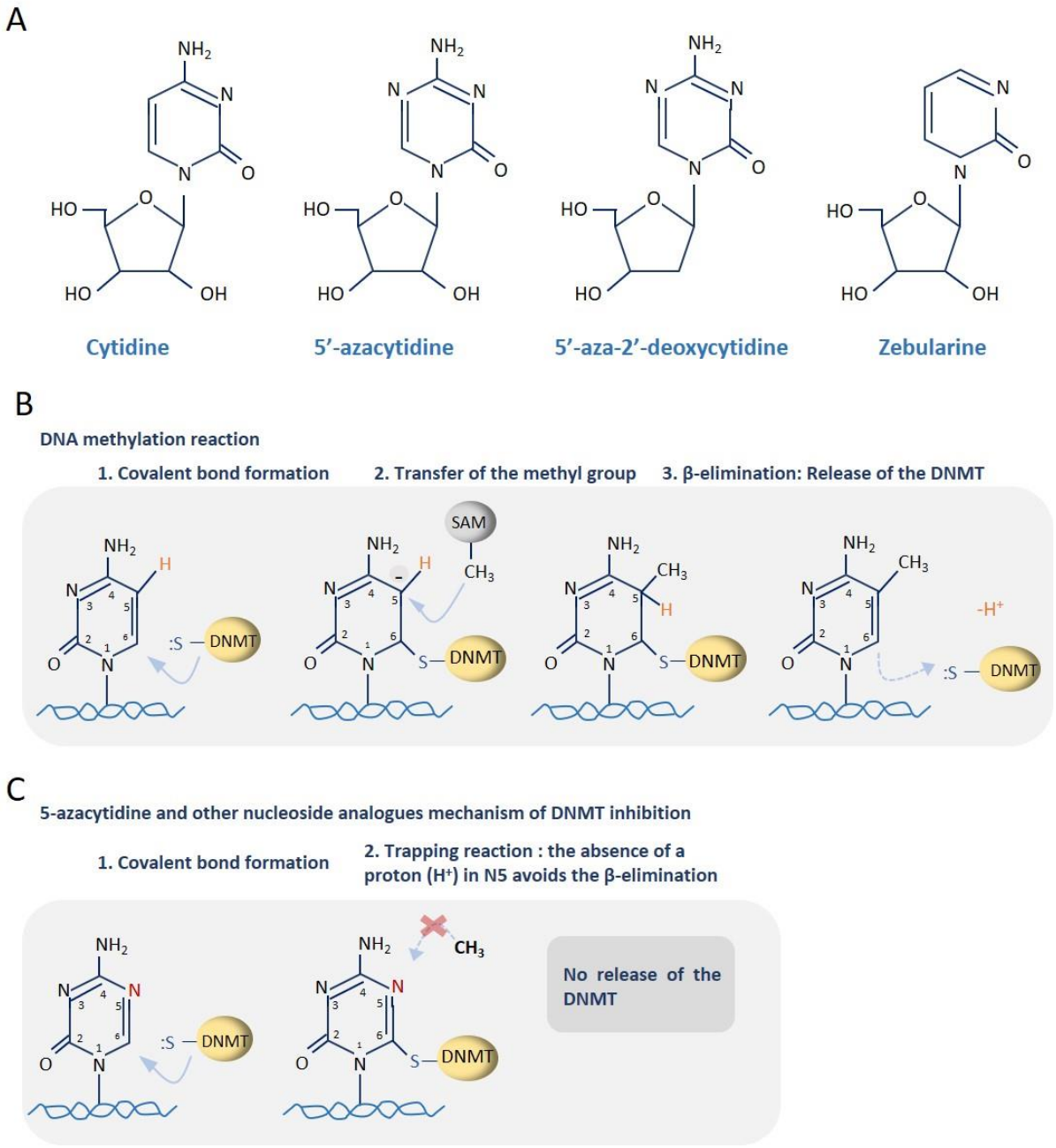
In the last years, targeted cancer therapies have emerged as an alternative to counterpart the cytotoxic side effects of traditional therapies. These therapies include monoclonal antibodies (mAb), small molecule inhibitors and immunotoxins.

Along with traditional treatments, mAbs are the most used therapy. mAbs target extracellular proteins. They can mediate the disruption of the interaction between ligands and receptors essential for cells, block a recognition site, participate in the activation of the immune system against cancer cells or deliver drugs to cancer cells in antibody-drug conjugates (Weiner, 2015). mAbs can also be used in conjugation with a toxin that is delivered to cancer cells and internalised. Toxins are generally from bacteria or plant origin. This antibody-based therapy is known as immunotoxins (Alewine et al., 2015; Weiner, 2015). The first immunotoxin approved by the FDA for use in clinics (in Cutaneous T-cell lymphoma, CTCL) is denileukin diftitox (Duvic and Talpur, 2008). Since then, other immunotoxins have been approved by the FDA and others are in ongoing trials (Alewine et al., 2015; Allahyari et al., 2017). Small molecule inhibitors inactivate proteins such as kinases in molecular pathways that are deregulated by cancer (Singh and Jadhav, 2018; Zhang et al., 2009). One advantage of these molecules compared to mAb is that they can enter in the cells. Even if these techniques are more specific and present lower side effects, they have some limitations. Tumors are heterogeneous and all cells might not express a defined protein. This applies also to differences among individuals. Moreover, mutations in targeted proteins would create a resistance to mAbs recognition. As this kind of therapies are frequently carried after standard treatments, immune system can be weakened.

2.3.3 DNA methylation as a target for cancer therapies

Because of its reversible nature and its role in cancer, DNA methylation is a promising target for therapeutic action. DNA methylation inhibition can lead to the re-expression of TSG promoters and the inhibition of aberrant proliferation. Such strategies targeting DNA methylation could be used in combination with conventional therapies. There are several DNMT inhibitors (DNMTi) classified as nucleoside and non-nucleoside inhibitors.

Before describing in details the use of DNMT inhibitors in cancer therapies, it is to note that other epigenetic drugs are used in cancer treatments. In particular, several histone deacetylases (HDAC) inhibitors have been approved by the FDA (**Introductory Table 4**): Vorinostat, approved for the treatment of CTCL, Romidepsin, for the treatment of CTCL and Peripheral T-cell lymphoma (PTCL), Belinostat, for the treatment of PTCL, and Panobinostat, approved for the treatment of relapsed/refractory multiple myeloma in combination with dexamethasone and bortezomib.



Introductory Figure 11. DNMT inhibition by cytosine analogues.

A. Structures of some of the principal DNMT inhibitors. 5'azacytidine (5aza) and 5'aza-2'-deoxycytidine (5azadC) are the only ones approved by the FDA for the treatment of AML and MDS.
B. Schematic representation of the DNA methylation reaction. The enzyme DNMT forms a complex with the DNA via a covalent union between a sulfhydryl group in the DNMT and the carbon 6 of cytosine ring. The methyl group ($-CH_3$) is transferred from the cofactor SAM and the enzyme is released by β-elimination.
C. Schematic representation of the DNMT inhibition with 5aza. The absence of a proton at the level of N5 position in the 5aza ring avoids the β-elimination and so the enzyme is irreversibly trapped.
Adapted from GNYSZKA et al., 2013; Schermelleh et al., 2005.

i. Nucleoside analogues DNMT inhibitors

5-azacytidine (5aza) and 5'-aza-2-deoxycytidine (5azadC)

5-azacytidine (5aza) and its analogue 5'-aza-2-deoxycytidine (5azadC) were synthetised for the first time in the 60's (Šorm et al., 1964). Both were shown to be active agents for the treatment of AML (Karon et al., 1973; Momparler et al., 1985; Richel et al., 1991; Vogler et al., 1976). At the same time, their inhibitory capacity against DNMTs was reported in several studies (Creusot et al., 1982; Friedman, 1981; Jones and Taylor, 1980; Taylor and Jones, 1979). In 1983 a correlation was found between the inhibition of methylation and the antileukaemic activity of 5azadC in a mouse model (Wilson et al., 1983). Both compounds are formed by a modified cytosine (nitrogen instead of carbon at position 5) and a ribose (5aza) or deoxyribose (5azadC) (**Introductory Figure 11A**). Because of that, 5azadC is incorporated exclusively in DNA whereas 5aza can be incorporated into RNA and DNA. 5aza and 5azadC are the only DNMT inhibitors approved by the FDA up to now. 5aza was approved in 2004 (Kaminskas, 2005) and 5azadC in 2006 (Kantarjian et al., 2006) by the FDA (an later by the European Medicines Agency) for the treatment of myelodysplastic syndrome (MDS), acute myeloid leukaemia (AML), and chronic myelomonocytic leukaemia (CMML) (**Introductory Table 4**).

The mechanism of action of these drugs is schematised in **Introductory Figure 11B**. In normal conditions, the DNMT covalently binds to the carbon 6 of cytosine and transfers the methyl-group (CH_3) from the cofactor SAM to the carbon 5 of cytosine, leading to the formation of 5mC. After the transfer, the enzyme is released by β-elimination. The nucleoside analogues 5aza and 5azadC are phosphorylated and incorporated into DNA during DNA replication. As these molecules have a nitrogen (N) at position 5 of the cytosine instead of a carbon, the methyl transfer does not take place and the β-elimination reaction is not performed. Thus, DNMTs are "trapped" on the DNA, which leads to passive loss of methylation (Schermelleh et al., 2005; Subramaniam et al., 2014) (**Introductory Figure 11C**). Several reports suggest that 5aza can also upregulate gene expression via mechanisms that are independent of DNA demethylation (Seelan et al., 2018).

Combination of drugs in therapy can act synergistically or can help to decrease the side effects of monotherapy. It has been shown that 5aza in combination with chemotherapeutical agents can improve its effect *in vitro* (Füller et al., 2015; Qin et al., 2015; Zhang et al., 2017a) and in patients (Clozel et al., 2013; Juergens et al., 2011).

There are several ongoing clinical trials combining 5aza treatment with other therapies (Coronel et al., 2011; Glasspool et al., 2014; Gore et al., 2017; Li et al., 2017; Sun et al., 2018b) (**Introductory Table 5**). In particular, combination of 5aza with immunotherapeutic approaches leads

Drug	Clinical trial ID	Condition	In combination with …	Trial phase
5'-azacytidine	NCT01935947	Advanced NSCLC	Entinostat (HDACi)	II
	NCT03206021	Children brain/solid tumors		I
	NCT02959437	Advanced Solid Tumors	Pembrolizumab and Epacadostat	I/II
	NCT02811497	Advanced Solid Tumors	Durvalumab	II
	NCT01386346	Resectable Oesophageal Cancer	Oxaliplatin, Epirubicin, Capecitabine	I
	NCT02260440	Chemo-refractory Metastatic Colorectal Cancer	Pembrolizumab (MK-3475)	II
	NCT02546986	Metastatic Non-small Cell Lung Cancer	MK-3475	II
	NCT02530463	MDS	Nivolumab and Ipilimumab	II
5'-aza-2'-deoxycytidine (decitabine)	NCT01627041	AML	Cytarabine, Daunorubicin Hydrochloride	II
	NCT02959164	Pancreatic Cancer and Sarcoma	Gemcitabine	I
	NCT03282825	Advanced breast cancer	Paclitaxel	I
	NCT02159820	Ovarian cancer	Carboplatin-Paclitaxel	II/III
	NCT01729845	Relapsed or Refractory AML or High-Risk MDS	Mitoxantrone Hydrochloride, Etoposide, and Cytarabine	II
5-fluoro-2'-deoxycytidine	NCT00978250	Head and Neck, NSCLC, bladder, breast cancer	Tetrahydrouridine	II
	NCT01534598	Advanced Solid Tumors	Tetrahydrouridine	I
4-thio-2-deoxycytidine	NCT03366116	Advanced Solid Tumors		I
	NCT02423057	Advanced Solid Tumors		I
SGI-110 (guadecitabine)	NCT02920008	AML		III
	NCT02907359	MDS or CMML		III
	NCT02348489	AML		III
	NCT02901899	Ovarian cancer	Pembrolizumab	II
	NCT03308396	Kidney Cancer	Durvalumab	II
	NCT01896856	Metastic colorectal cancer	Irinotecan	I/II
ASTX727 (decitabine and cytidine deaminase inhibitor), oral administration	NCT02103478	MDS		I/II
	NCT03306264	MDS and CMML		III

Introductory Table 5. Available DNMT inhibitors for cancer treatment in ongoing clinical trials (non-exhaustive list). Trial status: Active or recruiting on the 28/07/2018

to better outcomes and clinical trials are being carried out to analyse its efficacy in patients (Mazzone et al., 2017).

5aza and 5azadC present several limitations. First, they lack specificity of action, which induces a global demethylation genome-wide. At high doses, both molecules have cytotoxic effects. Moreover they have mutagenic potential (Jackson-Grusby et al., 1997). Concerning their pharmacokinetic properties, they are unstable in physiological media, they are vulnerable to deamination by cytidine deaminase and have a weak bioavailavility. Thus, research in new DNMT inhibitors (DNMTi) has been developed to overcome these limitations.

Other analogues

Zebularine is a nucleoside analogue that overcomes the lack of stability of 5aza and 5azadC (Zhou et al., 2002) and blocks the action of cytidine deaminase (Jeong et al., 1998) (**Introductory Figure 11A**). It forms a reversible covalent complex with DNA (Champion et al., 2010). Promising results demonstrating anti-tumor properties *in vitro* and in mice suggested that it was a potential good candidate for therapy (Herranz, 2005; Scott et al., 2007; Yoo et al., 2008). Sequential administration of 5aza followed by zebularine prevent the remethylation in demethylated regions, suggesting that a combination of both drugs could achieve sustained gene expression (Cheng et al., 2004; Lemaire et al., 2009). However, even if it was more stable, it could not enter clinical trials because high toxic doses were required (Johnson et al., 2006).

Another nucleoside demethylating agent is 5-fluoro-2'-deoxycytidine (FdCyd). Phase I trials combining it with tetrahydrouridine (cytidine deaminase inhibitor) for the treatment of AML, MDS or solid tumors have been completed (Newman et al., 2015) (NCT01041443, NCT00359606) and phase II trials are ongoing. Furthermore, 4-thio-2-deoxycytidine (TdCyd) is orally bioavailable and has a potential antineoplastic activity. It is currently starting to be tested in phase I clinical trials (NCT03366116, NCT02423057) (O'Sullivan Coyne et al., 2016) (**Introductory Table 5**).

Prodrugs of 5aza and 5azadC were synthesised in an attempt to improve their pharmacokinetics. Some examples are CP-4200, an elaidic acid ester that was shown to improve the therapeutic effect of 5aza in human cancer cells *in vitro* (Brueckner et al., 2010) and SGI-110 (guadecitabine), an oligonucleotide that improved cellular uptake and prevented cytidine deamination (Yoo et al., 2007). In SGI-110, the 5azadC molecule is linked via a phosphodiester bond to deoxyguanosine, which confers a longer half-life to the drug (Griffiths et al., 2013; Yoo et al., 2007). SGI-110 is in clinical trials for haematological diseases, ovarian and hepatocellular cancers (NCT02197676 NCT01696032, NCT01752933) and candidates are being currently recruited for its

phase III clinical trials study for AML, MDS or CMML treatment (NCT02920008, NCT02907359) (**Introductory Table 5**).

ii. Non-Nucleoside DNMT inhibitors

Non-nucleoside DNMT inhibitors have been discovered in screenings of natural compounds or synthetised drugs. These inhibitors do not need to be incorporated into the DNA sequence to perform their inhibitory effect, which means that the demethylation effect is achieved after only one round of cell division. They can be classified in several groups: DNMT binding molecules, oligonucleotides or SAM competitors. Most non-nucleoside compounds have not shown a demethylation effect *in vitro*, but their development could be interesting because they can target DNMTs in a selective manner.

One example of natural molecules is nanaomycin-A, which has an interesting profile because of its preference towards DNMT3B. It inactivates its catalytic domain, promoting the reactivation of silenced TSG (Kuck et al., 2010). In some cases, synthetic drugs used to treat other diseases have been repurposed after showing DNMT inhibitory properties. One example is hydralazine, a known drug used to treat hypertension, which was demonstrated to activate TSGs in cells without causing a global demethylation by reducing DNMT1 and DNMT3A expression (Ruiz-Magaña et al., 2016; Segura-Pacheco et al., 2003). Moreover, this compound is in clinical trials. Phase I trials in combination with valproate in solid tumors (NCT00996060) and phase II trials in combination with chemotherapy agents in cervical cancer (NCT00404326) and solid tumors (NCT00404508) have been completed. In Mexico, its utilisation in cervical cancer in combination with HDACi is authorised. The medicine has the name TRANSKRIP® and this was the first authorised epigenetic combined therapy (Candelaria et al., 2011). Other examples are procainamide and procaine, medicines used in arrythmia that show a demethylation effect in cancer cells. Procainamide is a competitive inhibitor of DNMT1, reducing its affinity for SAM and hemimethylated DNA (Lee et al., 2005; Villar-Garea et al., 2003). RG108 is a molecule found during a *in silico* screening and was shown to demethylate tumor suppressor genes promoters in cancer cells (Brueckner et al., 2005). The combination of procainamide with RG108 increased the cytotoxicity in cancer cells (Halby et al., 2012). High-throughput screening of chemical libraries against the complex DNMT3A-3L or the catalytic domain of DNMT3A allowed the identification of new chemical compounds of the families of flavones, quinones or propiophenones that were able to inhibit DNMTs. Specifically, some of them lead to a 5aza-phenotype in zebrafish (Ceccaldi et al., 2011), reactivated gene expression from a methylated promoter (Ceccaldi et al., 2013) or had cytotoxicity against cancer cells (Erdmann et al., 2015). The

quinoline based agent SGI-1027 is a SAM competitor and also lead to TSGs demethylation and reactivation (Datta et al., 2009; García-Domínguez et al., 2013).

Other agents showing a demethylation effect in cancer cell lines *in vitro* are the natural molecules genistein (Xie et al., 2014), curcumin (Liu et al., 2009) or polyphenol epigallocatechin-3-gallate (EGCG) (Fang et al., 2003).

Among the oligonucleotides able to inhibit DNMTs, MG98, an epi-mRNA targeting the 3'UTR region of DNMT1 showed very promising results in preclinical studies with the inhibition of DNMT1 and the reactivation of TSG, but clinical trials in patients were not efficient (Klisovic et al., 2008; Winquist et al., 2006). When used in combination with other molecules, its efficacy seems to increase (Amato et al., 2012).

An interesting approach is acting on several epigenetic marks at the same time. In this line, the dual targeting of the histone methyltransferase G9A and DNMT1 via the molecule CM-272 inhibits cell proliferation and promotes apoptose *in vitro*. Interestingly, the molecule also shows activity *in* vivo (San José-Enériz et al., 2017).

The list of compounds issued from screenings is very large. In a report that compared some of these non-nucleoside inhibitors with 5aza effect, only 5aza could promote reliable DNA demethylation (Chuang, 2005). As their ability to demethylate DNA is still under study, further research should be carried in the future to optimise these compounds or find the most advantageous combinations that could be applied in clinics.

3 TARGETED EPIGENOME EDITING

As we have seen in the last chapter, epigenetics is involved in many diseases and its remodelling can be considered as a therapeutic target because of its reversibility and impact on gene expression. The epigenetic pharmacological drugs currently in the market have a common disadvantage: they have genome-wide effects that modify not only the target genes of interest but all the genome. Even if some of these drugs are currently used in clinics, the fact that they have global effects is a main limitation.

Over the past few years, gene targeting technologies have emerged as promising tools for biotechnological application in research and therapeutics. These can be used for targeted DNA sequence editing but also as epigenome editing tools that alter gene function without affecting DNA-sequence. The main objective of epigenome editing is to rewrite epigenetic marks such as DNA methylation or histone marks in order to modulate gene expression. This is generally achieved by fusing proteins containing an effector domain to a programmable gene-specific DNA binding domain (DBD).

Epigenome editing can help to validate hypothesis involving the role of epigenetic marks such as DNA methylation in gene expression control and cell phenotype. Moreover, as most diseases have deregulated patterns of DNA methylation, epigenome editing can be developed with a therapeutic aim.

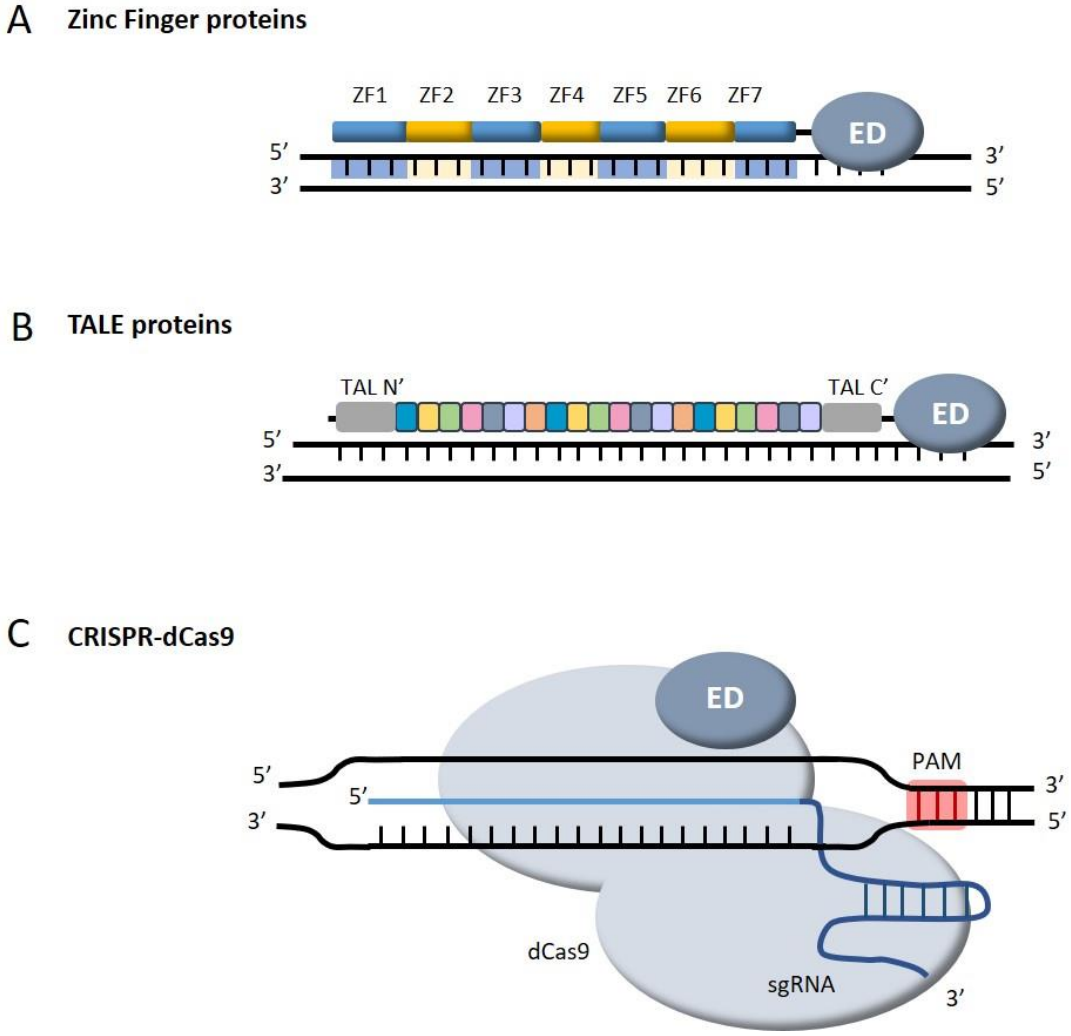
In the next section we will discuss the pros and cons of the different programmable genome editing technologies, the latest achievements in epigenome editing and its limitations.

3.1 Targeted genome editing tools

The most commonly used DNA binding domains for (epi)genome editing are zinc finger proteins (ZFPs), transcription activator-like effectors (TALEs) and the recently discovered Clustered Regulatory Interspaced Palindromic Repeats (CRISPR) system.

3.1.1 Zinc finger proteins

Zinc fingers are found in mammalian transcription factors. For example, they are encoded by 3% of human genes (Klug, 2010). Zinc finger proteins (ZFPs) are composed of modular zinc finger domains. Each finger domain contains 30 amino acids, from which 25 compose the module and the five others act as linker to the next module (Klug, 2010; Miller et al., 1985). Domains are composed of 2 β-sheets and one α-helix and their structure is stabilised by coordination of a Zn ion with two



Introductory Figure 12. Programmable DNA binding domains used in genome and epigenome editing.

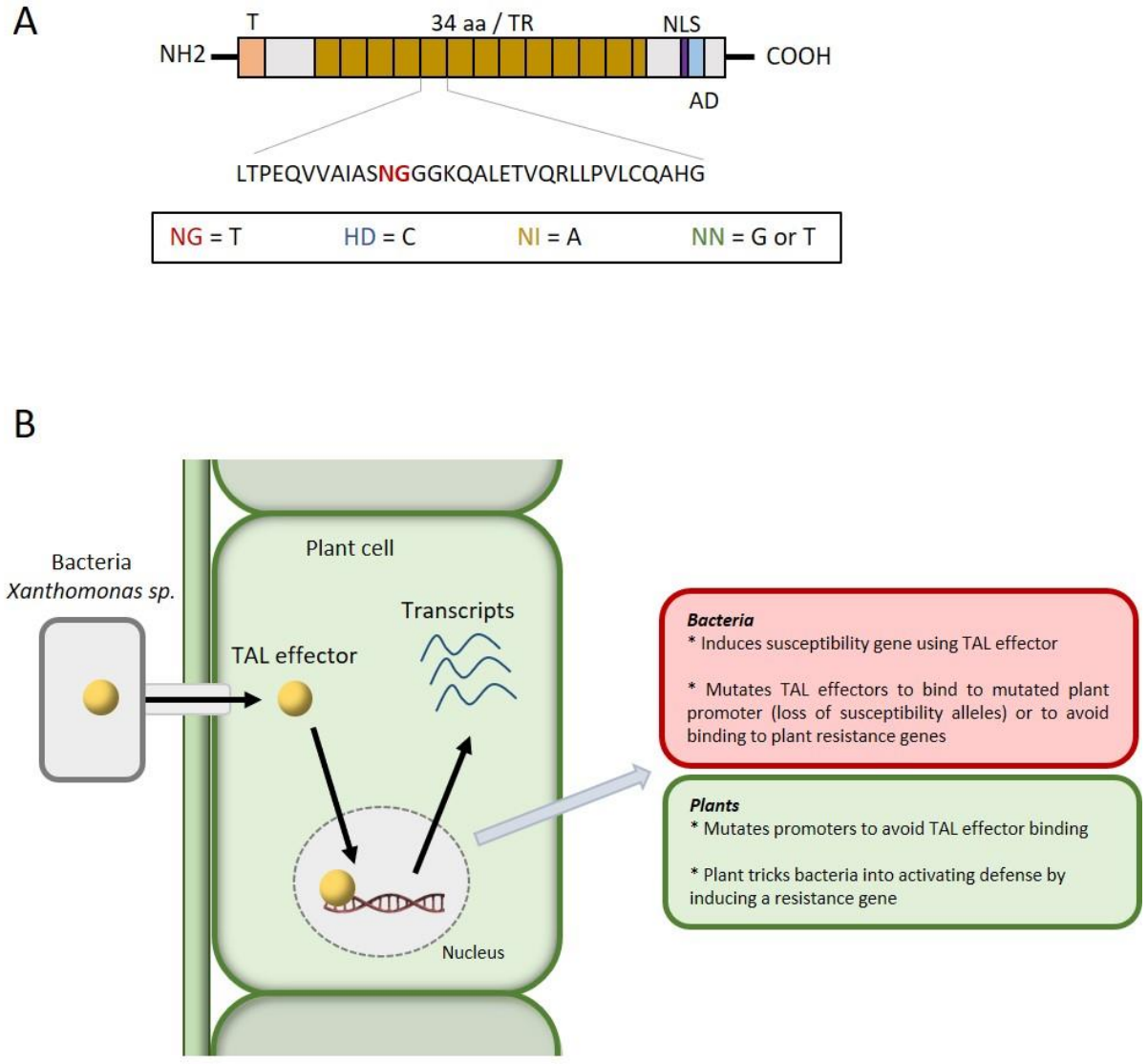
A. Zinc finger proteins are composed of zinc finger modules. Each domain recognises 3 nt in the DNA. Combinations of several modules can target longer DNA sequences.
B. TALE proteins are also composed of modules. In this case, each module recognises 1 nt according to a code.
C. CRISPR-dCas9 (catalytically dead Cas9) system is guided by a guide RNA of 20 nt recognising its complementary sequence in the genome. Recognition by the Cas9 protein requires a PAM sequence just downstream the target sequence. SpCas9 recognises the canonical PAM 5′ NGG 3′.
ED: Effector domain

histidines and two cysteines and a hydrophobic core. The α-helix harbours the recognition domain: 3 amino acids recognise 3 base pairs (bp) in the major groove of DNA (Pavletich and Pabo, 1991) (**Introductory Figure 12A**). Later, it was found that they also interact with a nucleotide on the neighbouring sequence (Fairall et al., 1993; Klug, 2010). Zinc finger domains can be associated in different tandem combinations to form larger modules that recognise larger sequences of DNA. For example, 6 zinc finger domains form a ZFP able to recognise 18 bp, which is considered the minimum length to be unique in the genome (Gersbach et al., 2014). There are libraries of artificially constructed fingers to facilitate the generation of ZFPs by modular assembly (Bhakta and Segal, 2010). ZFPs were the first DBD used in gene expression modulation when Choo et al. showed that a ZFP repressed the expression of an endogenous oncogene by transcriptional blockage (Choo et al., 1994).

ZFPs present some interesting characteristics. First, they are smaller than the other gene targeting tools (TALE, CRISPR-dCas9), which gives them an advantage for targeting compacted chromatin regions. Second, they are less immunogenic because they exist as endogenous components in mammalian cells, which is an asset to their use in clinics (Cano-Rodriguez and Rots, 2016). In fact, clinical trials with ZFP-nucleases are ongoing. A phase II trial for the treatment of HIV infections has been completed (NCT01252641). Other trials are ongoing (mostly in phase I) for the treatment of pathologies like transfusion-dependent β-thalassemia (NCT03432364), Mucopolysaccharidosis I (NCT02702115) or Human Papillomavirus-Related Malignant Neoplasm (NCT02800369). This raises the possibility that epigenome editing tools based on ZFPs will be used in clinics in the future. On the other side, ZFPs construction is laborious and expensive (Cano-Rodriguez and Rots, 2016).

3.1.2 TALE proteins

Xanthomonas sp are bacterial plant pathogens that use a secretion system to translocate virulence factors into the nucleus of the plants. These effectors modulate disease resistance genes: they regulate their expression by mimicking eukaryotic transcription factors to enhance bacterial proliferation, virulence and colonisation (**Introductory Figure 12B, 13A and B**). The AvrBs3/PthA or TAL (transcription activator-like) effectors (TALE) is the most common family of effectors in *Xanthomonas sp* (Boch and Bonas, 2010; Boch et al., 2009; Szurek et al., 2002). TALEs induce different reactions in the host plant, for example, the *pthA* gene from *X. citrii* causes cancer-like symptoms in citrus, such as hyperplasia or hypertrophy, leading to the rupture of epidermal layers and facilitating bacterial release (Duan et al., 1999; Sun et al., 2018a). Another example is the *AvrXa* gene in *X. oryzae*, which promotes bacterial growth in rice (Bai et al., 2000; Yang and White, 2004). Interestingly, some plants have developed resistance mechanisms by copying the TALE binding sites



Introductory Figure 13. TALE system in *Xanthomonas sp*.

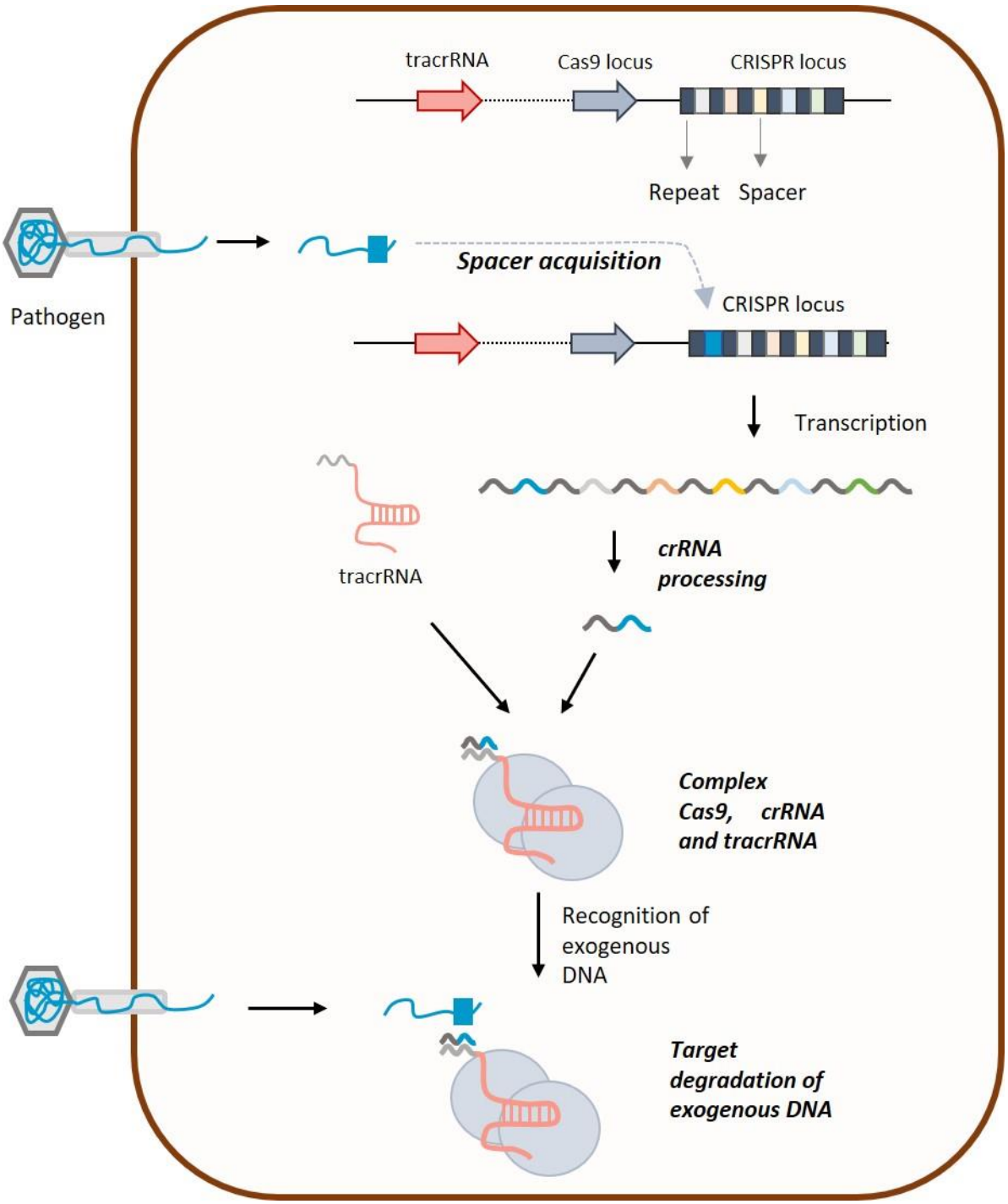
A. Schematic representation of the functional domains of a TAL effector from *Xanthomonas sp* bacteria. Each yellow square is a tandem repeat (TR) domain of 34 amino acids. The number of repeats can go from 1,5 to 33,5. The amino acids 12 and 13 of each TR form the RVD (repeat variable diresidue), responsible for the hybridisation to a specific nucleotide. The most common RVD and their targeted nucleotide are indicated. T, Translocation signal; NLS, Nuclear Localisation Signal; AD, Activation domain.
B. Schematic representation of the pathogenic mechanism of *Xanthomonas sp* in plants. Bacteria translocate the virulence factors into the nucleus of plant cells. There, they modulate the expression of disease resistance genes, with the aim of inducing local plant cell death and promote bacterial proliferation, virulence and colonisation. Adapted from Boch et al., 2009.

in resistance genes. Thus, the gene expression induced by a TALE can confer resistance to the infection (Gu et al., 2005) (**Introductory Figure 13B**). Similar TALE proteins have been found in *Ralstonia solanacerum* (Cunnac et al.; Li et al., 2013a) and *Burkholderia rhizoxinica* (de Lange et al., 2014). TALEs contain classic eukaryotic protein domains such as the nuclear localisation signal (NLS), required for translocation into the nucleus, or the acidic activation domain (AD) in the C-terminal end essential for transcriptional activation. Moreover, they also contain a translocation signal in the N-terminal end (**Introductory Figure 13A**).

TALEs are composed by tandem repeats (TR) of around 34 amino acids. Each repeat is responsible for the recognition of one nucleotide, and the combination of repeats in TALEs creates specificity to bind to a DNA sequence. The amino acids 12 and 13 in each repeat are variable and responsible of the specific recognition and binding to a nucleotide. These two amino acids are called repeat variable diresidues (RVD). The most common RVD in nature are NG, HD, NI and NN, which target T, C, A and G/A respectively (Boch et al., 2009) (**Introductory Figure 13A**). Another characteristic of the target sequence is that it must contain a thymine at the 5' end. The sequence just upstream of the first canonical tandem repeat in the TALE protein is similar to the sequence found on the repeats, so this region might bind to the conserved T. Mutation of the T reduces TALE binding and activity (Boch et al., 2009; Mak et al., 2012). Since the discovery of the code linking the RVD amino acids to the nucleotide sequence, it has become possible to engineer artificial TALEs by assembling different domains sequentially to obtain a DBD capable of recognising a unique sequence in the genome. 20,5 TR are frequently assembled to recognise sequences of 21 nt (22 nt if we consider the 5' T). Recently it has been shown that TALEs use a rotationally decoupled mechanism to search its target sequence in the DNA (Cuculis et al., 2016). Similarly to ZFP, TALE domains can be fused to nucleases or to effector domains and used in genome and epigenome editing.

Notably, the ability of TALEs to recognise 5mC has been studied. 5mC seems to prevent TALEs binding with the common cytidine RVD (HD). As the glycine from the NG pair accommodates the methyl-group of thymine, it was suggested and proved that the RVD NG was able to bind 5mC (Deng et al., 2012a). The RVD N* (where * is a deletion) is also able to bind 5mC (Valton et al., 2012). Other less common RVDs show specificity for 5mC and 5hmC, allowing discrimination among cytosine modification states (Zhang et al., 2017b).

An advantage of TALEs compared to ZFPs is that one module recognises one single base, giving more flexibility in sequence design and increased specificity. TALEs coupled to nucleases (TALENs) have been shown to be highly specific and have less off-target effects than ZFPs (Mussolino et al., 2011). However high DNA binding capacity of TALENs can also cause non-specific binding and off-



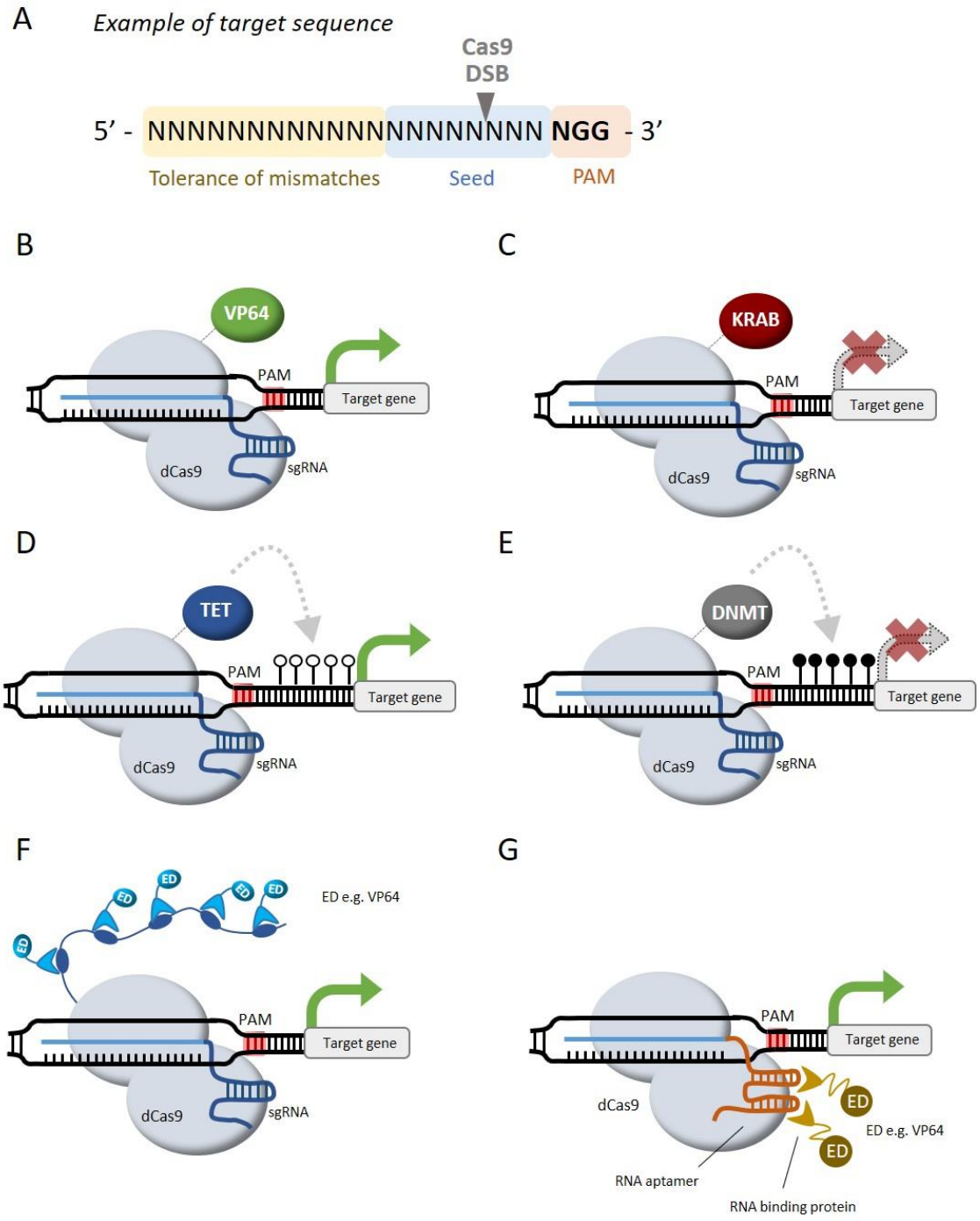
Introductory Figure 14. CRISPR-Cas9 system in bacteria (overview).

CRISPR class II system is the one used by *Streptococcus pyogenes*. Here we have represented a simplified schema showing its mechanism of action. Bacteria harbour a Cas9 locus coding for the Cas9 protein (among other factors), a CRISPR locus formed by *spacers* (foreign sequences) and *repeat* sequences and a tracrRNA locus. Exogenous sequences from invading pathogens are incorporated as *spacers* in the CRISPR locus. A crispr RNA (crRNA) is transcribed and hybridises the trans-activating crRNA (tracrRNA) via the repeats-derived sequence. The two RNA molecules guide the Cas9 endonuclease to the target sequence in the invading DNA. The Cas9 cleaves it, protecting bacteria of phage infections. Adapted from van der Oost et al., 2014.

target cleavage (Guilinger et al., 2014). As they are composed of many tandem repeats, one limitation of TALEs is that their amplification and cloning are laborious and expensive. They are big proteins, so their coding plasmids have a considerable size and can be difficult to transfect in some cell types. Another limitation in the choice of target sequences is the obligatory presence of a T in the 5' end. This constraint can be overcome by substituting some amino acids in the loop region of the repeat 1 (Tsuji et al., 2013). Because of their bacterial origin, TALEs are immunogenic. This is a limitation for their use in therapeutic applications. Nevertheless, several clinical trials are ongoing with the TALENs. For example, TALENs are used in CAR T-cell therapy (Chimeric Antigen Receptors), in which patients' immune cells are collected, genetically engineered to recognise tumor antigens and reinfused back in the patients. The Universal Chimeric Antigen Receptors (UCAR T-cells) are allogenic gene edited T-cells that are engineered to target specific antigens in tumors and in which essential genes triggering immune reject or conferring sensibility to conventional therapies have been inactivated via TALENs. For example, UCART19 (allogenic engineered T-cells expressing anti-CD19 CAR) are a treatment of CD19-expressing haematological malignancies. It is in phase I clinical trials for children and adults relapsed or refractory ALL (NCT02808442, NCT02746952). UCART123 is in an ongoing clinical trial for treatment of AML (NCT03190278) and relapsed or refractory Blastic Plasmacytoid Dendritic Cell Neoplasm (BPDCN) (NCT03203369). In June 2018, a new molecule, UCART22, has been approved by the FDA to enter clinical trials for the treatment of B-ALL. Some other applications involve treatment of HPV related cervical precancerous lesions (NCT03226470, NCT03057912 in comparison to CRISPR).

3.1.3 CRISPR-dCas9

The latest revolution in gene editing tools is the discovery of the CRISPR-Cas9 system. The Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) system is an adaptive immune system used by prokaryotes to defend themselves against invading pathogens through the acquisition and further recognition of their exogenous sequences (Barrangou et al., 2007; Mojica et al., 2005) (**Introductory Figure 14**). There are different types of CRISPR systems (van der Oost et al., 2014). The class II system is the one used by *Streptococcus pyogenes*, of which the CRISPR associated protein 9 (Cas9) is the most widely used in biotechnological applications (SpCas9). Briefly, the CRISPR locus in bacteria contains *spacers* (foreign sequences) and *repeat* sequences. Bacteria incorporate the exogenous sequences from viruses or phages as *spacers* in the CRISPR locus. Then, a crispr RNA (crRNA) from the CRISPR locus (formed by around 20 nt of the *spacer*-derived sequence and around 20 nt of the repeats-derived sequence) is transcribed and hybridises a trans-activating crRNA (tracrRNA) via the repeats-derived sequence. The two RNA molecules guide the Cas9 endonuclease to its target sequence in the invading DNA. If the same pathogen infects the bacteria, the RNAs will



Introductory Figure 15. Use of CRISPR-dCas9 as an epigenome editing tool.

A. Example of a targeted sequence by the sgRNA molecule in the genome. A sequence of around 20 nt is targeted. A PAM sequence (5' NGG 3' for SpCas9) just downstream of the targeted sequence is necessary for the recognition by the sgRNA. The part of the sequence closest to the PAM is called *seed* and is essential for the binding. The nucleotides in the 5' part of the sequence allow mismatches in the recognition, which can lead to off-target effects. The Cas9 endonuclease creates a double strand break 3 nt upstream the PAM.

B-E. Effector domains are fused directly to dCas9 and recruited to their target sequences via a sgRNA. **B.** An activator effector domain (e.g. VP64) targeted to a gene promoter activates gene expression. **C.** A repressor domain (e.g. KRAB) recruits chromatin modifiers and creates a heterochromatin state that silences gene expression. **D.** Epigenetic modifiers such as TET1 can demethylate CGIs in a gene promoter, thereby reactivating gene expression. **E.** Alternatively, a DNMT domain promotes methylation of CGI in gene promoters, leading to transcriptional repression.

F-G. Other systems based on CRISPR-dCas9 allow the recruitment of multiple proteins. **F.** SunTag system. dCas9 is fused to a GCN4 peptide array that is recognised by scFv antibodies anti-GCN4 fused to multiple effector domains (e.g. VP64). **G.** RNA aptamers in a scaffolding RNA that can be bound by RNA-binding proteins fused to effector domains.

detect the foreign DNA. An essential element for recognition is the Protospacer Adjacent Motif (PAM), a three nucleotides sequence just downstream of the targeted sequence (Shah et al., 2013). As the PAM is not present in the CRISPR bacterial locus, the bacteria can differentiate between its own and exogenous genetic material. In the case of *S. pyogenes*, the PAM motif is 5' NGG 3'. The Cas9 has two nuclease domains RuvC and HNH that cleave double-stranded DNA, inducing the degradation of the foreign DNA (**Introductory Figure 14 and 15A**). Cas9 system was adapted to create a genome editing tool for biotechnological applications (Doudna and Charpentier, 2014; Jinek et al., 2012). A chimeric single guide RNA (sgRNA) generated from the combination of the crRNA and tracrRNA guides the Cas9 protein to its target sequences. In the last years, it has been used for a large range of applications (Cho et al., 2013; Cong et al., 2013; Jinek et al., 2013; Mali et al., 2013a). Cas9 cleaves 3 nucleotides before the PAM sequence. Once a DSB is produced, the genome can be repaired by two pathways. Firstly, the Non-Homologous End Joining (NHEJ) DNA repair pathway that can result in insertions or deletions leading to the disruption of a gene. Secondly, the Homology Directed Repair (HDR), that in presence of a homologous sequence (e.g. a donor plasmid) can introduce new specific sequences into a targeted position in the genome. Later, an inactive version of Cas9 (dead Cas9, dCas9) was engineered that contains two silencing mutations in both nuclease domains (D10A and H841A). dCas9 does not induce DSB at the target site but retains the ability to target specific sequences through the sgRNA, therefore it can be fused to effectors and used as a DBD (Jinek et al., 2012; Qi et al., 2013) to modulate gene expression, to perform epigenome editing and to decipher functions of regions in the genome (**Introductory Figure 12C and 15B-G**).

The main difference between the dCas9 system and the other DBD systems described before is that dCas9 is based in nucleotide-nucleotide interactions instead of protein-DNA interactions. This has major advantages, for example, a simple and fast design and implementation. The design of sgRNAs is cheaper than the construction of protein modules. However, the CRISPR-Cas9 system also has some limitations. Because some mismatches are allowed in the sequences, CRISPR-Cas9 systems have off-target effects. A very recent study shows that DNA breaks induced by CRISPR-Cas9 promote deletions extending long distances (several kilobases) and DNA damage in distant regions from the cut site in several mouse and human cell lines (Kosicki et al., 2018). Off-target effects can be improved by introducing point mutations in the Cas9 protein that weaken DNA binding activity and inhibit off-target activity at sites containing mismatches while preserving on-target activity (Slaymaker et al., 2016). In addition, the size of the Cas9/dCas9 protein is considerable, so plasmids coding for its sequence have a very large size. The requirement of a PAM sequence next to the recognition site also limits, to some extent, the possibilities of targeted sequences. Several strategies can be designed to overcome these problems, such as the modifications of the Cas9 protein or the

Effect	Mechanism of effect	Effector domain	References
Activation of gene expression	Recruitment of transcriptional activation complex and recruitment of histone modifications (H3K27ac, H3K4me)	VP64	* ZFP (Beltran et al., 2007; Ji et al., 2014; Liu et al., 2001) * TALE (Gao et al., 2013; Perez-Pinera et al., 2013a; Zhang et al., 2011a) * dCas9 (Chavez et al., 2015; Farzadfard et al., 2013; Gilbert et al., 2013; Liu et al., 2016b; Maeder et al., 2013a; Mali et al., 2013b; Perez-Pinera et al., 2013b) * SunTag (Gilbert et al., 2014; Tanenbaum et al., 2014)
		VP48 and VP160	* dCas9 (Cheng et al., 2013)
		p65	* ZFP (Ji et al., 2014; Liu et al., 2001) * TALE (Cuculis et al., 2016) * dCas9 (Gilbert et al., 2013; Konermann et al., 2015)
		VPR (VP64-p65-Rta)	dCas9 (Chavez et al., 2015)
Repression of gene expression	Prevention of transcription complex formation. Recruitment of other repressive modificators (histone methyltransferases for H3K27me3 and H3K9me3 and histone deacetylases)	KRAB	* ZFP (Beerli et al., 1998; Stolzenburg et al., 2015) * TALE (Cong et al., 2012; Gao et al., 2014b; Zhang et al., 2015b) * dCas9 (Gao et al., 2014b; Gilbert et al., 2013, 2014; Kearns et al., 2015; Klann et al., 2017; Ma et al., 2014; O'Geen et al., 2017; Thakore et al., 2015)
		Sin3a	TALE (Konermann et al., 2013)
		SID	TALE (Cong et al., 2012)
Gene promoter methylation is linked with repression of gene expression	DNA methylation	DNMT3A	* ZFP (Kungulovski et al., 2015; Nunna et al., 2014; Rivenbark et al., 2012; Stolzenburg et al., 2015) * TALE (Bernstein et al., 2015; Li et al., 2015a) * dCas9 (Lin et al., 2018; Liu et al., 2016b; McDonald et al., 2016; Vojta et al., 2016) * dCas9 SunTag (Huang et al., 2017; Pflueger et al., 2018)
		DNMT3B	* dCas9 (Lin et al., 2018)
		DNMT3A/3L	* ZFP (Siddique et al., 2013) * TALE (Bernstein et al., 2015) * dCas9 (Amabile et al., 2016; Saunderson et al., 2017; Stepper et al., 2017)
		M.SssI	dCas9 (Xiong et al., 2017)
H3K9me Transcription repression	H3K9me	G9a (EHMT2)	* ZFP (Heller et al., 2014; Snowden et al., 2002) * TALE (Cho et al., 2015) * dCas9 (O'Geen et al., 2017)
		GLP	* ZFP (Kungulovski et al., 2015)
		SUV39H1	* ZFP (Snowden et al., 2002) * dCas9 (O'Geen et al., 2017)
Transcription repression	H3K27me3	EZH2	*dCas9 (O'Geen et al., 2017)
Transcription repression (enhancer specific)	H3K4 demethylation and reduction of H3K27ac	LSD1	* dCas9 (Kearns et al., 2015) * TALE (Mendenhall et al., 2013)
Gene promoter demethylation is linked with activation of gene expression	DNA demethylation	TET1	* TALE (Li et al., 2015a; Maeder et al., 2013b) * dCas9 (Choudhury et al., 2016; Liu et al., 2018), scRNA (Xu et al., 2016) * SunTag (Gallego-Bartolomé et al., 2018; Morita et al., 2016)
		TET2	* ZFP (Chen et al., 2014; Huisman et al., 2016)
		TDG	* ZFP (Gregory et al., 2012)
Transcription activation	H3K4me3	PRDM9	* dCas9 (Cano-Rodriguez et al., 2016)
Transcription activation (sustainability)	H3K79me	DOT1L	* dCas9 (Cano-Rodriguez et al., 2016)
Transcription activation (ideal for enhancers)	H3K27ac	p300	* ZFP, TALE, dCas9 (Hilton et al., 2015), dCas9 (Garcia-Bloj et al., 2016), dCas9 screening (Klann et al., 2017)
Study of chromatin structure	Recruitment of enhancer-associated endogenous Ldb1	Ldb1 self-association domain	* ZFP (Deng et al., 2014)

Introductory Table 6. Gene expression modulation and epigenome editing with ZF, TALE and CRISPR-dCas9 coupled to effector domains

use of shorter gRNAs. We will take a closer look in *Section 3.3.2 Specificity of the dCas9 system*. Finally, similar to TALE proteins, CRISPR is a system originated from bacteria, meaning that its immunogenicity could be a limitation for its therapeutic use. Nevertheless, several clinical trials with the Cas9 system have been approved. For example, there are ongoing trails for the knockout of PD-1 in T-cells and further re-infusion in the patients of several types of cancers (NCT03081715, NCT02863913, NCT02793856) or the disruption of the gene encoding CISH that weakens immune cells in metastatic gastrointestinal epithelial cancer (NCT03538613). However, there is still some reluctance to its use in clinics. Recently, the FDA placed a CRISPR-Cas9-based clinical trial for the treatment of sickle cell disease on hold.

3.2 Epigenome editing applications

3.2.1 Modulation of gene expression

DNA binding domains can be fused to transcription activators or repressors to modulate gene expression. These tools selectively reprogram gene expression, which can be interesting to understand the regulation of a specific gene or to study therapeutic applications. In the **Introductory Table 6** we find a list of examples in which these systems have been used. The most widely used transcriptional activators are VP64, a tetrameric repeat of the herpes simplex VP16 minimal activation domain (amino acids 437-447: DALDDFDLDML) connected by GS-linkers (Beerli et al., 1998), and p65, a subunit of nuclear factor NF-kappa-B (Schmitz and Baeuerle, 1991). The optimal target distance for an efficient activation with TALE-VP64 is around -120 to -80 bp of the TSS, whereas for dCas9-VP64 it is -150 to -90 bp (Hu et al., 2014). Other reports indicate a wider range of action for dCas9-mediated gene activation going from -400 to -50 bp from the TSS (Gilbert et al., 2014). VP64 can activate gene expression by multiple mechanisms. For example, it has been shown that VP64 can recruit histone modulators such as p300 (H3K27 acetylation) (Gao et al., 2013, 2014b) and promote DNA demethylation (Gao et al., 2013). Interestingly, other reports show that the strong reactivation of the methylation-repressed *Oct4* gene by TALE or dCas9-VP64 do not correlate with promoter demethylation, meaning that VP64-dependent re-expression is not always sufficient to revert silencing epigenetic marks. Strategies acting directly on the epigenome should be developed to achieve a stable reprogramming (Hu et al., 2014). The fusion of different transcriptional activators in one artificial protein (e.g. dCas9-VP64-p65-RTa) can improve the transcriptional activation (Chavez et al., 2015). In this line, targeting several effectors to the same target region usually shows a synergistic effect improving gene activation (Chavez et al., 2015; Hu et al., 2014; Maeder et al., 2013b; Perez-Pinera et al., 2013a). Another way to recruit several proteins is via the SunTag system (*See Section 3.3.1 Efficiency of the dCas9 system*).

Comparison of TALE and dCas9 mediated gene activation revealed that TALE proteins can be more efficient than CRISPR-dCas9 system when targeting the same genes (Gao et al., 2014b). Nevertheless, one main advantage of dCas9 is the possibility of targeting multiple genes simultaneously. The multiplexing is much easier with the CRISPR-dCas9 tool because of the simplicity of the design and synthesis of gRNAs compared to laborious and time-consuming cloning of protein modules. In fact, it is possible to create libraries of sgRNA allowing to study gene regulation of gene families or at the whole-genome level for genome-wide dCas9-based screening (Joung et al., 2017; Konermann et al., 2015).

In contrast, DBD fusion proteins can also be used to inhibit gene expression and they represent an interesting alternative to interference RNA. Targeting DBDs alone (e.g. dCas9 without effector domain) can decrease gene expression by steric constraints (Gao et al., 2014b), but recruitment of endogenous repressors and histone modification actors leads to a more robust silencing. Krüppel-associated Box (KRAB) is a repressor domain widely used in epigenome editing. Its targeting induces heterochromatin formation, increasing histone repressive marks such as H3K27me3 and H3K9me3 and decreasing active histone marks H3K4me3 or H3K27ac (Kearns et al., 2015; Ma et al., 2014; Thakore et al., 2015). KRAB can also stimulate *de novo* DNA methylation in embryonic development (Wiznerowicz et al., 2007), but DNA methylation as a secondary effect of targeted gene repression has not been demonstrated in cellular systems up to now. KRAB performs a strong but transient repression, suggesting that other mechanisms should be implicated in a stable gene silencing (O'Geen et al., 2017). The optimal region to target the effectors and achieve the best inhibition is found -50 to +300 bp from the TSS and a peak of silencing is achieved when targeting 50-100 bp downstream (Gilbert et al., 2014). Other silencing effectors such as mSin interaction domain (SID) are also able to repress gene expression in an efficient manner when fused to TALEs (Cong et al., 2012).

The potential applications of TALE or dCas9-based technologies for targeted gene modulation are multiple. For example, targeted transcriptional activation has been used to improve cell reprogramming (Chavez et al., 2015; Gao et al., 2013) or for reactivation of silenced tumor suppressor genes in cancer cells (Beltran et al., 2007).

3.2.2 Epigenome editing of DNA methylation and histone marks

Activators and repressors have been shown to efficiently modulate gene expression. However, robust and stable changes in expression levels are not achieved, possibly because it is necessary to modulate epigenetic marks such as DNA methylation and histone post-translational modifications to achieve stable reprogramming of transcription. Until recently, it was only possible to reprogram these marks with drugs having a genome-wide impact. In the last years, a huge quantity of studies has tried to take advantage of programmable DNA binding technologies to target epigenetic effectors to one or several genes. This can help to interrogate the functional link between epigenetic marks of regulatory elements, gene expression and cell phenotype, as well as to improve the persistence in time of gene-specific transcriptional reprogramming, which is a key if we think about future clinical applications. An overview of these studies is presented in the next paragraphs.

i. Epigenetic repression

As we have seen in the *section 3.2.1*, transcription repressors such as KRAB are not capable of inducing permanent gene silencing. DNA methylation is known to act as a long-term repressor of gene expression, so its use in epigenome editing could overcome these limitations.

DNMT3A has been fused to DBD in different contexts: first zinc-finger proteins, then TALEs and more recently dCas9 (**Introductory Table 6**). When DNMT3A is fused to dCas9, it induces methylation with a peak at around 5 days after transfection (Kungulovski et al., 2015; McDonald et al., 2016; Vojta et al., 2016). Methylation is higher around the binding site of the DBD. For example, when the DBD is dCas9, a peak of methylation is found within 30-50 bp from the sgRNA binding site (McDonald et al., 2016; Vojta et al., 2016). Multiple gRNAs targeting the same region can help to extend the area of methylation (McDonald et al., 2016; Vojta et al., 2016). The increase of methylation within targeted regions is around 20-50% on average, but individual CpGs can show more than 60% increase in methylation. This methylation on gene promoters is linked with gene expression silencing (Bernstein et al., 2015; McDonald et al., 2016; Rivenbark et al., 2012; Stepper et al., 2017; Vojta et al., 2016). The addition of DNMT3L to DNMT3A effectors can improve the efficiency of methylation (Bernstein et al., 2015; Saunderson et al., 2017; Siddique et al., 2013; Stepper et al., 2017). The effects on the stability of the induced methylation are variable. Some reports indicate that methylation and gene repression are progressively lost after transitory transfection of ZF-DNMT3A or dCas9-DNMT3A (Kungulovski et al., 2015; Vojta et al., 2016). Interestingly, both studies performed transitory transfections in cancer cells (HEK293 and SKOV3 respectively), so the pathways involved in maintenance of methylation can be impaired. On the other hand, several reports show a robust persistence of methylation and gene repression (up to 50 days) after targeted DNMT3A methylation

even in *in vivo* xenografts models (McDonald et al., 2016; Rivenbark et al., 2012; Stolzenburg et al., 2015). In all these cases, the fusion proteins containing the effector domains are integrated into the genome and inducible by doxycycline. Even if the effector domains are apparently not expressed when the doxycycline is removed, we cannot exclude the possibility that imperceptible leaks of the inducible expression lead to minimal expression of the DBD-DNMT3A constructs sufficient to maintain the repressed state. Amabile et al. showed that permanent methylation-dependent silencing of a gene can be achieved in transient transfections by combining targeting of DNMT3A, DNMT3L and KRAB in HEK293T cells and an AML cell line (Amabile et al., 2016). Besides an increase in DNA methylation, the authors observed also an increase in H3K9me3 and a diminution in H3K4me3. This could mimic the sequential association of epigenetic marks during early development, where DNA methylation promotes a long-term repression of gene expression. Furthermore, reactivation of the gene was only achieved by using DNMT inhibitors or targeted dCas9-TET1, suggesting that DNA methylation is the main factor responsible of the silencing. These studies show that the persistence of targeted DNA methylation can be locus and cell type dependent. One locus can allow repression more easily than others depending on the chromatin context.

Another main concern of targeted DNA methylation by DBD-DNMT3A constructs is the specificity of the DNA methylation and the possible methylation in off-target regions. We will develop this point and how we can overcome this problem with the SunTag strategy in *section 3.3.2 Specificity of the dCas9 system*.

Alternatively, epigenetic repression can be achieved through the induction of histone repressive marks. ZFPs, TALEs or dCas9 fused to histone demethylase LSD1 reduce the presence of H3K4me2 and H3K27ac at enhancers, leading to the silencing of associated genes (Kearns et al., 2015; Mendenhall et al., 2013). H3K9 methylases such as EHMT2 or GLP promote the deposition of H3K9me2/3 at its targeted sites, leading to the silencing of associated genes (Heller et al., 2014; Snowden et al., 2002). The repression linked to H3K9me3 deposition is not stably maintained, probably as a result of the secondary recruitment of acetyltransferases that induces gene expression (Kungulovski et al., 2015). However, a combination of epigenetic modifiers such as EZH2, DNMT3A and DNMT3L can induce long-term repression (O'Geen et al., 2017).

ii. Epigenetic activation

DBDs can be fused to epigenetic effectors that catalyse DNA or histone modifications such as TET family proteins or p300 in order to activate silenced genes (**Introductory Table 6**).

Human TET1 (hTET1) is the most widely used effector domain. Targeted demethylation of CpG islands on gene promoters leads to reactivation of gene expression (Choudhury et al., 2016; Liu et al., 2016b; Maeder et al., 2013b; Morita et al., 2016; Xu et al., 2016). The level of re-expression varies according to the level of promoter demethylation. Sometimes, a slight demethylation in a specific CpG is sufficient to reactivate a gene. For example, the demethylation of the cytosine at -266 bp from the TSS of the *HBB* gene is critical for its re-expression (Maeder et al., 2013b), suggesting that demethylation of specific CpGs in enhancers or near genes TSS can make methylation-sensitive TF binding sites accessible. In general, the demethylation takes place more efficiently on the closest CpGs to the DBDs binding site (Choudhury et al., 2016; Maeder et al., 2013b), spreading in distances of around 200bp in average (Maeder et al., 2013b; Morita et al., 2016). dCas9-TET1 showed a higher efficiency than the TALE-based tool when used in the same context, however this was an isolated example (Liu et al., 2016b). The use of a SunTag dCas9-TET1 increases the efficiency and decreases dramatically the off-target effects (Morita et al., 2016). As TET proteins catalyse the conversion of 5mC into 5hmC, some of these studies also analysed the presence of this intermediate, showing an increase of the 5hmC at the boundaries of targeted sequences (Choudhury et al., 2016; Liu et al., 2016b). Targeted demethylation switches the chromatin status to an active conformation. Liu et al. demonstrated that a decrease in DNA methylation at the *FMR1* promoter was accompanied by the presence of RNA polymerase II, the deposition of H3K4me3 and H3K27ac and a diminution of H3K9me3 (Liu et al., 2018). Interestingly, targeted DNA demethylation by human TET1 works also in plants (Gallego-Bartolomé et al., 2018). Transitory transfections in human cells with TALEs or dCas9-hTET1 show a peak of demethylation and gene reactivation around 3-4 days after transfection, but these effects are progressively lost when the cells are cultured (Maeder et al., 2013b; Xu et al., 2016). In Fragile X-syndrome-induced pluripotent stem cells (FXS-iPSC) expressing a dCas9-hTET1 construct in a stable manner, the re-expression of *FMR1* gene is maintained for at least two weeks after the inhibition of the dCas9 molecule (Liu et al., 2018). In this case, the targeted region was a repeated sequence near the promoter, so one individual gRNA could target many repeated sequences. In the same study the authors injected methylation-reprogrammed neuronal precursor cells (NPCs) in mice brains and showed that the reprogramming was stably maintained for at least 3 months. Even if these results are achieved with a constitutively expressed dCas9-TET1, the efficiency of the technique *in vivo* is promising (Liu et al., 2018). The catalytic domain of mouse TET proteins has also been tested. Mouse TET1 (mTET1) was shown to induce a partial demethylation of *Foxp3* in T regulatory

cells, but this demethylation was not sufficient to stabilise its expression (Okada et al., 2017). Mouse TET2 has also been shown to efficiently reprogram gene expression through gene promoter demethylation (Chen et al., 2014; Huisman et al., 2016).

Other systems induce changes in histone marks that enhance gene activation. dCas9, TALEs and ZFPs coupled to p300 efficiently deposit H3K27ac in promoter and enhancers, reactivating gene expression (Hilton et al., 2015). The ability of p300 to promote transcription from enhancers makes it a suitable tool for functional studies. Combination of activating effectors such as p300, VP64 and VPR in a promoter dramatically increased efficiency. For example, combining dCas9-p300 or dCas9-VPR with TALE-VP64 reactivated the epigenetically silenced gene *MASPIN* more than 1000-fold compared to controls in HEK293T cells (Garcia-Bloj et al., 2016). Cano-Rodriguez et al. suggested that the presence of H3K79me3 is necessary to the stable reprogramming of genes (Cano-Rodriguez et al., 2016). Indeed, they targeted the deposition of H3K4me3 to promoters via dCas9-PRDM9 and the deposition of H3K79me3 via dCas9-DOT1L. Stable reactivation depended on the methylation context of the promoter: low methylated-promoters correlated with stable epigenetic reactivation, whereas hypermethylated promoters are only transiently activated. For example, strong epigenetic reactivation of *PLOD2* (slightly methylated promoter and silenced in C33A cells) using both editors was stable even 20 days after transfection (Cano-Rodriguez et al., 2016).

3.2.3 Other applications of epigenome editing tools

As epigenetic marks correlate with transcription status, epigenome editing tools can be used to decipher the mechanisms of epigenetic regulation, describe the functions of candidate regulatory regions, study the crosstalk between epigenetic marks or identify new genes involved in cell activity. Some examples are presented below.

Functional analysis of candidate regulatory regions

Targeting epigenetic enzymes to candidate regulatory regions can help to validate their functions. These strategies could be used to validate methylation-silenced cryptic promoters, non-coding genes or methylated-regions involved in alternative splicing. Concerning the latter, several recent studies point towards a link between DNA methylation and splicing through modulation of the elongation rate of the RNApol II or recruitment of splicing factors to methylated sites (Lev Maor et al., 2015). To validate these models, targeting methylation or demethylation to specific sequences in alternative splicing control regions could help to better understand the role of DNA methylation in splicing. Another application of dCas9-based methods is the validation of candidate enhancers. For example, targeting of LSD1 or p300 to enhancers provides a rapid and powerful approach to validate

the function a enhancer elements and identify their target genes without disruption of the genetic architecture (Hilton et al., 2015; Kearns et al., 2015; Mendenhall et al., 2013).

Genome-wide functional screening

In this line, genome-wide screens can be designed with the aim to expose unknown regulatory elements or genes involved in defined cellular pathways. The CRISPR-Cas9 system has been adapted to use in combination with genome-wide sgRNAs libraries targeting all genes to perform genome-scale CRISPR-Cas9 knockout screenings (Chen et al., 2015c; Wang et al., 2014b). In complement, genome-wide loss and gain of function screenings based on dCas9-mediated recruitment of repressors or activators have also been developed. As an example, cells expressing dCas9-SunTag-VP64 or dCas9-KRAB were transduced with sgRNAs libraries and treated with pulses of cholera and diphtheria toxins, allowing the identification of genes regulating the resistance or sensibility to these compounds (Gilbert et al., 2014). Klann et al. created stable cells coding for either dCas9-p300 or dCas9-KRAB and used a lentiviral sgRNA library to identify novel regulatory elements involved in the regulation of β-globin and HER2 genes in human cells (Klann et al., 2017). Finally, another report screened for candidate genes conferring resistance to BRAF inhibitor by performing a genome-scale activation screen with dCas9-VP64 and a sgRNA library targeting all RefSeq genes (Konermann et al., 2015). *BRAF* is an oncogene activated frequently in cancer that leads to the constitutive activation of mitogen activated protein kinases (MAPK) pathways (Konermann et al., 2015).

Chromatin organisation

Genome editing tools can also be used to manipulate chromatin organisation by inducing chromatin looping and study its effects on gene regulation. For example, targeting of the self-association domain of Ldb1 by artificial ZF to the promoter of β-globin leads to the formation of a control region-promoter loop that results in the reactivation of the silenced embryonic globin gene in adult mice cells (Deng et al., 2014). More recently, a new CRISPR-dCas9 based tool has been developed to manipulate chromosomal looping in mammalian cells (Morgan et al., 2017). In this system, dCas9 proteins are fused to proteins of the plant ABA signalling pathway, and juxtaposition of two chromosomal loci can be induced by addition of ABA which induces protein dimerisation.

Another way to study the organisational changes in the nucleus is by fusing the dCas9 to fluorescent proteins like GFP or mCherry. This enables the study of spatiotemporal dynamics of genomic loci in living cells, the visualisation of repetitive sequences or the observation of telomere dynamics (Chen et al., 2013a; Fu et al., 2016). The labelling of multiple loci through sgRNAs scaffolds

combined to different fluorescent proteins allows simultaneous imaging of multiple loci to study their dynamics at the same time (Ma et al., 2016).

3.3 Future challenges in epigenome editing

The future challenges with epigenome editing tools are focused on the optimisation of the techniques to improve their efficiency and specificity. CRISPR-dCas9 being the most versatile, easy to use and adaptable tool, it holds the most potential for future biotechnological applications, and consequently several strategies have been developed to improve its efficiency and specificity. We discuss some of them in the following section.

3.3.1 Efficiency of the dCas9 system

SunTag system

The SunTag system was developed to improve CRISPR-dCas9 efficiency. This system allows the recruitment of multiple proteins to a target sequence. It consists in a repetitive peptide array of variable number (4-10 peptides GCN4 seems to be the most widely used) fused to the dCas9 and recognised by single chain anti-GCN4 scFv antibody fused to the effector domain (**Introductory Figure 15F**). The advantage compared to direct fusion proteins is that it allows for the recruitment of multiple effector proteins at the same site. It was used for the first time coupled to the transcriptional activator VP64, showing a strong activation of gene expression (Tanenbaum et al., 2014). The system has also been efficiently used coupled to TET1, DNMT3A or fluorescent proteins (Huang et al., 2017; Morita et al., 2016; Pflueger et al., 2018; Ye et al., 2017). However, as explained below, its main advantage could be a higher specificity (Pflueger et al., 2018).

Aptamers

Guide RNAs can be extended to form scaffold RNAs, which harbour secondary structures that can be bound by RNA-binding proteins. Thus, multiple effector domains can be recruited to a specific region by fusing them to the RNA-binding proteins. For example, two copies of the MS2 RNA loop original from bacteriophages can be added to the 3' end of the sgRNA and expressed in combination with a dCas9 and a MS2-VP64 constructions (Mali et al., 2013b). Depending on the number of hairpins, more or less proteins can be recruited. Other aptamers that can be used are PP7, Com and PUF binding site (PBS) (Cheng et al., 2016; Konermann et al., 2015; Zalatan et al., 2015) (**Introductory Figure 15G**).

Spaciotemporal control

Other strategies aim to control the expression of Cas9 or dCas9 and the effector recruitment upon an external signal like chemical or light induction. dCas9 can be divided into two domains fused to binding proteins that dimerise in the presence of an external drug like rapamycin (Zetsche et al., 2015a) or blue light (Nihongaki et al., 2015).

3.3.2 Specificity of the dCas9 system

Several genome-wide studies performing ChIP-seq have shown that dCas9 binds multiple genomic sequences and that the 5 nt close to the PAM proximal region in the gRNA (called "seed" sequence) are essential to the binding. This binding takes place mostly in open chromatin sequences and suggests that off-target events could be frequent when working with dCas9 fused to effector domains. In contrast, cleavage by Cas9 endonuclease is rarely observed in off-target sites, meaning that a more extensive annealing with target DNA could be required for catalytic activity (Duan et al., 2014; Kuscu et al., 2014; O'Geen et al., 2015; Wu et al., 2014). In fact, dCas9-DNMT3A direct fusion has been shown to be highly unspecific in several recent publications. Meissner's laboratory performed WGBS and demonstrated that a doxycycline inducible dCas9-DNMT3A construction (in the absence of sgRNA) resulted in a widespread gain of methylation in mouse ES cells (Galonska et al., 2018). By hybrid capture probes, they showed that the presence of sgRNAs did not have any impact on reducing the off-target activity. The disadvantage of this technique is that they studied regions of 20 kb on either side of the sgRNAs, so the observed effects could also be a consequence of the spreading of targeted methylation (Galonska et al., 2018). In this line, Luo's team also showed that the expression of dCas9-DNMT3A or dCas9-D3B had substantial off-target effects. Even if it was not a global gain of methylation, they found more than 1000 DMRs. Interestingly, ChIP-seq and WGBS showed a low correlation between dCas9 peaks and DMRs (Lin et al., 2018). Widespread off-target activity of the direct fusion protein dCas9-DNMT3A was also observed by Lister's team (Pflueger et al., 2018). To try to optimise the system, they made the hypothesis that, in contrast to direct fusion proteins, indirect recruitment of DNMT3A by using a modular dCas9 system utilising the SunTag array would help to favour on-target DNA methylation and reduce spurious off-target DNA methylation. This system did not improve the efficiency of on-target methylation compared to the dCas9-DNMT3A system, but it dramatically reduced the off-target effects. ChIP-seq showed very few off-target peaks of dCas9 that partially matched with the sgRNA sequence, meaning that an optimised sgRNA design could eliminate off-targets (Pflueger et al., 2018). This reduction in off-target effects by using dCas9-SunTag-DNMT3A system has also been described in other reports (Huang et al., 2017). As these wide-genome studies have all been carried out with the DNMT3A effector, we cannot exclude the possibility that the observed off-target effects are dependent of the DNMT3A

activity. Morita et al. showed that dCas9-hTET1 off-target activity was also reduced with the SunTag system (Morita et al., 2016).

Several other strategies can be implemented to improve the specificity of the system.

trugRNAs

Off-target effects of CRISPR-Cas9 or dCas9 system can be diminished by playing with the length of sgRNAs. Canonical sgRNAs are 20-nucleotides long, and the PAM proximal sequence (seed) is critical for Cas9 specificity (**Introductory Figure 15A**). Mismatches in this zone can avoid nuclease activity whereas mismatches in 5' region are more tolerated (Jinek et al., 2012). The use of shorter sgRNAs of 17-18 nt (truncated sgRNAs, trugRNAs) reduces the off-target effect without affecting the on-target efficiency (Fu et al., 2014).

Cas9 modifications

Some studies have reported that mutations in the Cas9 residues that are essential for interactions with the sgRNA or the target DNA sequence can improve specificity. These mutations promote weaker interactions among the complex, leading to a reduction in off-target interactions. For example, SpCas9-HF and eSpCas9 are variants with residue substitutions in the hydrogen bonding between SpCas9 and the DNA (Kleinstiver et al., 2016) and with neutralisation of charges essential to stabilise the interaction (Slaymaker et al., 2016), respectively. Other engineering strategies of SpCas9 protein are focused on creating new variants with novel PAM preferences different from NGG. The NGG requirement is stringent and could limit the sequences that can be targeted. New SpCas9 variants can recognise new PAM sequences (NGCG, NGAG) and decrease the off-target effects observed with canonical PAMs NGG, NAG or NGA (Kleinstiver et al., 2015).

Cas9 orthologues from other bacterial and archaeal species can also overcome some limitations. For example, *Staphylococcus aureus* Cas9 (SaCas9) has similar efficiency than SpCas9 with the advantages of presenting a smaller size appropriate for *in vivo* Adeno-associated virus (AAV) delivery and a different PAM. Other orthologues like *Neisseria meningitidis* Cas9 or *Streptococcus thermophilus* Cas9 have complex PAM so, even if they are less efficient, their use results in less off-target effects (Cebrian-Serrano and Davies, 2017; Ran et al., 2015). Other class II systems like Cas12a (Cpf1) use just one molecule of RNA to guide the protein to its target sequence. Moreover, it uses a T-rich PAM, which could simplify genome editing in AT-rich genome organisms (Zetsche et al., 2015b).

Finally, it has been recently discovered that the CRISPR system can be used to target RNA sequences and used as a more specific alternative to interference RNA. Indeed, Cas13 orthologue is

able to cleave transcripts. A catalytically dead version of Cas13 (dCas13) fused to adenosine deaminase acting on RNA type 2 can perform RNA editing by changing the adenosine nucleoside by an inosine (Abudayyeh et al., 2017; Cox et al., 2017). These RNA-editing tools are interesting biotechnological tools to perform fundamental research or for future clinical applications aiming at mutating transcripts containing pathogenic mutations.

3.3.3 Stability of dCas9-induced reprogramming

Even if the number of studies involving epigenetic editing is increasing, just a few have followed the stability of the induced marks in time. As already explained through this chapter, some reports found that the induced marks are lost once the effector domains are not expressed anymore, whereas others demonstrate stable changes. The effects on stability can be dependent on the chromatin context, the cell type and the tendency of individual locus to become silenced or activated. They can also be influenced by the technical approaches chosen or the time of exposure to the effector domain. Interestingly, the two articles that showed the strongest stable epigenetic silencing or reactivation (with effects still present even 20-50d after transfection) showed a well-established crosstalk among DNA methylation and histone marks (Amabile et al., 2016; Cano-Rodriguez et al., 2016). Future research should focus on unravelling the kinetics of epigenetic changes and the interplay between different epigenetic marks involved in stable epigenetic reprogramming. Studying optimal combinations among known epigenetic effectors and other yet to be tested could improve the efficiency and stability of the reprogramming.

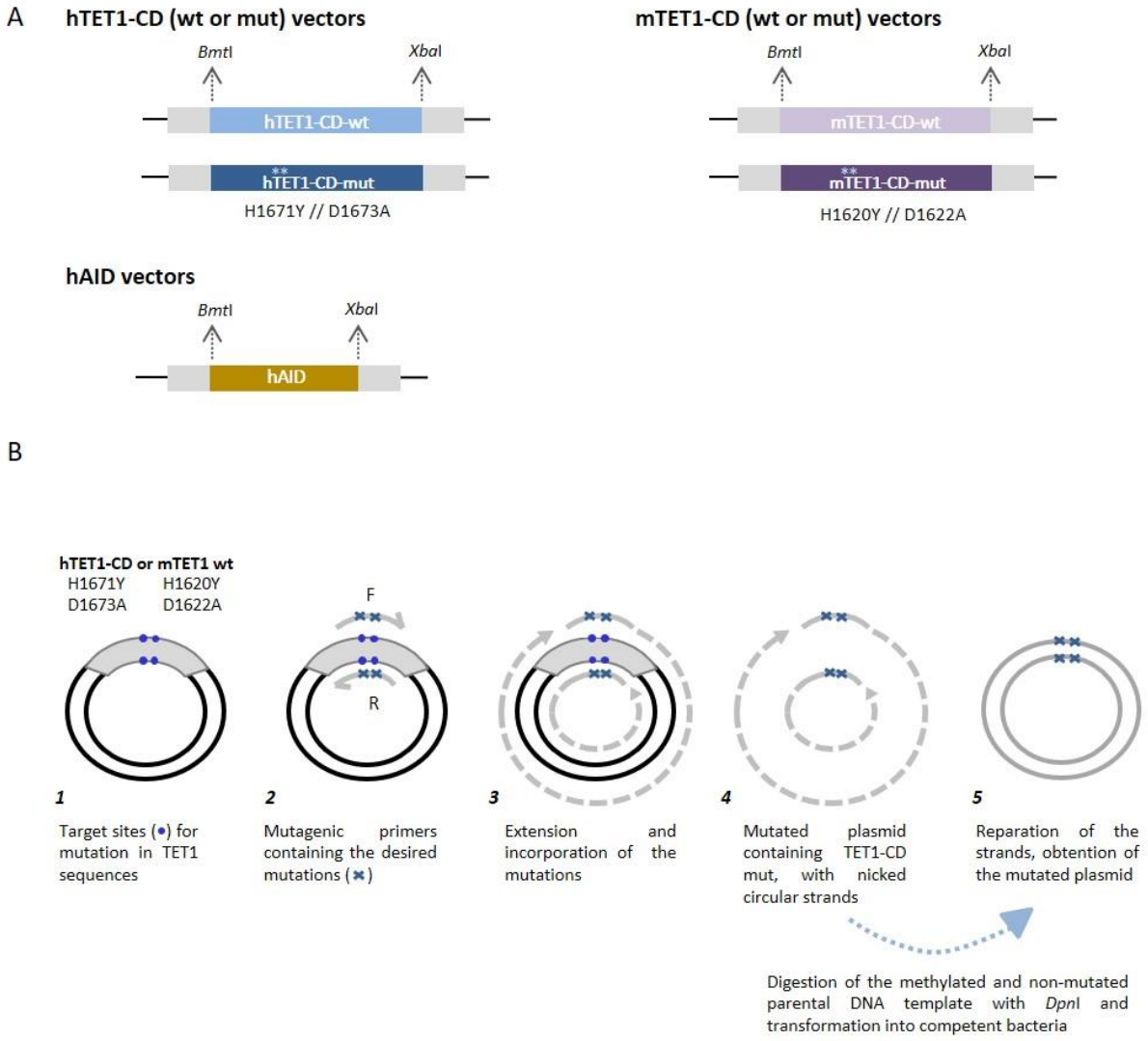
3.3.4 *In vivo* delivery of dCas9 for clinical applications

Given the aberrant patterns of methylation in cancer and other diseases, epigenome editing has emerged as a potential therapeutic field with powerful clinical applications. Targeted gene expression modulation has been shown to reprogram cancer cells proliferation *in vitro* and to be efficient *in vivo*. However, there are still barriers that need to be overcome before clinical use in humans, especially those related to the *in vivo* delivery.

Adeno-associated virus vectors (AAV) are suitable for gene therapy because of their efficiency and low immunogenicity. However, their cloning capacity is limited to around 4,7 kb. Since SpCas9 is too big to be incorporated in combination with the sgRNA, smaller Cas9 orthologues such as Cas9 from *Staphylococcus aureus* (SaCas9) are proposed as a promising alternative. SaCas9 combined with adenovirus delivery was shown to efficiently modify the *Pcsk9* gene in the mouse liver (Ran et al., 2015), or correct dystrophin expression by excision of the mutated exon in postnatal muscle tissue *in vivo* in a mouse model of Duchenne muscular dystrophy (DMD) (Nelson et al., 2016; Tabebordbar et al., 2016). Alternatively, SpCas9 and the sgRNA can be delivered separately in two

different AAV-vectors, which were also able to efficiently correct dystrophin expression *in vivo* in mouse (Long et al., 2016a). More recently, delivery of Cas9 via non-viral methods has been used to bypass the challenges of viral delivery. For example, cationic lipid-mediated delivery was used to deliver Cas9:gRNA complexes *in vivo* to correct deafness by disrupting the dominant deafness-associated allele of the *Tmc1* gene in a mouse model of human genetic deafness (Gao et al., 2017). Next generation delivery systems such as nanoparticles can also be considered (Sun et al., 2015). Systemic or targeted *in vivo* delivery of Cas9 systems is also a parameter that needs to be taken into consideration. AAVs can be either systemic or can be used for targeted delivery, as some of these vectors provide tissue tropism. In this line, tissue-specific promoters could also be used to limit the expression of the effector domains in the cells of interest. Up to now, there are no examples of clinical trials with epigenetic editing tools. However, as we detailed at the beginning of this chapter, clinical trials involving DBD fused to nucleases or CRISPR-Cas9 system have been approved. Most of these trials are based on *ex vivo* strategies consisting on the collection of autologous or allogenic blood cells, correction of the target gene in the laboratory and transplantation into the patient. Some clinical trials designed to treat cervical HPV related malignancies use suppositories or plasmids in gel containing the editing tools. Results of these clinical trials and future improvements in methods for *in vivo* delivery will give us more clues about the possible use of epigenetic effectors for human clinical studies.

METHODS



Methods Figure 1. TET1 constructions and hTET1 directed mutagenesis.
A. Human and mouse TET1 constructions (wt and mutant) and human AID.
B. Schematic representation of the directed mutagenesis on human and mouse TET1-CD to eliminate their catalytic activities.

METHODS

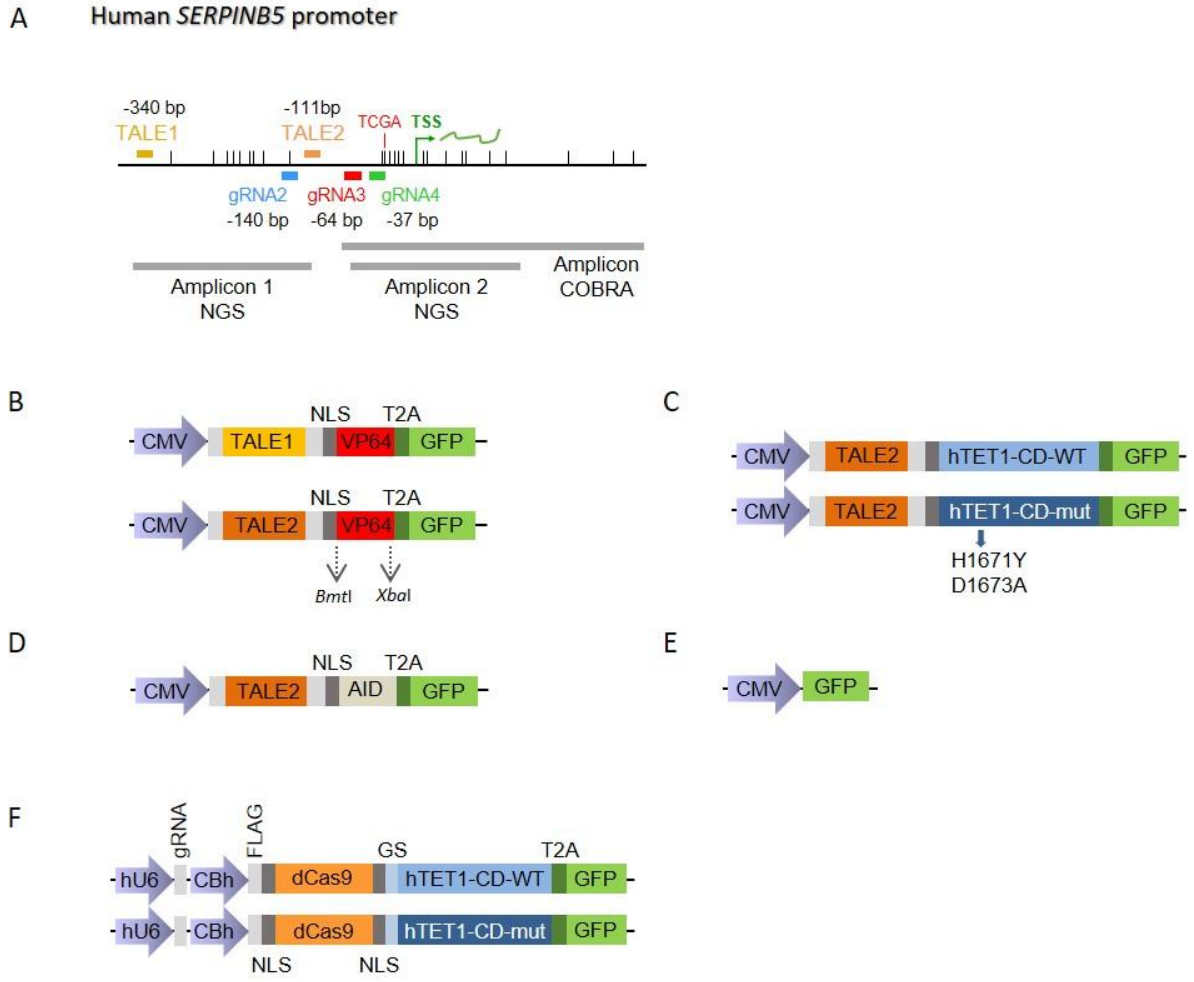
1 Vectors

In the first part of the thesis I have used two strategies to target human *SERPINB5* promoter with programmable DNA binding modules: TALE and CRISPR-dCas9-based methods. In the second part of the thesis I have optimised the dCas9 system. I will describe all the plasmids used during the project in the next pages and the genetic engineering that I performed on them.

1.1 Vectors coding for effector domains

Sequences coding for the wild-type (wt) human TET1 catalytic domain (hTET1-CD) (2157 bp, 719 amino acids) and the wild-type (wt) mouse TET1 catalytic domain (mTET1-CD) (2019 bp, 673 amino acids) were synthesised by Integrated DNA Technologies (IDT) preceded by an NLS sequence. Both sequences were flanked by *Bmt*I and *Xba*I sites (**Methods Figure 1A**). As we wanted to use catalytic mutants as a control in our experiments, hTET1-CD wt and mTET1-CD wt were inactivated by directed mutagenesis (**Methods Figure 1B**). The two point mutations that eliminate catalytic activity of TET1 are H1671Y and D1673A (human) and H1620Y and D1622A (mouse), as previously described (Nakagawa et al., 2015; Tahiliani et al., 2009). As the two residues were close to each other, I mutated them in one single step by directed mutagenesis using primers containing the two desired mutations. The primer sequences are described in **Methods Table 1**. 50 ng of plasmid were used for the reaction in combination with 125 ng of each primer, 1X buffer HF, 0,15mM dNTP, 3% DMSO and 1U of Phusion High Fidelity DNA polymerase (Thermo Scientific) in a final volume of 30µL. The reaction was performed using the following conditions: 2min at 98°C, followed by 25 cycles of 1min30 at 98°C, 1min30 at 52°C, 7min30 at 72°C and a final step of 15min at 72°C. The samples were digested with *Dpn*I to eliminate the non-mutated template plasmid prior bacteria transformation. Plasmids sequences were verified by Sanger sequencing (GATC).

We also investigated a possible role of Activation-induced cytidine deaminase (AID) in targeted demethylation. For this purpose, human AID sequence (hAID, 594 bp) was also synthesised by IDT preceded by the NLS sequence and flanked by *Bmt*I and *Xba*I sites (**Methods Figure 1A**).



Methods Figure 2. TALE and CRISPR-dCas9-based strategies for targeting of the *SERPINB5* promoter.

A. Schematic illustration of the *SERPINB5* promoter region and TALE and sgRNAs binding sites. The vertical bars represent the CpGs. The TALE1 and TALE2 binding sites are located -340 bp and -111 bp from the TSS, respectively. The sgRNA2, sgRNA3 and sgRNA4 bind at -140, -64 and -37 bp from the TSS, respectively.
B-H. Schematic representation of the vectors used in the transfection experiments: TALE-VP64 (**B**), TALE2-hTET1-CD (wt and mut) (**C**), TALE2-hAID (**D**), eGFP plasmid (transfection control) (**E**), dCas9-hTET1 (wt and mut) – GFP (**F**).

1.2 Vectors used in transitory transfection experiments

First part of the thesis

1.2.1 TALE-VP64 plasmids

Plasmids targeting *SERPINB5* promoter coding for TALE1-VP64-T2A-GFP and TALE2-VP64-T2A-GFP were purchased from Labomics. TALE1 binds a region of 21 nucleotides located at -340 bp relative to the Transcription Start Site (TSS), whereas TALE2 binds a sequence located at -111 bp relative to the TSS (**Methods Figure 2A, Figure 2B and Table 2**). VP64 is a transcriptional activator consisting on a tetrameric repeat of the herpes simplex VP16 minimal activation domain (amino acids 437-447: DALDDFDLDML) connected by GS-linkers. After its first use in 1998 (Beerli et al., 1998) it has been widely used as a molecular biology tool to study gene expression. The expression of TALE1-VP64 and TALE2-VP64 is driven by a CMV promoter. The plasmids contain *Bmt*I and *Xba*I restriction sites flanking the VP64 sequence.

1.2.2 TALE-hTET1 and TALE-hAID plasmids

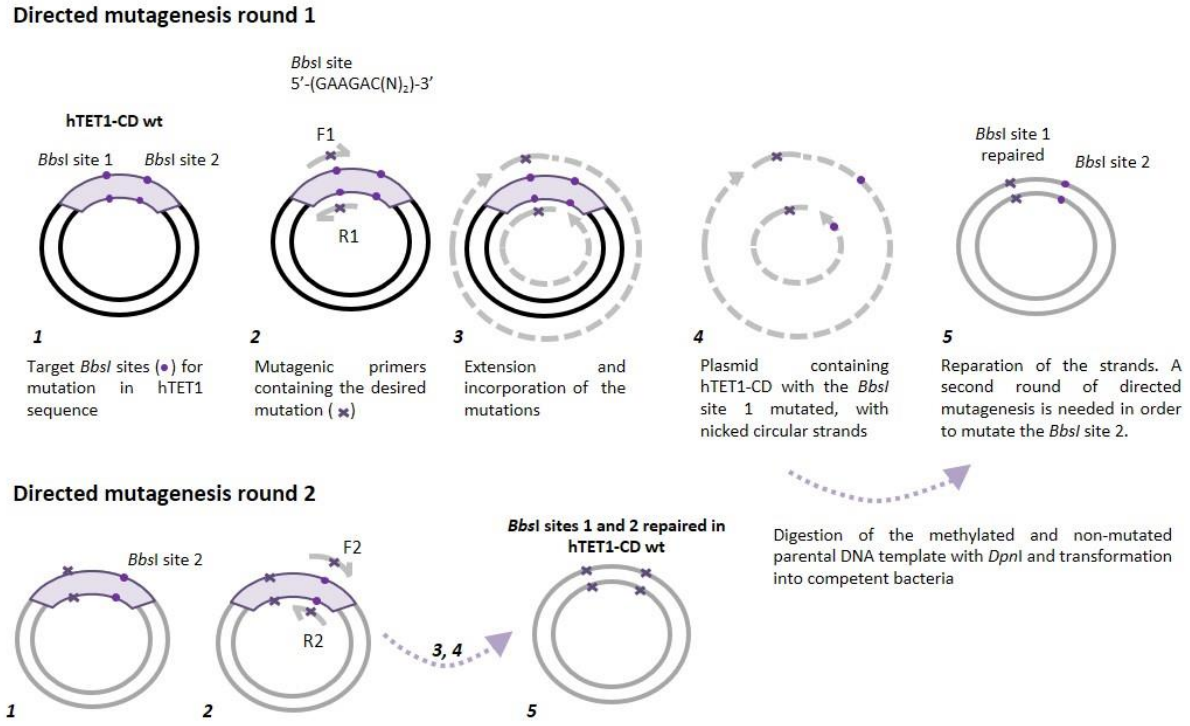
The construction of the TALE2-hTET1 (wt and mut) plasmids and TALE2-hAID was performed by substitution of the NLS-VP64 sequence in the TALE2-VP64 plasmid by restriction enzymes digestion with *Bmt*I and *Xba*I. The final fusion proteins formed were TALE2-NLS-hTET1-CD (wt_or_mut)-T2A-GFP and TALE2-NLS-hAID-T2A-GFP (**Methods Figure 2C and 2D**). We verified the plasmid sequences by Sanger sequencing (GATC).

1.2.3 pEGFP-C1 vector

pEGFP-C1 vector (**Methods Figure 2E**) was used as a control of the efficiency of the transfections.

1.2.4 dCas9-hTET1-GFP plasmids

The dCas9-hTET1-CD-GFP plasmids were constructed based on the pdCas9-DNMT3A-EGFP plasmid (Addgene #71666). The original plasmid was digested with the restriction enzymes *Bam*HI and *Fse*I to remove the DNMT3A fragment and replace it with the catalytic domain of human TET1 (wt or mut) that were amplified by PCR using forward and reverse primers introducing *Bam*HI and *Fse*I restriction sites (**Methods Figure 2F**). The final fusion proteins formed were FLAG(x3)-dCas9-NLS-GS_linker-hTET1_CD(wt_or_mut)-T2A-eGFP. The expression of the constructions was driven by a CBh promoter. We verified the plasmid sequence by Sanger sequencing (GATC). The CRISPR-dCas9 plasmid contains a site for the cloning of the guide RNA with the enzyme *Bbs*I that cuts at 5'-(GAAGAC$(N)_2$)-3' and 3'-(CTTCTG$(N)_6$)-5'. As the human TET1 CD contains two undesired *Bbs*I



Methods Figure 3. Elimination of *Bbs*I restriction sites in hTET1-CD-wt.

Schematic representation of the strategy used to eliminate two *Bbs*I sites present in hTET1-CD by two rounds of directed mutagenesis.

restriction sites, I performed two rounds of site-directed mutagenesis as described in **Methods Figure 3** with two different couples of primers. The primers were designed to induce silent mutations and are shown in **Methods Table 1**. The first couple of primers was designed to mutate the codon 436/719 AAG (K) by AAA (K) in the first *Bbs*I site. The second couple of primers was designed to mutate the codon 470/719 AGA (R) by AGG (R) in the second *Bbs*I site. 1-3 ng of plasmid were used for the reaction in combination with 1µM of each primer, 1X buffer HF, 0,2mM dNTP and 0,5U of Phusion High Fidelity DNA polymerase (Thermo Scientific) in a final volume of 20µL. Reaction was performed using the following conditions: 2min at 98°C, followed by 30 cycles of 10s at 98°C, 3min at 72°C, 10min at 72°C and a final step of 10min at 72°C. hTET1-CD-wt containing the *Bbs*I mutated sites was then mutated in order to inactivate its catalytic activity by a directed mutagenesis strategy as described before (**Methods Figure 1B**) and incorporated into the dCas9 plasmid.

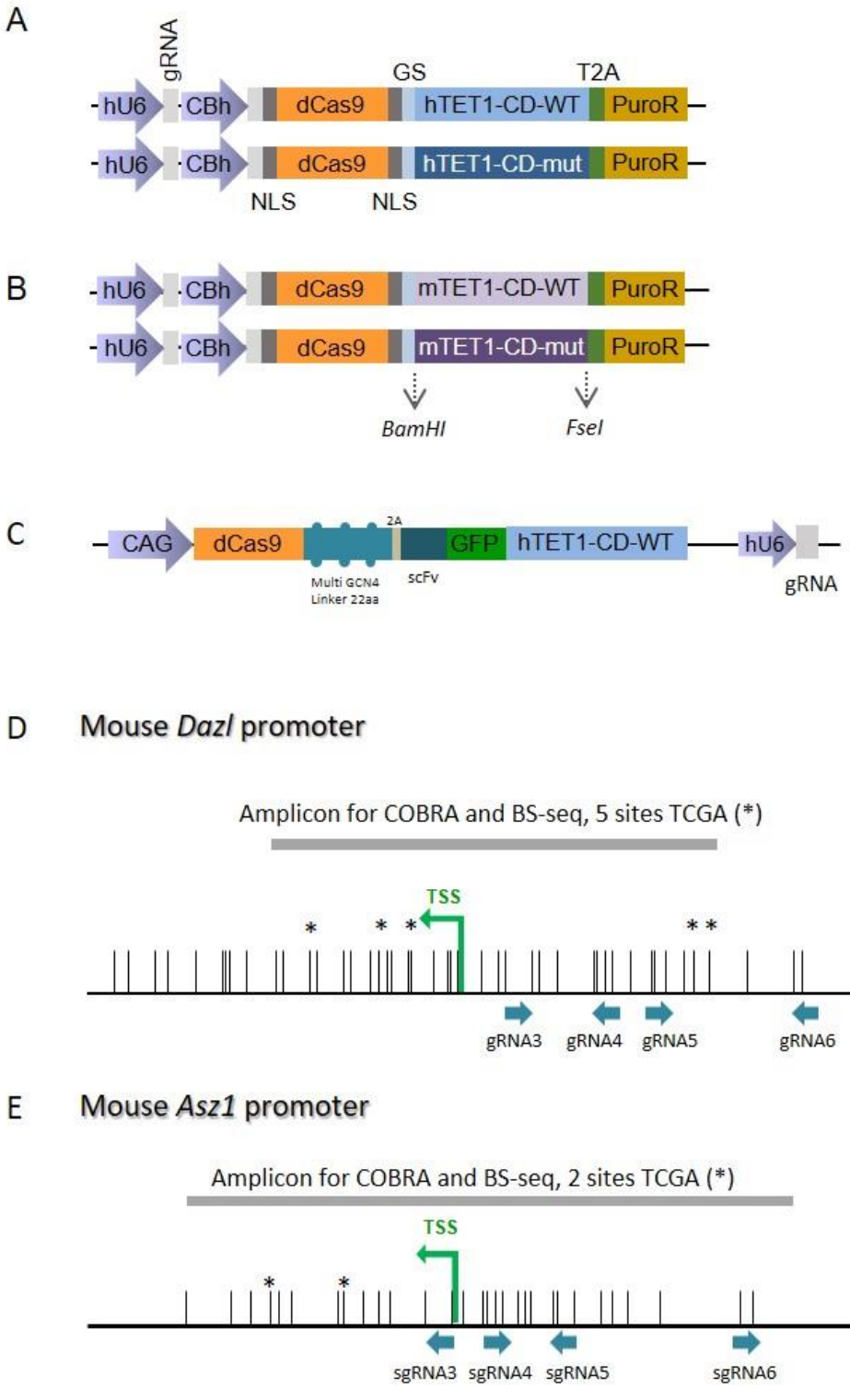
Three gRNAs targeting the *SERPINB5* promoter were designed using http://crispr.mit.edu, synthesised as pairs of oligonucleotides (IDT), annealed and cloned into the *Bbs*I sites of the pdCas9-hTET1-CD (wt or mut) plasmids. sgRNA2, sgRNA3 and sgRNA4 binding sites are located at -140 bp, -64 bp and -37 bp from the TSS (**Methods Figure 2A and Table 2**). The expression of the gRNAs was driven by a hU6 promoter. We verified the cloning of the gRNAs by Sanger sequencing (GATC).

Second part of the thesis

1.2.5 dCas9-hTET1-Puro and dCas9-mTET1-Puro plasmids

The EGFP sequence in the pdCas9-DNMT3A-GFP plasmid (Addgene #71666) was substituted with the puromycin sequence from PX459-V2 plasmid (Addgene #62988) by digestion with the restriction enzyme *Eco*RI. The pdCas9-DNMT3A-Puro was digested with the restriction enzymes *Bam*HI and *Fse*I to remove the DNMT3A fragment and replace it with the catalytic domain of human or mouse TET1 (wt or mut) (**Methods Figure 1A**) that were amplified by PCR using forward and reverse primers introducing *Bam*HI and *Fse*I restriction sites. The hTET1-CD had been previously engineered to eliminate two undesired *Bbs*I restriction sites as described in section *1.2.4* (**Methods Figure 3**) and mutated in order to inactivate its catalytic activity by a directed mutagenesis as described in section *1.1* (**Methods Figure 1B**). The final fusion protein formed was FLAG(x3)-dCas9-NLS-GS-hTET1_CD(wt or mut)-T2A-Puro (**Methods Figure 4A**) and FLAG(x3)-dCas9-NLS-GS-mTET1_CD(wt or mut)-T2A-Puro (**Methods Figure 4B**). The expression of the constructions was driven by a CBh promoter. We verified the plasmid sequences by Sanger sequencing (GATC).

The sgRNAs targeting *Dazl and Asz1* promoters were designed using http://crispr.mit.edu (**Methods Figure 4D and 4E, Table 4**), synthesised as pairs of oligonucleotides (IDT), annealed and cloned into the *Bbs*I sites of the pdCas9-hTET1-CD (wt or mut) or pdCas9-mTET1-CD (wt or mut) plasmids. The



Methods Figure 4. CRISPR-dCas9-TET1 based strategy to target the mouse *Dazl* and *Asz1* promoters.

A-C. Schematic representation of the vectors used in the transitory transfections: dCas9-hTET1 (wt and mut) – PuroR (**A**), dCas9-mTET1 (wt and mut) – PuroR (**B**), dCas9-hTET1 wt SunTag (**C**). **D-E.** Schematic illustration of *Dazl* and *Asz1* promoter regions and sgRNAs binding sites. The vertical bars represent the CpGs.

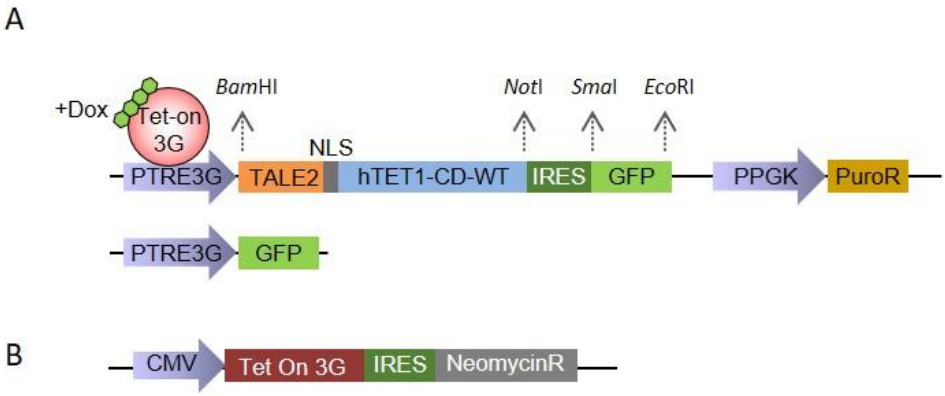
expression of the sgRNAs was driven by a hU6 promoter. We verified the cloning of the sgRNAs by Sanger sequencing (GATC).

1.2.6 dCas9-hTET1 SunTag plasmids

For experiments using the SunTag system, we used the all-in-one pPlatTET-gRNA2 vector (Addgene #82559) that codes for dCas9 fused to the GCN4 peptide arrays, the scFv of the anti-GCN4 antibody fused to TET1-CD, and contains a cloning sites for gRNAs (**Methods Figure 4C**) (Morita et al., 2016). Cloning of the gRNAs in the pPlatTet-gRNA2 vector was made according to a Gibson assembly strategy as follows. 20 bp gRNA sequences were designed via https://crispr.mit.edu (sequences are the same as for dCas9-TET1 direct fusion proteins, **Methods Table 4**). 19bp of the selected target sequences (3' end) were incorporated into two 60 mer oligonucleotides as indicated in **Methods Table 5**. The two oligos were annealed and extended to make a 100bp double stranded DNA fragment using Q5 High Fidelity Polymerase (NEB). 1,25 µM of each oligonucleotide were used in combination with 1X Q5 buffer, 1mM dNTPS and 0,4 U of Q5 High Fidelity Polymerase in a final volume of 10 µL. Reaction was performed upon the following conditions: 1min at 98°C, 3 cycles of 10s at 98°C, 30s at 65°C, 10s at 72°C and a final step of 10min at 72°C. pPlatTet-gRNA2 vector (Addgene #82559) was linearised using *AflII* restriction enzyme (NEB). The 100 bp fragments were incorporated into the digested plasmid using Gibson assembly. 300 ng of digested pPlatTet-gRNA2 plasmid were combined with 10 ng of 100 bp fragments and either 5 µL of Gibson assembly mastermix (NEB) in a final volume of 10µL or 2µL of 5X Fusion Takara mastermix in a final volume of 10µL. The assembling reaction was prepared on ice and incubated 30min at 50°C. Plasmids were amplified via bacteria transformation and sequences were verified by Sanger sequencing (GATC).

1.3 Lentiviral vectors for stable inducible MDA-MB-231 cell lines

In order to create an inducible MDA-MB-231 stable cell line expressing TALE2-hTET1 under the control of a doxycycline-inducible promoter, we subcloned the TALE2-hTET1 sequence in the MCS-1 and the EGFP sequence in the MCS-2 of the lentiviral pLVX-TRE3G-IRES plasmid (Tet-On 3G System, Clontech #631362) (**Methods Figure 5A**). A plasmid containing only the EGFP sequence cloned in the MCS-2 site was used to create a control cell line (**Methods Figure 5A**). The EGFP coding sequence (720 bp) was synthesised by Integrated DNA Technologies (IDT) flanked by the desired restriction sites (*SmaI* and *EcoRI*). TALE2-hTET1-CD wt was amplified by PCR using forward and reverse primers introducing *BamHI* and *NotI* restriction sites. We amplified the sequence with a High-fidelity DNA polymerase (Phusion DNA polymerase, Thermo Scientific) and we cloned it into the MCS-1 of the pLVX sequences. We verified the plasmid sequences by Sanger sequencing (GATC). The pLVX-Tet-On 3G plasmid (Clontech) codes for the rtTA transactivator protein under the control of a CMV promoter



Methods Figure 5. Schematic representation of the vectors for stable inducible MDA-MB-231 cells.

pLVX-TRE3G-TALE2-hTET1-IRES-GFP and pLVX-GFP **(A)**, pLVX-Tet-On 3G plasmid **(B)**.

(**Methods Figure 5B**).

2 Cell culture and transitory transfections

2.1 MDA-MB-231

MDA-MB-231 cells (ATCC: HTB-26) were cultured in RPMI 1640 medium (Gibco, Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (FBS) and 50 mg/ml gentamycin sulfate (Gibco, Thermo Fisher Scientific). Cells were cultured in a humidified atmosphere at 37ºC supplemented with 5% CO_2. All cell lines have been tested for mycoplasma contamination. For DNA methylation inhibition experiments, MDA-MB-231 cells were treated with 5-aza-2'-deoxycytidine (Sigma) at 0.2, 2 or 5µM and harvested for analysis 24h, 48h and 72h after the beginning of the treatment.

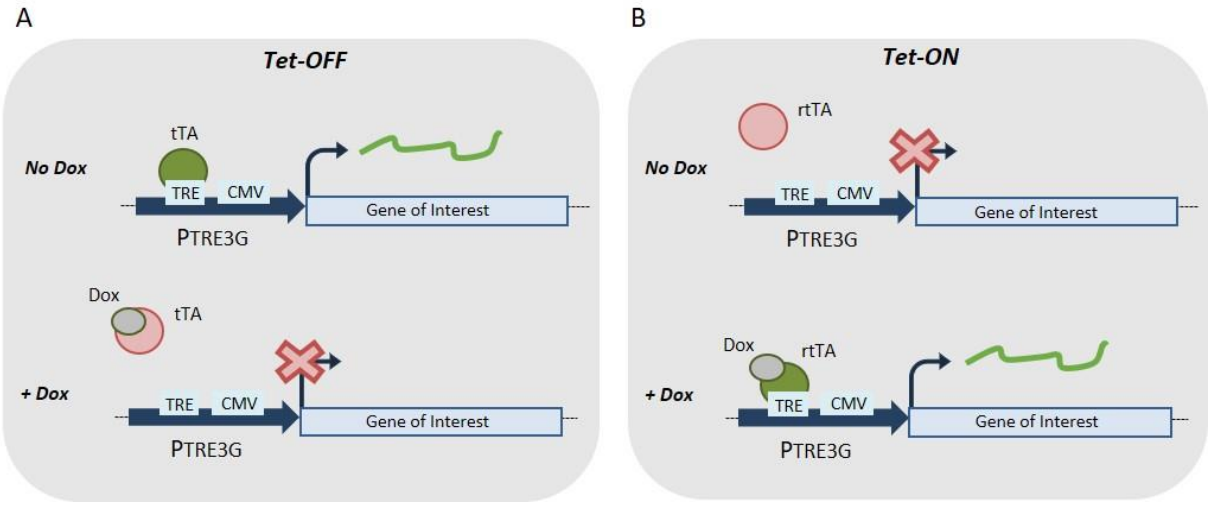
2.2 MEFs

Mouse Embryonic Fibroblasts (MEFs) cells were derived from C57BL/6J E13.5 mouse embryos and cultured in DMEM (Gibco, Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (FBS) and 1% PenStrep (Gibco, Thermo Fisher Scientific). Cells were cultured at 37ºC supplemented with 5% CO_2. All cell lines have been tested for mycoplasma contamination.

2.3 Transitory transfections

MDA-MB-231 cells were transfected with TALE-VP64-T2A-GFP, TALE-TET1-CD-T2A-GFP (WT and mut), TALE-AID-T2A-GFP and dCas9-TET1-CD-T2A-GFP (WT and mut) plasmids using the Trans-IT BrCa transfection reagent (Mirus Bio) according to the manufacturer's instruction. In brief, 60-80% confluent cells in 100 mm dishes were transfected with 15µg of plasmid. As a control, we also transfected cells with a plasmid coding for GFP only (pEGFP-C1, Clontech). GFP-positive and negative cells were sorted 48h after transfection for DNA methylation and gene expression analysis using a BD FACS Vantage cell Sorter (BD Biosciences). All transfections were performed in triplicate.

MEF cells were transfected with dCas9-hTET1 (WT and mut) or dCas9-mTET1 (WT and mut) plasmids using Polyethylenimine (PEI) transfection reagent. A 2mg/mL PEI stock solution was prepared by dissolving PEI 40000 MAX (Polysciences, MW 40000) in DNAse/RNAse free water. The pH of the solution was adjusted to 7.0 with NaOH and the solution was filtered (0,22µm). In brief, 10 µg of plasmid and 20µL of PEI were diluted in 250 µl of 150mM NaCl each, and then combined. The DNA/PEI mixtures were incubated for 30 min at RT. The complexes were added to 70-80% confluent MEF cells in 100-mm dishes. The medium was changed 4 or 6 hours later. 24 hours after transfection the cells were selected with 3µg/µL of puromycine (Gibco, Thermo Fisher Scientific) for 48h. Non-transfected cells were used as control of transfection. All transfections were performed in triplicate.



Methods Figure 6. Schematic representation of the conditional expression systems Tet-ON and Tet-OFF.

A. Tet-OFF system. In the absence of dox the tTA binds the Tet Response Element (TRE) ant triggers the activation of the gene of interest. In the presence of dox the tTA changes its conformation and it is not able to bind the TRE anymore.
B. Tet-ON system. The tTA protein is mutated in this system (rtTA) and presents the opposite binding properties. In the absence of dox the rtTA does not bind the TRE, whereas the presence of dox changes its conformation to bind the TRE sequence and activate expression of the gene of interest.

3 Generation of stable inducible MDA-MB-231 cell lines

3.1 Presentation of the Tet-ON Tet-OFF system

Inducible systems allow the precise spatiotemporal control of gene expression at specific times or in a specific tissue. The most used inducible gene expression systems depend either on Cre recombinase or on exogenous ligands like tetracycline or doxycycline. In the latter case, a gene of interest is placed under the control of a promoter inducible by the presence of an exogenous ligand such as tetracycline or doxycycline (a synthetic tetracycline derivative).

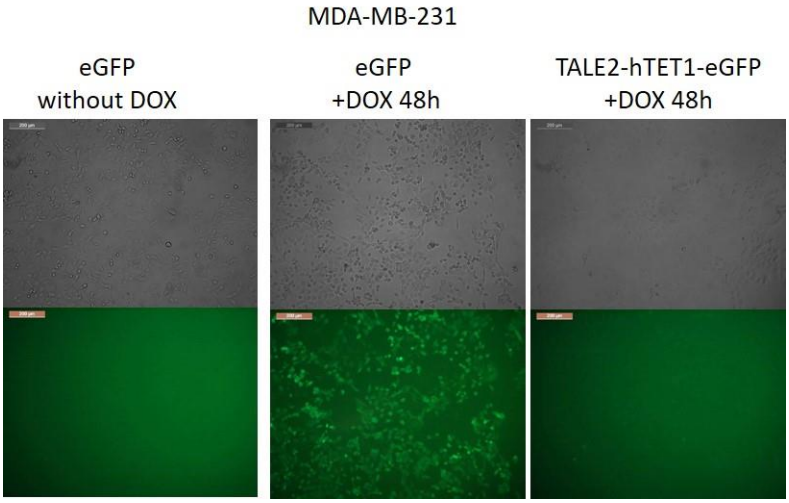
Tet-ON and Tet-OFF systems are the most commonly used conditional expression systems in mammalian cells. Drug treatment can be used to either activate or repress gene expression respectively, always in a reversible manner. These systems are based on the tetracycline resistance operon encoded in *E. coli*.

The Tet-OFF system is based on a fusion protein called tetracycline-controlled transactivator (tTA) (Gossen and Bujard, 1992). tTA is composed by the tetracycline repressor from *E. coli* (TetR) and the transactivator domain VP16 and it is constitutively expressed in the cells. tTA recognises and binds the tetracycline response elements (TRE). In this way, it activates the expression of a gene of interest under the control of a minimal CMV promoter combined with the TRE. In the presence of tetracycline or doxycycline, tTA loses its binding abilities to the TRE, and the expression of the gene of interest is then inactivated (**Methods Figure 6A**).

In the Tet-ON system, the TetR is mutated so the DNA binding properties are reversed compared to the tTA (Gossen et al., 1995). The fusion protein resulting is called reverse tetracycline-controlled transactivator (rtTA) and it is constitutively expressed in cells. In the absence of doxycycline, rtTA is not able to bind the TRE, so the gene of interest is repressed. In the presence of doxycycline, rtTA changes its conformation and binds to the TRE, activating the minimal promoter and the gene expression as a consequence (**Methods Figure 6B**).

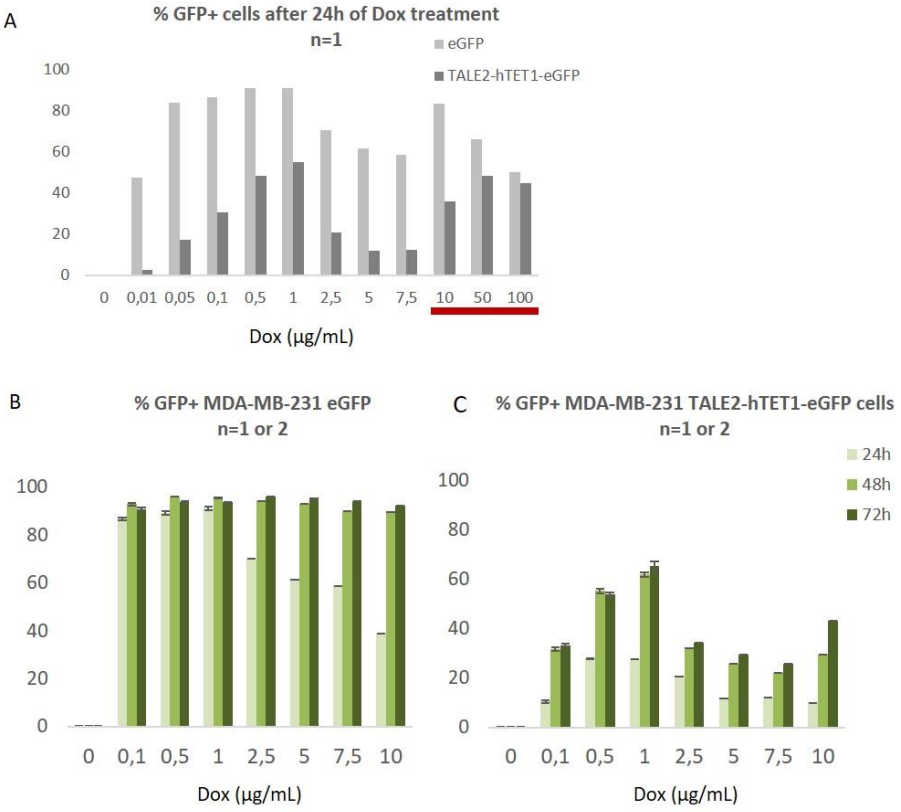
Both systems present some limits. In both cases we have to take into account the stability of the transactivator protein. It is important to choose a promoter constitutively and efficiently active in our cells. The rapidity and the efficiency of the induction are also important parameters. If we are working with a polyclonal population, the gene integration will be different in each cell and the induction can vary from one cell to another.

Concerning the Tet-ON system, one of the main problems is the background expression of the promoter in absence of doxycycline. As the transactivator protein is constitutively expressed in the



Methods Figure 7. Validation of MDA-MB-231 stable inducible cell lines.

Cells were treated with doxycycline (10 µg/mL) during 48h and the GFP signal was analysed under a fluorescence microscope.



Methods Figure 8. Doxycycline induction in MDA-MB-231 cells.

A. MDA-MB-231 cells expressing inducible EGFP or TALE2-TET1-CD-IRES-GFP were treated for 24h with increasing Dox concentrations ranging from 0 to 100 µg/mL, and the percentage of GFP-positive cells was quantified by flow cytometry. Concentrations abve 10 µg/mL were toxic for the cells, which is indicated by the red bar.
B-C. MDA-MB-231 cells expressing inducible EGFP (**B**) or TALE2-TET1-CD-IRES-GFP (**C**) were treated with Dox concentrations ranging from 0 to 10 µg/mL, and the pourcentage of GFP-positive cells was analysed 24, 48 and 72h after induction.

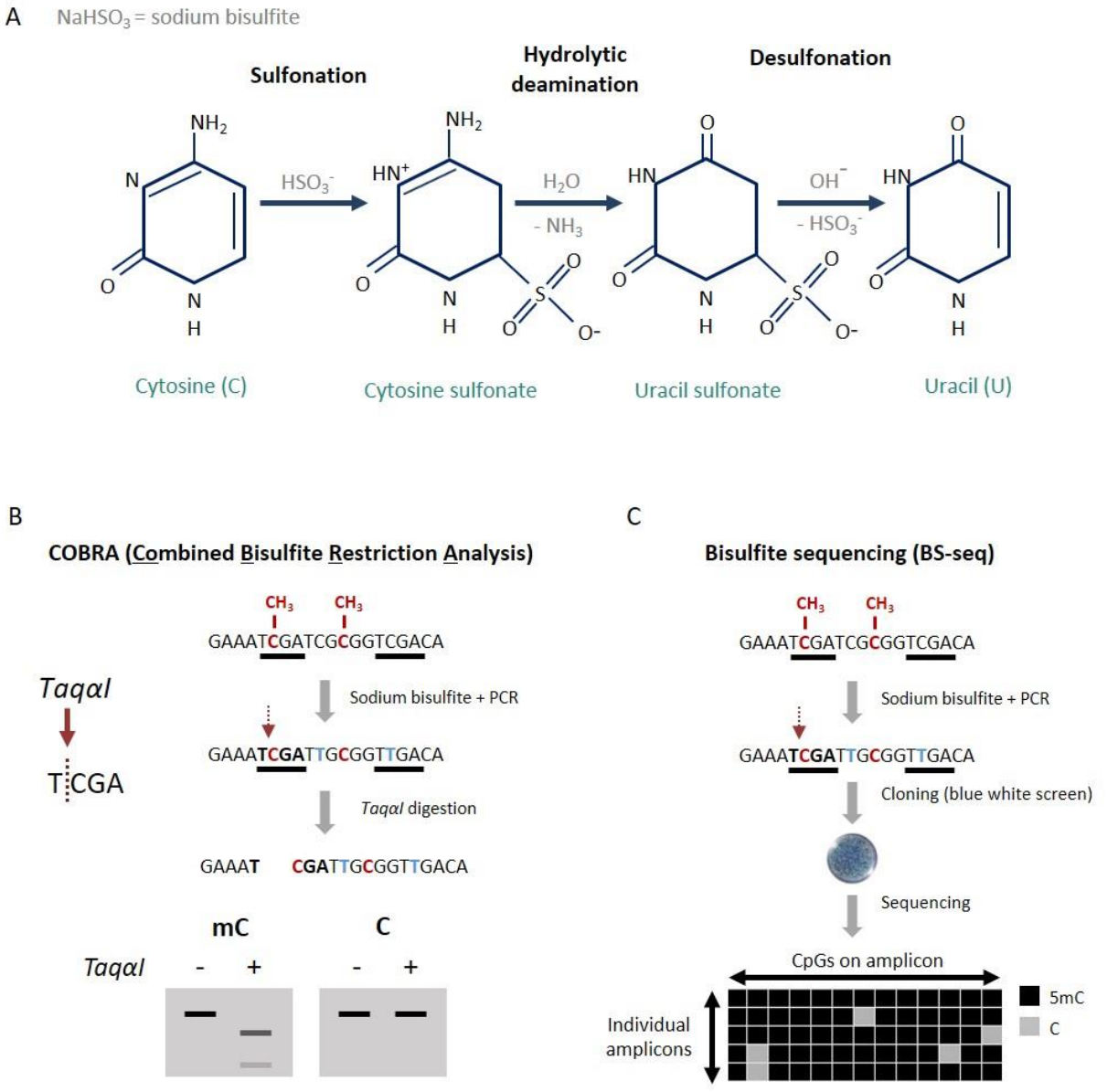
cells, it can have some residual binding activity to the TRE. In this purpose, there have been several improvements of the system in the last years. Zhou et al., 2006 used HIV-rtTA virus variants capable of acquiring random mutations in the rtTA gene. Some of these mutations reinforced the transcriptional activation (up to 7-fold) and the doxycycline sensitivity (up to 100-fold more) of the rtTA. They also improved the basal expression of the promoter. These are called the Tet-On 3G systems and we used one of them in our project (Clontech).

The Tet-ON 3G promoter (PTRE3G) shows a very low residual activity and responds to very low doxycycline concentrations (beyond toxic levels). It contains 7 repeats of a 19 bp tet resistance operator sequence (TRE sequence) and it is located upstream of a minimal CMV (Citomegalovirus) promoter. In presence of doxycycline, rtTA binds to the TRE and it enhances the expression of the downstream gene. PTRE3G does not contain any endogenous transcription factor binding site.

3.2 Inducible cell lines and Dox treatment

We used the Tet-ON 3G system to create stable MDA-MB-231 cells expressing EGFP or TALE2-TET1-CD-IRES-GFP under the control of a Doxycycline-inducible promoter PTRE3G. The lentiviral vectors have been described in the previous section *1.3*. MDA-MB-231 inducible cells were generated by lentiviral transduction. Production of lentiviral particles in HEK293T cells and transduction of MDA-MB-231 cells were performed at the GIGA-Viral Vectors platform (University of Liège, Belgium). Cells were transduced with the pLVX-Tet3G plasmid and twice with the pLVX-TRE3G-TALE2-TET1-CD-IRES-EGFP or the pLVX-TRE3G-IRES-EGFP (coding for EGFP alone). After transduction, stably infected cells were selected with Puromycin and G418 and kept as a pool. To validate the sensitivity of the cells to doxycycline, the cells were treated with doxycycline (10 µg/mL) during 48h and the GFP signal was analysed under a fluorescence microscope. Pictures of the cells are shown in **Methods Figure 7**.

In order to choose the optimal doxycycline concentration, I treated the cells at different concentrations and we analysed the GFP fluorescence signal by flow cytometry using a FACS Calibur (BD Biosciences). First, we treated the cells for 24h at concentrations ranging from 0 to 100 µg/mL. Concentrations beyond 10 µg/mL were toxic for the cells. The activation of control GFP cells was efficient at concentrations between 0,05 and 1 µg/mL, whereas the activation of TALE2-hTET1 cells was only efficient at a concentration of 0,5 and 1 µg/mL (**Methods Figure 8A**). We then treated the cells with dox concentrations from 0 to 10 µg/mL and we analysed the eGFP signal at 24h, 48h or 72h after induction (**Methods Figure 8B and 8C**). Based on the results, we chose the concentration 1 µg/mL for further experiments because it was the lowest concentration having an efficient induction and it did not have any toxic effects in the cells.



Methods Figure 9. Sodium bisulfite DNA conversion for DNA methylation analysis.

A. Sodium bisulfite conversion. Non-methylated cytosines are converted in uracil after a deamination reaction.
B. Combined Bisulfite restriction analysis (COBRA). After DNA bisulfite conversion, the region of interest is PCR amplified and digested with the restriction enzyme *Taq*αI (T^CGA). If the TCGA site was methylated, methylated cytosines will remain as C after bisulfite conversion and the *Taq*αI restriction enzyme is able to cut them. If the TCGA site was not methylated, non-methylated cytosines are converted in thymines after bisulfite conversion, so the restriction site disappears.
C. Bisulfite sequencing (BS-seq). After DNA bisulfite conversion and PCR amplification, the PCR product is cloned and several clones are sequenced by Sanger sequencing and compared to the non-converted sequence. In the results figure, each line represents the sequence of a single clone, and each square in this line represents an individual CpG methylated (black) or unmethylated (grey).

Cells were maintained in culture with a working concentration of antibiotics (0,25 µg/mL Puromycin and 300 µg/mL G418) (Invivogen). For the induction experiments, cells were treated with selection concentrations of antibiotics (0,5 µg/mL Puromycine and 800 µg/mL G418) during 3 days prior to doxycycline treatment. Cells were treated with Doxycycline (1 µg/mL, Clontech) for up to 14 days. The medium and Doxycycline were renewed every day. Non-treated cells were cultured in parallel to serve as controls. Doxycycline was then removed from the medium and the cells were cultured up to 30 days after Dox removal. The induction was followed at several time points during and after treatment by monitoring GFP fluorescence by flow cytometry using a FACS Calibur (BD Biosciences).

4 T7 endonuclease assays.

Genomic DNA was extracted from MDA-MB-231 cells transfected with plasmids coding for Cas9 and sgRNA 3 and 4 targeting the *SERPINB5* promoter. The regions flanking the sgRNA target site in the *SERPINB5* promoter were amplified by PCR with the Platinum Taq DNA polymerase (Invitrogen, Thermo Fisher Scientific) from 2 ng of DNA using the following PCR conditions: 20 cycles of 30s at 95°C, 30s at 64–54°C (with a 0.5°C decrease per cycle), 50s at 72°C followed by 35 cycles of 30s at 95°C, 30s at 54°C, 50s at 72°C. 200ng of the purified PCR products were denatured-annealed following the reaction: 95°C 5 min; -2°C/s to 85°C; -0,1°C/s to 25°C; 12°C. The products were digested with 10U of T7 endonuclease I for 20 min at 37°C. The endo I was inactivated with 25mM EDTA and the products were run in a 1,5% agarose gel. The primers used for the T7 endonuclease assays are given in the **Methods Table 2**.

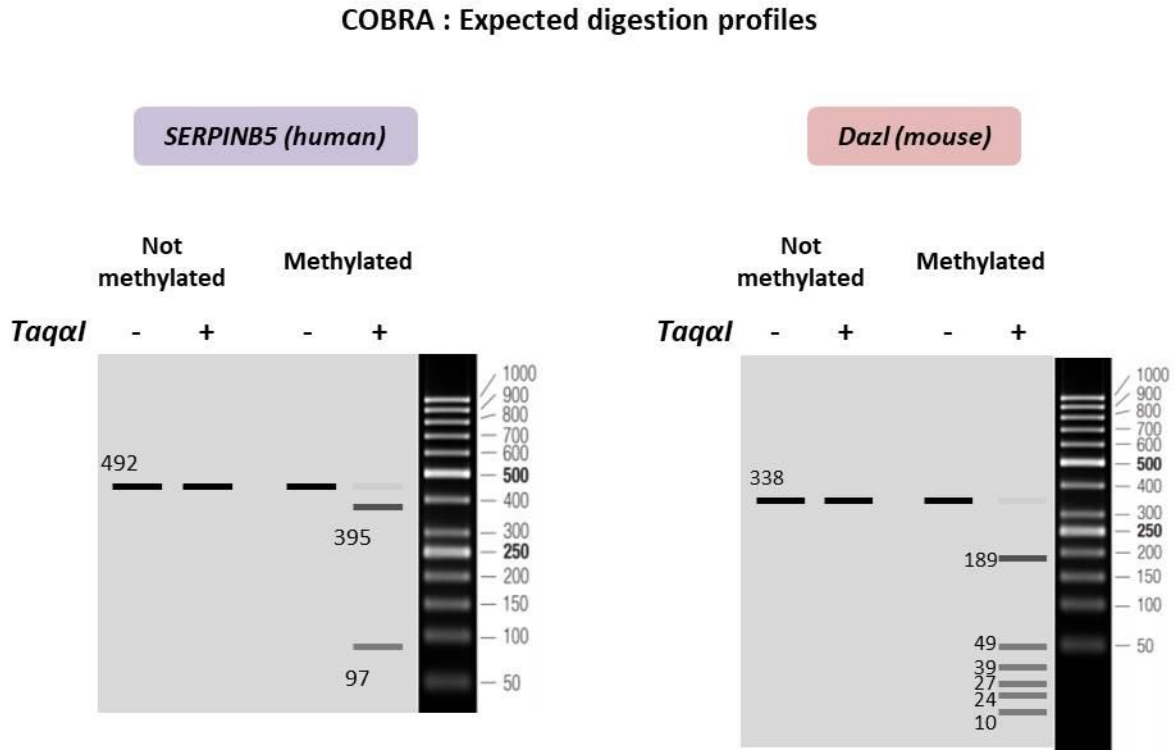
5 Locus specific methylation analysis

We analysed the methylation status of specific regions on the genome by *locus* specific techniques that are based on the conversion of DNA with sodium bisulfite.

5.1 DNA conversion with sodium bisulfite

Sodium bisulfite conversion of DNA allows to differentiate methylated and unmethylated cytosines. It is the gold standard technique for the study of DNA methylation. Sodium bisulfite deaminates unmethylated cytosines to uracil, whereas methylated or hydroxymethylated cytosines (5mC and 5hmC) remain unchanged (**Methods Figure 9A**). Conversion can be followed either by PCR amplification or high throughput sequencing methods to study single cytosines methylation status in a specific locus or in the whole genome.

In brief, unmethylated cytosines are deaminated to uracils during the conversion. The uracils are amplified as thymines in the PCR, whereas 5mC (or 5hmC) are amplified as cytosines. In this way



Methods Figure 10. Expected COBRA gel profiles of the amplified fragments of *SERPINB5* and *Dazl* after TaqαI digestion.

we can compare a bisulfite treated sample to the reference genome and obtain cytosine methylation status patterns. One of the limits of the technique is that bisulfite conversion does not discriminate between methylated or hydroxymethylated cytosines.

In my project I used the EpiTect Bisulfite Kit (Qiagen) according to the manufacturer's instructions to perform all DNA bisulfite conversions.

5.2 COBRA

COBRA (Combined Bisulfite Restriction Analysis) is a quantitative technique allowing to study DNA methylation status at a single locus (Xiong and Laird, 1997). After DNA bisulfite conversion and PCR amplification of the target region, a restriction enzyme digestion is performed. The enzyme used is *TaqαI*, which has a relatively frequent site in CG-rich regions (T^CGA). Methylated cytosines will remain as C, so *TaqαI* will be able to cut them. Non-methylated cytosines will be converted in thymines after bisulfite conversion and PCR amplification, so the restriction site will disappear. The relative amounts of digested and undigested product are analysed by electrophoresis on an agarose gel (**Methods Figure 9B**).

First part of the thesis

Genomic DNA was extracted from the MDA-MB-231 cells with the AllPrep DNA/RNA Mini Kit (Qiagen). Between 200 and 1000 ng of genomic DNA were bisulfite converted using the EpiTect Bisulfite Kit (Qiagen) according to the manufacturer's instructions. For COBRA, the *SERPINB5* promoter (human) region was amplified by PCR with the Platinum Taq DNA Polymerase (Invitrogen, Thermo Fisher Scientific) using the following conditions: 20 cycles of 30s at 95°C, 30s at 62–52°C (with a 0.5°C decrease per cycle), 50s at 72°C followed by 35 cycles of 30s at 95°C, 30s at 52°C, 50s at 72°C. The amplicon was 492 bp long and had 16 CpG. The CpG number 2 (CpG-2) was contained in the T^CGA site (**Methods Figure 2A**). 40 ng of PCR amplicon was digested with 10U *TaqαI* restriction enzyme (Thermo Fisher Scientific) for 1h at 65°C, and an equal amount of PCR amplicon was used for the undigested control. The samples were then purified with the QIAquick PCR purification kit (Qiagen) and migrated on an 2.5% agarose gel. The expected digeston profile with *TaqαI* for *SERPINB5* amplicon is represented in **Methods Figure 10**. Primers to amplify human *SERPINB5* from bisulfited DNA are shown in **Methods Table 3**.

Second part of the thesis

Genomic DNA was extracted from the MEF cells with the AllPrep DNA/RNA Mini Kit (Qiagen). Between 100 and 200 ng of genomic DNA were bisulfite converted using the EpiTect Bisulfite Kit

(Qiagen) according to the manufacturer's instructions. *Dazl* promoter (mouse) region was amplified by PCR with the Platinum Taq DNA polymerase (Invitrogen, Thermo Fisher Scientific) using the following conditions: 20 cycles of 30s at 95°C, 30s at 58–48°C (with a 0.5°C decrease per cycle), 50s at 72°C followed by 35 cycles of 30s at 95°C, 30s at 48°C, 50s at 72°C. The amplicon was 338 bp long and had 5 sites TCGA (**Methods Figure 4D**). The expected digeston profile with *TaqαI* for *Dazl* amplicon is represented in **Methods Figure 10**. Primers used to amplifications from bisulfited DNA are shown in **Methods Table 5**.

5.3 Bisulfite sequencing (*mDaz* and *mAsz1l*)

Sanger bisulfite sequencing is a quantitative technique that allows to study the methylation status of the entire amplicon at the CpG level (Frommer et al., 1992). The first steps of bisulfite sequencing (BS-seq) are the same as COBRA: genomic DNA is converted with sodium bisulfite conversion and the region of interest is amplified by PCR. *Dazl* and *Asz1* promoter regions were amplified as described above (Tm for *Asz1* 62°C-52°C, with a 0.5°C decrease per cycle). The amplicon for *Dazl* was 338 bp long and had 32 CpG and for *Asz1* 612 bp long with 29 CpGs (**Methods Figure 4D and E**). We then cloned the bisulfited PCR products in the pCR2.1 vector (TA Cloning Kit, Invitrogen) and transformed *E. coli* DH10B (NEB 10-beta bacteria from NEB, φ80dlacZΔM15). Positive colonies containing the insert were identified by blue/white screening. Plasmid DNA was purified from the colonies and sequenced by the GATC sequencing platform. Each colony gives us the information of an individual cell methylation profile. Sequences were aligned with the BISMA software (Rohde et al., 2010) and filtered to remove clonal biases. We compared the obtained sequences to the non-converted reference sequence and we established the methylation status. The results are given in the form of **Methods Figure 9C**, where each line represents the sequence of a single clone, and each square in this line represents an individual CpG methylated (black) or unmethylated (grey).

5.4 High-throughput bisulfite amplicon sequencing (*hSERPINB5*)

Bisulfite converted DNA was amplified with two pairs of primers in the *SERPINB5* promoter using the same PCR conditions as for COBRA. The first pair amplifies a 233 bp fragment containing 9 CpG, whereas the second amplicon is 237 bp long and contains 13 CpG. Hence, we can study 22 CpGs close to the TALE2 target region (**Methods Figure 2A**). The forward primers were barcoded at the 5' end with 4 nucleotides to allow multiplexing of samples in the sequencing runs. The primers used for bisulfite amplicons are given in the **Methods Table 3**. The amplicons were purified using AMPure XP beads (Beckman Coulter), pooled and sequencing libraries were prepared and sequenced (2 x 125 bp) on a HiSeq 4000 by GATC Biotech. We obtained a sequencing depth between 10,000 and 50,000X for each experimental sample. Reads were trimmed with TrimGalore (v0.4.2) to remove adapter

sequences and low-quality ends with a Phred score below 20, and aligned to the expected amplicon sequences with BSMAP (v2.74) allowing a maximum of 4 mismatches. We extracted methylation scores as the ratio of the number of Cs over the total number of Cs and Ts and combined CpG methylation ratios from both strands. For single allele methylation analysis, we mapped the reads with Bismark because it returns methylation information on single reads.

6 RT-qPCR assays

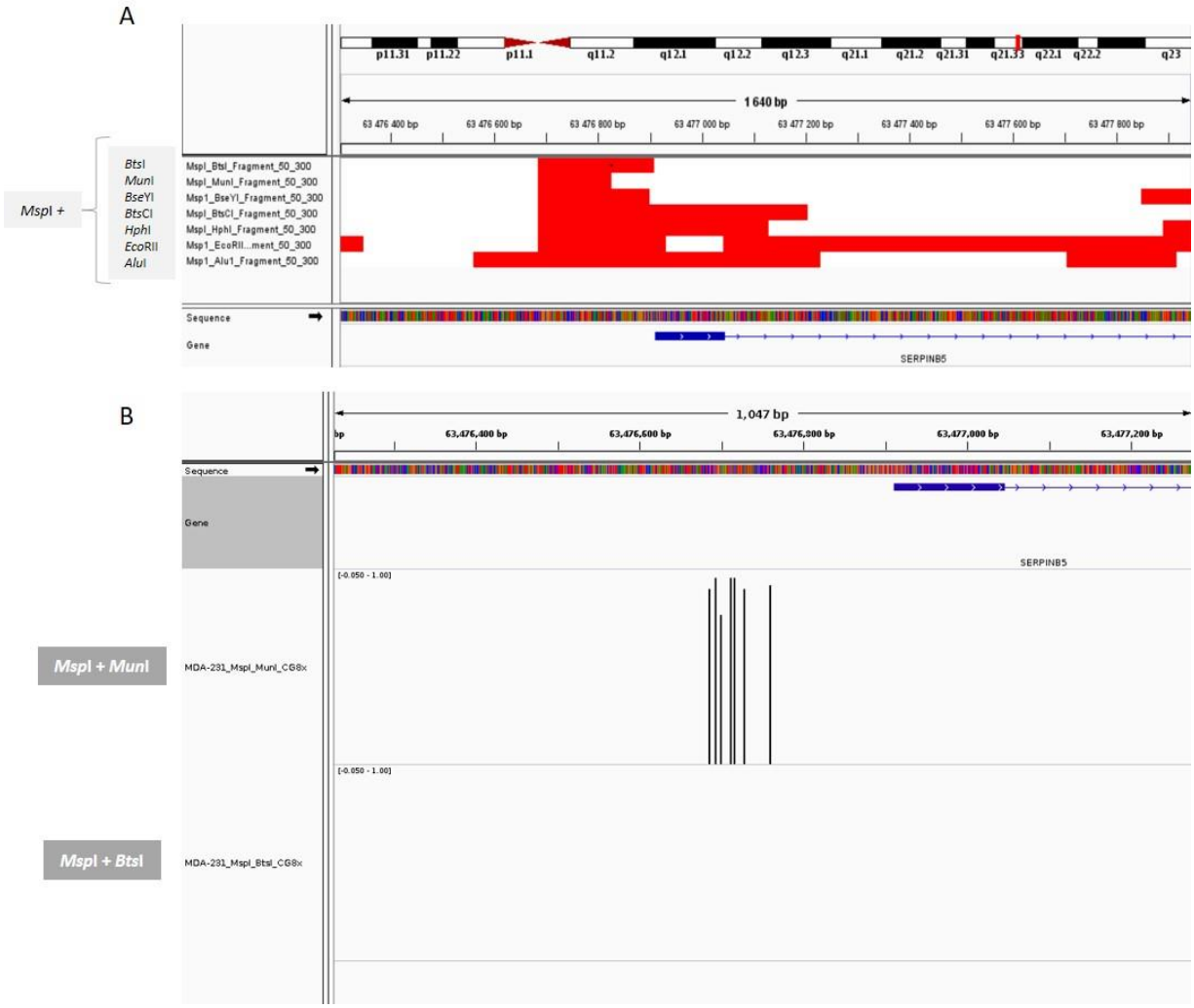
Total RNA was extracted together with DNA samples with the AllPrep DNA/RNA Mini Kit (Qiagen). RNA samples were reverse transcribed with the Maxima First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). Quantitative PCR reactions were performed with the KAPA SYBR FAST kit (Clinisciences) on a StepOnePlus Real-Time PCR System (Life Technologies, Thermo Fisher Scientific). Amplification cycling conditions were as follows: 95 °C for 3 min and 40 cycles (95 °C for 20 sec, 60 °C for 30 s). All samples were assayed in triplicate reactions.

The mRNA expression of *SERPINB5* in MDA-MB-231 cells was normalised in comparison to the housekeeping genes *ACTB*, *RPL13A* and *SDHA* and compared to the non-transfected or non-treated cells. Expression of other tumor suppressor genes (*DAPK1* and *CDH1*) was estimated under the same conditions. We systematically amplified in parallel no-RT controls to exclude the presence of contaminating genomic DNA. Primer sequences for all genes are given in the **Methods Table 3**.

In the 5aza treatment experiments in mouse cells, the mRNA expressions of *Dazl* and *Asz1* in MEF cells were normalised in comparison to combinations of the housekeeping genes *Hprt*, *Rpl13a*, *Mrpl32* and *Gus* and compared to the non-treated cells. For the transfections, the housekeeping genes used for normalisation were *Rpl13a*, *Gus* and *B2m* and the fold-changed was established according to the cells transfected with the plasmid without gRNA. We systematically amplified in parallel no-RT controls to exclude the presence of contaminating genomic DNA. Primer sequences for all genes are given in the **Methods Table 5**.

7 Western blot

Total proteins were extracted with the RIPA Lysis Buffer System (Clinisciences) supplemented with Halt Protease and Phosphatase Inhibitor Cocktail (Thermo Fisher Scientific). Proteins were run on 10% SDS-polyacrylamide gels. The proteins were transferred to a 0,2 µm Nitrocellulose membrane (BioRad) using the Trans-Blot Turbo Blotting System (BioRad). The membranes were washed in Tris-buffered saline solution (TBS 10X, Euromedex) with 0,05% Tween 20 and blocked with 5% nonfat milk in TBS 1X solution with Tween 20 for 1 h. The membranes were incubated with primary antibodies overnight at 4°C (SERPINB5) or 1 h at RT (VINCULIN), then washed and incubated for 1h



Methods Figure 11. RRBS double enzyme digestion test.

A. Predicted restriction fragments in the *SERPINB5* promoter after *in silico* double digestion with different couples of enzymes.
B. Screenshots of RRBS results over the *SERPINB5* promoter after performing double digestion RRBS protocols in MDA-MB-231 cells. The *SERPINB5* promoter is not covered after *Msp*I + *Bts*I double digestion, whereas double digestion with *Msp*I + *Mun*I covers 7 CpGs on *SERPINB5* promoter as expected.

with horseradish peroxidase-conjugated secondary antibodies followed by chemiluminescence detection using the ECL detection reagent (Amersham, GE Healthcare). SERPINB5 was detected with a mouse monoclonal antibody (BD Biosciences #554292, 1:1000) and a secondary sheep anti-mouse antibody (GE Healthcare Lifescience #RPN4201, 1:30000). Vinculin was used as a loading control and detected with a rabbit anti-vinculin antibody (Abcam #ab129002, 1:10000) and a secondary goat anti-rabbit antibody (Sigma #A0545, 1:50000).

8 Global methylation assay: RRBS

RRBS (Reduced Representation Bisulfite Sequencing) is a bisulfite-based technique that profiles DNA methylation of DNA enriched in CG-rich regions of the genome, that is promoters and other relevant regions (Gu et al., 2011a; Meissner et al., 2005). In brief, genomic DNA is digested with the *Msp*I methylation insensitive restriction enzyme (C^CGG). The resulting fragments are ligated to methylated Illumina adaptors and size selected by gel electrophoresis to enrich for small, CpG-rich fragments (40-220 bp). Then, the fragments are bisulfite converted and amplified. High throughput sequencing is performed with an Illumina HiSeq, and the sequencing reads are aligned to the reference genome. The methylation score is estimated for each cytosine by calculating the ratio between the number of sequenced reads containing a C at one specific position and the total number of reads at this position. RRBS is a highly sensitive technique and allows us to quantify the methylation status of around 2.5 million CpG sites in the human genome, covering 10% of the genome and all CpG islands.

One of the limits of the RRBS technique is the use of the restriction enzyme *Msp*I, which is used to enrich for CpG-rich sequences. Although this restriction site is present in most of the CpG islands, some do not contain two adjacent *Msp*I sites and therefore they are lost in the selection procedure. This is the case with the human *SERPINB5* promoter that is not covered by the classical RRBS protocol. To overcome this problem, we tested a double digestion protocol with two different restriction enzymes.

8.1 Double digestion tests

It has been shown that a double-enzyme protocol for RRBS increases the CpG coverage compared to the single-enzyme method (Wang et al., 2013). We performed *in silico* simulation of double enzyme digestion on the human genome (hg38). We selected methylation insensitive enzymes whose restriction sites are present in the *SERPINB5* promoter sequence and that in combination with *Msp*I would result in a restriction fragment between 40-220 bp. This identified 7 possible restriction enzymes: *Alu*I, *Bse*YI, *Bts*CI, *Bts*I, *Eco*RII, *Hph*I and *Mun*I (**Methods Figure 10A**).

Digestion with MspI alone generates 750,000 predicted restriction fragments between 40 and 220bp. Since we wanted a combination of enzymes that results in less than 2 million fragments (to stay in the same range of sequencing), we choose to test the combinations MspI + BtsI (5' - GCAGTG_NN^ - 3') (2 million fragments) and MspI + MunI (5' - C^AATT_G - 3') (900,000 fragments). I adjusted the first steps of the RRBS protocol to include these double digestions, and I prepared two test libraries with MDA-MB-231 genomic DNA in order to validate both approaches.

Since *Bts*I is active at a different temperature and produces DNA overhangs in the opposite strand orientation compared to *Msp*I, we first performed the digestion with *Bts*I and we treated with the klenow to eliminate the 3'→5' ends of *Bts*I digestion. Then we digested with *Msp*I and we performed the end repair and A-tailing with klenow exo- prior to adaptor ligation. However, when we analysed the sequencing results of the *Msp*I + *Bts*I test library, we were not able to sequence the *SERPINB5* promoter, suggesting that the end repair step after *BtsI* digestion did work work efficiently (**Methods Figure 10B**).

In contrast, *Msp*I and *Mun*I produce DNA overhangs in the same strand orientation, therefore they can be used in combination without the necessity to modify the end repair step of the protocol. We just needed to add the nucleotide dTTP (1mM) during the end repair step. After sequencing of the *Msp*I + *Mun*I test library, we obtained methylation information for 7 CpGs in the *SERPINB5* promoter, as predicted by the in sillico restriction digestion (**Methods Figure 10B**). We decided to choose the double restriction enzyme digestion with *Msp*I and *Mun*I for further experiments.

8.2 Preparation of RRBS libraries

We prepared RRBS libraries from non-transfected cells and cells transfected with TALE2-hTET1-CD (wt or mutant) and selected by FACS 48h later. RRBS libraries were prepared according to a published protocol (Gu et al., 2011a) with modifications.

We performed RRBS in three biological replicates of non-transfected cells and cells transfected with TALE2-TET1-CD-WT or TALE2-TET1-CD-mut. RRBS libraries were prepared as described (Auclair et al., 2014), with some modifications. Briefly, 50 ng genomic DNA were digested for 5h with *Msp*I and *Mun*I (Thermo Fisher Scientific) followed by end-repair, A-tailing and ligation to methylated adapters with T4 DNA ligase (Thermo Fisher Scientific) in Tango 1X buffer. Small fragments (150 to 400 bp) were selected by electrophoresis on a 3% (w/v) agarose 0.5X TBE gel and purified with the MinElute gel extraction kit (Qiagen). Bisulfite conversion was performed by two rounds of conversion with the EpiTect kit (Qiagen) according to the manufacturer's instructions. RRBS libraries were amplified by 14 cycles of PCR with PfU Turbo Cx hotstart DNA polymerase (Agilent) and purified with AMPure magnetic beads (Beckman Coulter). The libraries were paired-end sequenced (2x75 bp) with

a HiSeq4000 by Integragen SA (Evry, France). Sequencing reads were trimmed with TrimGalore (v0.4.2) and aligned to the human hg38 genome with BSMAP (v2.74). For the mapping, we allowed a maximum of two mismatches. We combined CpG methylation ratios from both strands and only retained CpGs with a minimum sequencing depth of 8X. To compare methylation profiles between samples, we generated pairwise plots with methylation values calculated in 400 bp windows containing at least 3 CpGs. The dendrogram was generated based on the Euclidean distances between CpG methylation scores with the dist and hclust functions in R. To identify DMRs, we used the eDMR algorithm from the methylKit R package (Li et al., 2013). We identified DMRs with a change in methylation >15%, at least 4 differentially methylated CpGs (DMCs), and a p-value <0.01 (t-test) between the triplicates. DMRs were considered promoter proximal if they are located -500 to +500 bp from an annotated RefSeq TSS.

9 Growth and Apoptosis assays

We assessed proliferation and apoptosis in the doxycycline inducible cell lines. To quantitatively study the percentage of cell death (early apoptotic, late apoptotic or necrotic cells) we performed Annexin-V PE and 7-AAD staining using commercial kits (BD Bioscience) according to the manufacturer's instructions. The percentages of apoptotic cells were measured by flow cytometry using a FACSCelesta flow cytometer (BD Biosciences) 7, 10 and 14 days after the beginning of the Dox treatment, and 10 days after Dox removal. A positive control was induced at each time point with campthotecin (Sigma) at 8µM for 4h at 37ºC to determine the flow cytometry settings. To measure cell proliferation, Dox induced cells (9 days after the beginning of the treatment) were seeded in 96-well plates (2,000 cells per well) and cultured for 1 to 5 days. The medium containing Dox was refreshed every day. The cells were incubated with the WST-1 Cell Proliferation Reagent (Roche) for 4h at 37°C, and then the absorption was measured at 440 nm using the Asys UVM 340 Microplate Reader (Biochrom). All experiments were performed in biological triplicate.

Primers directed mutagenesis hTET1-CD and mTET1-CD	
hTET1 Directed mutagenesis Forward	5' – GACTTCTGTGCTCATCCC**TAC**AGG**GCC**ATTCACAACATGAATAA – 3
hTET1 Directed mutagenesis Reverse	5' – TTATTCATGTTGTGAAT**GGC**CCT**GTA**GGGATGAGCACAGAAGTC – 3
mTET1 Directed mutagenesis Forward	5' – GGATTTTTGTGCCCATTCT**TAC**AAG**GCC**ATTCACAACATGCACAAC – 3'
mTET1 Directed mutagenesis Reverse	5' – GTTGTGCATGTTGTGAAT**GGC**CTT**GTA**AGAATGGGCACAAAAATCC – 3'
Primers directed mutagenesis *Bbs*I sites in hTET1-CD	
BbsI site 1 - Forward	5' – GTCCCC**GAAAAC**TGCTTCAGCCACACCAGCTCCACTGAA – 3'
BbsI site 1 - Reverse	5' – GCTGAAGCA**GTTTTC**GGGGACCAGGAGAAGCCTGGAGATG – 3'
BbsI site 2 - Forward	5' – CTTCGG**GAAGGC**TCAGTGGTGCCAATGCAGCTGCTGCTGAT – 3'
BbsI site 2 - Reverse	5' – CACCACTGA**GCCTTC**CCGAAGGCATCGTACAGTGGGGAGTG – 3'

Methods Table 1. Primers used for directed mutagenesis.

TALE sequences targeting human *SERPINB5* promoter	
TALE1 binding sequence	TAGCAGAATGAGCTGCTGCAGT
TALE2 binding sequence	TTGGAAGCTGTGCAGACAACAG
sgRNA targeting human *SERPINB5* promoter (without PAM)	
sgRNA 1	TTCCTGCCCGAACATGTTGG
sgRNA 2	ATTGTGGACAAGCTGCCAAG
sgRNA 3	GTAGGAGAGGAGTGCCGCCG

Methods Table 2. TALE and sgRNA (CRISPR-dCas9) binding sequences in the human *SERPINB5* promoter.

Amplification *SERPINB5* promoter for COBRA and BS-seq (bisulfited DNA)	
SERPINB5 BS-Seq Forward	TTTTGGAAGTTGTGTAGATAATAGTAATTTT
SERPINB5 BS-Seq Reverse	CCCTCTACAATATTCCTAATACTCAAAAA
Amplification *SERPINB5* promoter for amplicon Next Generation Sequencing (bisulfited DNA)	
Forward region-1 (**NNNN** = tag)	NNNNTGTTGTAGTTTATATAAAAAGAATGGAGA
Reverse region-1	ATCTACACAACTTCCAAAAAACCTC
Forward region-2 (**NNNN**= tag)	NNNNAATTGTGGATAAGTTGTTAAGAGGTTTG
Reverse region-2	CTACCCCACCTTACTTACCTAAAATCAC
T7 endonuclease assay	
Primer gSERPINB5 F	CAGGCCTGAGTAATCCTAATCACAG
Primer gSERPINB5 R	GAACACTGCAGAGGGAGTAAAATC
RT-qPCR	
SERPINB5 primer F	GCCACGTTCTGTATGGGAAA
SERPINB5 primer R	TGGACTCATCCTCCACATCC
SDHA primer F	GCTCATGCATGTGTCCATGT
SDHA primer R	CAGCCACTAGGTGCCAATCT
RPL13A primer F	CCGAGAAGAACGTGGAGAAG
RPL13A primer R	GGCAACGCATGAGGAATTA
ACTB primer F	AGGTGTGGTGCCAGATTTTC
ACTB primer R	GGCATGGGTCAGAAGGATTC
DAPK1 primer F	GGCGAGCTGTTTGACTTCTT
DAPK1 primer R	GATTTGAAGGGAGTGCAGGT
CDH1 primer F	CTCGACACCCGATTCAAAGT
CDH1 primer R	GGCGTAGACCAAGAAATGGA

Methods Table 3. Primers sequences used in the first part of the thesis.

sgRNA targeting mouse *Dazl* promoter (without PAM)	
sgRNA 3	GACGTGCTACAGCCAATAGG (+)
sgRNA 4	ACGCACTCCGTGGGCGACGT (-)
sgRNA 5	GCATGCGCGACATCTACGTA (+)
sgRNA 6	AGGGGCTTCGGACACGGTCC (-)
sgRNA targeting mouse *Asz1* promoter (without PAM)	
sgRNA 3	GGAAGAACAGCTTGGAGCCA (-)
sgRNA 4	GTGGCCTTGCGCAAGCGCCT (+)
sgRNA 5	GTGAAAGGCCAGCTCGTGGG (-)
sgRNA 6	GTGGAACGAGCTAAATACGA (+)

Methods Table 4. sgRNAs sequences in the mouse *Dazl* and *Asz1* promoters.

Primers for the cloning of sgRNAs into the pPlatTET-gRNA2 vector (N = 19nt of gRNA sequence) Sequences marked in purple and orange are reverse complement of each other	
pPlatTet-gRNA2 gRNA Insert F	5' – TTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGNNNNNNNNNNNNNNNNNNN -3'
pPlatTet-gRNA2 gRNA Insert R	5'- GACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACNNNNNNNNNNNNNNNNNNNC – 3'
Primers for the amplification of mouse *Dazl* promoter for COBRA and BS-seq (bisulfited DNA)	
Dazl BS-Seq Forward	GTGAGGTGGTTGTTTAGT
Dazl BS-Seq Reverse	ATCCATCTATCTATCAATAAAACTC
Asz1 BS-Seq Forward	TAAGTTAAAGTTTTGGGAAAAAGTTAGG
Asz1 BS-Seq Forward	TTATTTCTAATAAAAATCCATCTCACCC
Primers RT-qPCR (5' → 3')	
Dazl-2 primer F	AACAACCCCCTGAGATGAGTT
Dazl-2 primer R	TTCAGAAGAAGTCGGGCAGT
Dazl-3 primer F	TCAATTCCTCCAACAAAAACG
Dazl-3 primer R	ATCAGCAACCACAAGTCAAGG
Asz1 primer F	CCCGAATCCACACACTGTAAC
Asz1 primer R	CGGGAAAATGACCCAACTC
Mael primer F	CCATTTTGGGTGTCACACTG
Mael primer R	TCTGTTTGGAATCCAGCTCA
Hprt primer F	GCAGTACAGCCCCAAAATGG
Hprt primer R	GGTCCTTTTCACCAGCAAGCT
Rpl13a primer F	GAGTCCGTTGGTCTTGAGGA
Rpl13a primer R	GGTTACGGAAACAGGCAGAA
Mrpl32 primer F	AGGTGCTGGGAGCTGCTACA
Mrpl32 primer R	AAAGCGACTCCAGCTCTGCT
Gus primer F	GATTCAGATATCCGAGGGAAAGG
Gus primer R	GCCAACGGAGCAGGTTGA
B2m primer F	TATGCTATCCAGAAAACCCCTCAA
B2m primer R	GTATGTTCGGCTTCCCATTCTC

Methods Table 5. Primers sequences used in the second part of the thesis.

RESULTS

RESULTS (1)

Objectives

The first part of my thesis is focused on the study of DNA methylation as a therapeutic target. DNA methylation is a reversible modification and is a target of choice in the development of therapeutic strategies in cancer. CpG islands in promoters of tumor suppressor genes are frequently aberrantly methylated and repressed in cancer. Thus, erasing methylation in these regions would allow the re-expression of endogenous genes involved in essential pathways such as control of cell cycle and the induction of apoptosis. Up to now, several DNMTs inhibitors have been developed and tested in clinics. For example, the 5'-azacytidine (5aza) molecule, a cytidine analogue incorporated into DNA and able to *trap* DNMTs in the chromatin, is approved for the use in Acute Myeloid Leukaemia and Myelodysplastic syndromes. However, its use is limited because the loss of methylation induced by these inhibitors is non-specific, leading to a global demethylation of the genome, secondary effects and high toxicity in non-cancer cells.

My objective in this first part was to develop targeted demethylation strategies addressed to tumor suppressor gene promoters without affecting the global methylation profile of the cells. My work was aimed to answer the following questions:

(1) Can we induce a specific demethylation of a tumor suppressor gene promoter in cancer cells?

(2) Is the epigenetic reprogramming durable in time?

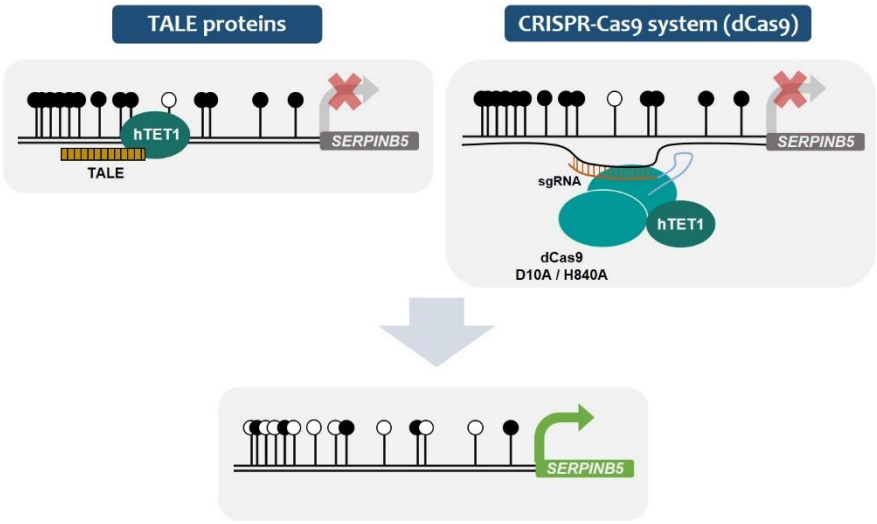
As a model gene to test epigenome editing strategies we have used *SERPINB5* (also known as *MASPIN*), a gene with tumor suppressor properties in epithelial cells and whose promoter has been reported to be hypermethylated in breast cancer (Domann et al., 2000; Futscher et al., 2004; Maass et al., 2002; Oshiro et al., 2003). It is also downregulated in prostate cancer (Abraham et al., 2003; Sheng et al., 1996). Re-expression of *SERPINB5* in *in vivo* or *in vitro* studies has been shown to inhibit tumor growth, cell migration and invasion, metastasis, angiogenesis and promote apoptosis (Latha et al., 2005; Lee et al., 2012b; Seftor et al., 1998; Shi et al., 2007; Zhang et al., 2005). The physiological role of *SERPINB5* as a gene with tumor suppressor properties has been questioned (Teoh et al., 2014), however a more recent study using a novel *Serpinb5* mouse knockout model showed that *Serpinb5* protects from the apparition of cancerous phenotypes in many tissues, confirming its tumor suppressor function (Dzinic et al., 2017).

During my PhD, I have used the breast cancer cell line MDA-MB-231. Once I had validated that *SERPINB5* promoter was methylated on these cells and its expression was sensible to 5aza

treatment, I constructed plasmids coding for engineered chimeric proteins aiming to perform epigenome editing on these cells. In order to induce the targeted demethylation of the *SERPINB5* promoter, I have developed fusion proteins containing a DNA binding domain TALE fused to the catalytic domain of the human TET1 (hTET1), the enzyme catalysing the hydroxylation of methylated cytosines. During the second year of my PhD, the field of genome and epigenome editing was revolutionised with the discovery of a flexible and cheaper alternative to TALEs: the CRISPR-Cas9 system. Using the catalytically inactive version of Cas9 (dCas9), we created the dCas9-hTET1 fusion proteins and we designed sgRNAs targeting *SERPINB5* promoter in order to compare its efficiency to the results we had obtained with the TALE-based system.

As TALE seemed to achieve a more robust demethylation and re-expression, we created an inducible and reversible system coding for TALE-hTET1 under the control of a promoter sensible to doxycycline. This system allowed us to study the kinetics of the demethylation and re-expression in time. Moreover, we were able to study the stability of the demethylation in the absence of the effector domain.

I have demonstrated that the experimental strategy leads to a specific and durable reprogramming of *SERPINB5* gene in breast cancer cells. These results could open the door to research aiming to a therapeutic application.



Objectives Figure 1. Schematic representation of the objectives of the first part of the thesis.

Manuscript:

Argüeso-Lleida A, Dumas M, Cleroux E, Guibert S, Weber M (2018). Durable reactivation of a repressed tumor suppressor gene by targeted DNA demethylation in human cancer cells. *Submitted*.

MANUSCRIPT

Durable reactivation of a repressed tumor suppressor gene by targeted DNA demethylation in human cancer cells

Andrea Argüeso-Lleida[1,2], Michael Dumas[1,2], Elouan Cleroux[1,2], Sylvain Guibert[1,2], Michael Weber[1,2]*

[1] University of Strasbourg

[2] CNRS, UMR7242 Biotechnology and Cell Signaling, 300 Bd Sébastien Brant, 67412 Illkirch Cedex, France

* Correspondence: michael.weber@unistra.fr

Abstract

Aberrant hypermethylation of CpG-rich gene promoters is a hallmark of cancer, contributing to the epigenetic silencing of tumor suppressor genes (TSGs). Targeted DNA demethylation with programmable DNA binding proteins constitutes a promising strategy to induce specific reactivation of TSGs in cancer cells, however it is unknown if these approaches can induce persistent reactivation of TSGs. Here we demonstrate that targeted demethylation induces durable reactivation of the repressed SERPINB5 gene in breast cancer cells. We first compared different targeting tools and found that TALE is more efficient than dCas9 for TET1-mediated demethylation of SERPINB5, with minimal off-targets. Using a doxycycline inducible system, we show that targeted reactivation of SERPIN5 by transient expression of TALE-TET1 is sustained for up to 30 days in culture. Our work highlights the potential of targeted demethylation tools for inducing specific and durable reactivation of TSGs in cancer cells.

Keywords: epigenome editing, *SERPINB5*, TALE, dCas9, TET1

Introduction

In mammals, DNA methylation occurs on cytosines of CpG dinucleotides and plays critical roles in embryonic development and tumorigenesis (Smith and Meissner, 2013). DNA methylation is catalysed by DNA methyltransferases (DNMTs) and reversed by ten-eleven translocation (TET) proteins. TET proteins catalyse the oxidation of 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) and other oxidised derivatives, which induces their replacement with unmodified C by replication-dependent dilution or through the action of DNA glycosylases such as thymine-DNA glycosylase (TDG) and base excision repair (Wu and Zhang, 2017). In normal cells, dense methylation of CpG-rich gene promoters is restricted to a small number of genes, mainly imprinted genes and germline-specific genes (Auclair et al., 2014). In contrast, aberrant CpG island promoter DNA methylation is frequent in all types of cancer cells and contributes to the epigenetic silencing of tumor suppressor genes (TSGs) involved in many cellular processes such as DNA repair, apoptosis, cell cycle control and cell migration.

SERPINB5 (also known as MASPIN) is a member of the serpin family of serine protease inhibitors. Numerous studies have demonstrated that SERPINB5 has tumor suppression functions in epithelial cancers by promoting apoptosis and inhibiting invasion and metastasis of cancer cells (Cher et al., 2003; Seftor et al., 1998; Sheng et al., 1996; Shi et al., 2002; Zhang et al., 1997; Zou et al., 1994). Knockout of *Serpinb5* is embryonically lethal in mice (Gao, 2004), whereas mice with a heterozygous deletion of *Serpinb5* develop prostatic hyperplasia (Shao et al., 2008). Moreover *Serpinb5*-deficient mice that survive until adulthood develop cancerous phenotypes in the mammary gland, prostate, lung, intestine, and skin (Dzinic et al., 2017). In normal tissues, *SERPINB5* expression is inhibited in non-epithelial cells by dense promoter CpG methylation (Futscher et al., 2002). The *SERPINB5* gene is rarely mutated in cancer but instead its expression is frequently repressed by aberrant promoter DNA methylation in breast cancer cells (Domann et al., 2000; Maass et al., 2002).

Because of its reversible nature, DNA methylation is a promising target for anti-cancer therapies. Especially, inhibiting CpG island promoter methylation could restore expression of tumor suppressor genes in cancer. Numerous studies are evaluating the clinical benefits of global DNA methylation inhibitors such as 5-aza-2'-deoxycytidine (5azadC), but their use is limited because of lack of specificity and high toxicity. Recent developments in targeted epigenome editing offer new perspectives to correct DNA methylation at specific genes in cancer cells. These strategies consist in fusing programmable DNA binding modules to epigenetic effector domains to reprogram the epigenome at specific sites (Kungulovski and Jeltsch, 2016). In particular, targeted recruitment of TET enzymes has emerged as an attractive approach to demethylate and reactivate silenced TSGs in cancer cells. Using artificial zinc finger proteins (ZFPs), transcription activator-like effectors (TALEs) or

catalytically inactive Cas9 (dCas9), TET catalytic domains have been targeted to hypermethylated gene promoters in human cells, resulting in successful removal of DNA methylation and gene reactivation (Chen et al., 2014; Choudhury et al., 2016; Huisman et al., 2016; Maeder et al., 2013; Xu et al., 2016). However, these studies did not perform genome-wide methylome analysis to evaluate the specificity of the demethylation. Another point that was not addressed is the durability of TET-induced reactivation of TSGs, which is of crucial importance for potential future therapeutic applications. Only a handful of studies addressed the stability of targeted epigenome editing. These showed that ZFP-mediated recruitment of DNMT3A can induce stable DNA methylation at some but not all genes (Kungulovski et al., 2015; Rivenbark et al., 2012; Stolzenburg et al., 2015). Another study showed that targeted gene activation by artificial ZFPs can be stable for several days in culture only in DNA hypomethylated loci (Cano-Rodriguez et al., 2016).

Here we addressed if targeted demethylation can induce durable reactivation of TSGs in cancer cells, using as a model the methylated *SERPINB5* gene. We first compared different strategies and showed that TALE is more efficient than dCas9 for inducing TET-dependent demethylation and reactivation of *SERPINB5* in breast cancer cells. We then created an inducible system to study the durability of the epigenetic reprogramming, and found that reactivation of SERPINB5 is detectable at the mRNA and protein level for up to 30 days in culture. Our data highlight the potential of targeted demethylation strategies to induce long lasting reactivation of TSGs in cancer cells.

Results

Engineering of TALEs targeting the SERPINB5 promoter

The expression of *SERPINB5* has previously been reported as being silenced by DNA methylation in MDA-MB-231 cells (Primeau et al., 2003; Wozniak et al., 2007). To recapitulate these findings in our experimental settings, we treated MDA-MB-231 cells with the DNA methylation inhibitor 5-aza-2'-deoxycytidine (5azadC) at different concentrations (0.2, 2 or 5µM) during three days. As shown in **Figure S1**, *SERPINB5* expression is strongly upregulated in the cells treated with 5azadC, at higher levels compared to other methylation-repressed tumor suppressor genes such as *DAPK1* and *CDH1*. Thus *SERPINB5* is a good candidate gene to perform targeted demethylation experiments. In order to compare different strategies for targeted demethylation of the *SERPINB5* gene, we first engineered TALE modules targeting different locations of the *SERPINB5* promoter. TALE1 binds a region of 21 nucleotides located at -340 bp relative to the Transcription Start Site (TSS), whereas TALE2 binds a sequence at -111 bp relative to the TSS (**Figure 1A**). To evaluate the efficacy of the TALEs, we fused them to the transcriptional activator VP64 and a green fluorescent protein (GFP) reporter (**Figure 1B**), and transfected MDA-MB-231 cells to assess *SERPINB5* expression

on GFP-positive and GFP-negative cells 48h after transfection. *SERPINB5* mRNA expression was upregulated more than 150 times in cells transfected with TALE2-VP64 compared to untransfected cells or cells transfected with a GFP control vector (**Figure 1C**). In contrast, TALE1-VP64 had only a modest impact on SERPINB5 expression (**Figure 1C**). Based on these results, we selected TALE2 for further experiments.

TALE-mediated targeted demethylation of SERPINB5

Having identified a TALE targeting *SERPINB5*, we fused TALE2 with the catalytic domain of the human TET1 protein (TET1-CD-WT) or a catalytic inactive mutant (TET1-CD-mut) (**Figure 1B**) and tested the effects of both constructs on *SERPINB5* expression and promoter methylation in MDA-MB-231 cells 48h after transfection. TALE2-TET1-CD-WT was able to upregulate *SERPINB5* expression level around 50 times compared to non-transfected cells (**Figure 1D**). In contrast the catalytic mutant TALE2-TET1-CD-mut had no significant impact on *SERPINB5* expression (**Figure 1D**), demonstrating that the upregulation is specifically due to the catalytic activity of TET1. Because previous studies suggested that the Activation-induced cytidine deaminase (AID) protein might be able to trigger DNA demethylation (Ramiro and Barreto, 2015), we also tested the possibility to fuse TALE2 with the human AID protein. We found that it does not efficiently restore *SERPINB5* expression, as TALE2-hAID was able to reactivate *SERPINB5* expression only 5 times compared to control cells (**Figure S2**).

We then performed different analysis to measure DNA methylation in the *SERPINB5* promoter. First, we performed Combined Bisulfite Restriction Analysis (COBRA) to evaluate cytosine methylation in a TCGA site located close to the TSS (**Figure 1A**). Demethylation is observed in MDA-MB-231 cells expressing TALE2-TET1-CD-WT, but not in cells expressing TALE2-VP64 and untransfected cells (**Figure 1E**). The fact that TALE2-VP64 does not impact DNA methylation indicates that demethylation is not just a consequence of transcriptional activation. To achieve a quantitative and accurate quantification of methylation, we performed Next Generation Sequencing (NGS) of two bisulfite PCR amplicons in the *SERPINB5* promoter covering 22 CpGs (**Figure 1A** and **Figure S3**). This revealed that TALE2-TET1-CD-WT induced a global decrease in methylation of 16% over the amplicons compared to the control cells (**Figure 1F**). A closer inspection of DNA methylation levels at individual CpGs showed that demethylation occurred approximately 100 bp upstream and 200 bp downstream of the TALE2 binding site (**Figure 1G**). The CpGs around the TALE2 binding site show the highest demethylation, in particular CpGs 8 to 11 that lose more than 30% methylation compared to the control cells (**Figure 1G**). Altogether, we show that TALE2-TET1-CD induces efficient demethylation and reactivation of *SERPINB5* in breast cancer cells.

Comparison of TALE and dCas9-mediated demethylation of SERPINB5

Next, we compared the ability of TALE and dCas9 based methods to induce epigenetic reactivation of the *SERPINB5* promoter. We transfected MDA-MB-231 cells with a plasmid coding for the catalytically inactive Cas9 (dCas9) fused to the catalytic domain of TET1 (dCas9-TET1-CD-WT) or the catalytically dead mutant (dCas9-TET1-CD-mut) together with a GFP reporter and a sgRNA targeting the *SERPINB5* promoter (**Figure 2A**). Fluorescence-positive and negative cells were sorted 48h after transfection. We tested independently three different sgRNAs targeting a sequence close to the TALE2 binding site (**Figure 2B**), as well as a condition without sgRNA as control. Only one sgRNA (sgRNA2) was able to induce *SERPINB5* expression compared to control cells, but with a lower magnitude than TALE2-TET1-CD-WT (**Figure 2C**). In contrast, the sgRNA3 and sgRNA4 did not induce a significant derepression of *SERPINB5* expression (**Figure 2C**). This is not due to an absence of targeting activity of these gRNAs because they are able to induce Cas9-mediated DNA repair as tested by T7 endonuclease assay (**Figure S4**). We then evaluated the demethylation induced by the sgRNA2. As shown by the COBRA results in the **Figure 2D** and consistent with the expression results, the sgRNA2 induced only a minor decrease of methylation compared to the TALE2-TET1-CD-WT construct. This is consistent with other studies reporting low efficacy of dCas9-TET1 in mammalian cells (Morita et al., 2016). Taken together, our results indicate that TALE is more efficient than dCas9 for inducing targeted epigenetic reactivation of *SERPINB5* in MDA-MB-231 cells.

Genome-wide specificity of demethylation by TALE2-TET1

Our next goal was to investigate the genome-wide specificity of the demethylation induced by TALE2-TET1. To this end, we performed Reduced Representation Bisulfite Sequencing (RRBS) in cells transfected with TALE2-TET1-CD-WT and compared the methylation profiles to the ones in cells transfected with TALE2-TET1-CD-mut and non-transfected cells. Because the *SERPINB5* promoter sequence does not contain two *Msp*I restriction sites, we adapted the RRBS protocol to perform a double digestion with *Msp*I and *Mun*I (see Methods). We performed RRBS in triplicates, capturing between 2 and 3 million CpGs in each sample. The overall methylation profiles of cells expressing TALE2-TET1-CD-WT were indistinguishable from the ones of the control cells (**Figure 3A**). Compared to the control cells, cells expressing TALE2-TET1-CD-WT did not show a global decrease in DNA methylation, indicating that TALE2-TET1-CD-WT is site-specific and does not modify genome-wide methylation (**Figure 3B**). The RRBS results confirmed the demethylation of the *SERPINB5* promoter, whereas other tumor suppressor genes such as *CDH1* are not affected (**Figure 3C**). To precisely quantify the impact on DNA methylation, we identified differentially methylated regions (DMRs) with more than 15% methylation difference ($p<0.01$) between cells expressing TALE2-TET1-CD-WT and

control cells. This identified only 68 DMRs showing differential methylation in TALE2-TET1-CD-WT cells compared to non-transfected cells and TALE2-TET1-CD-mut cells, which are all hypomethylated (**Figure S5A**). Only 6 of these DMRs were located in gene promoters (-500 to +500bp relative to the TSS) (**Figure S5B**), including the *SERPINB5* promoter (**Supplementary Table 1**). These DMRs could result from off-target binding, stochastic events or consequences of *SERPINB5* re-expression. Altogether this indicates that TALE2-TET1-CD-WT has minimal off-target activity and can induce durable and specific demethylation of *SERPINB5* in MDA-MB-231 cells.

TALE-mediated epigenetic reprogramming of SERPINB5 is durable in time

We then focused on the TALE2-TET1 strategy to test if the epigenetic reprogramming of *SERPINB5* in durable over time. For this purpose, we used lentiviral transduction to create stable MDA-MB-231 cells expressing TALE2-TET1-CD fused to an IRES-GFP under the control of a doxycycline-inducible promoter (Tet-On 3G System, **Figure 4A**). As a control we also created stable inducible MDA-MB-231 cells expressing GFP alone (**Figure 4A**). To validate the induction, we monitored GFP fluorescence in cells treated with and without Dox and observed strong GFP signal upon Dox treatment (**Figure S6**). Dox treatment for 14 days (**Figure 4B**) resulted in a progressive re-expression of *SERPINB5* in inducible TALE2-TET1-CD cells but not in GFP control cells (**Figure 4C**). To test if *SERPINB5* is also induced at the protein level, we performed western blot analysis and observed strong accumulation of SERPINB5 protein in Dox-treated TALE2-TET1-CD cells (**Figure 4D**). We then tested if the re-expression of *SERPINB5* is durable in time by monitoring *SERPINB5* expression after removing Dox from the medium, which efficiently and rapidly switches off the expression of the transgene (**Figure S6**). *SERPINB5* expression is detected after Dox removal and subculturing the cells for up to 30 days, even though the level of re-expression slightly goes down with time (**Figure 4E**). We observed the same results at the protein level by western blot (**Figure 4F**), and this was repeated in several independent experiments (**Figure S7**). Thus the de-repression of *SERPINB5* remains durable for at least 30 days when the expression of TALE2-TET1-CD is switched off in the cells. We then measured DNA methylation of the *SERPINB5* promoter by amplicon sequencing at several time points of Dox induction and after Dox removal. Dox treatment in cells expressing GFP only did not alter DNA methylation of *SERPINB5* (**Figure 4G**). In contrast, Dox treatment of cells expressing TALE2-TET1-CD lead to a decrease in methylation from 89% in control cells to 81% after 14 days (**Figure 4G**). The methylation profiles are similar to what we had already observed in transitory transfections, with the highest demethylation occurring at CpGs 8-11 close to the TALE2 binding site (**Figure 4H**). After Dox removal, the methylation slightly increased but even after 30 days the methylation levels were not restored to initial levels (**Figure 4G-H**), which coincides with

persistent expression of *SERPINB5*. The increase in methylation after Dox removal could be caused by a slow remethylation or a competitive advantage of *SERPINB5*-negative cells. Altogether, we conclude that targeted promoter demethylation induces durable *SERPINB5* expression for at least 30 days in MDA-MB-231 cells.

Impact of SERPINB5 demethylation on the cellular phenotype

As it has been previously shown that re-expression of *SERPINB5* in cancer cells could induce apoptosis and inhibit cell proliferation (Beltran et al., 2007; Garcia-Bloj et al., 2016), we examined if stable *SERPINB5* re-expression has a phenotypical impact in our system. We quantified apoptotic cells by flow cytometry after Annexin V and 7-AAD staining in our MDA-MB-231 cells expressing inducible TALE2-TET1-CD or GFP at several time points during Dox treatment and 10 days after Dox removal. As shown in **Figure S8**, the induction of TALE2-TET1-CD and re-expression of *SERPINB5* had no significant effect on the proportion of early or late apoptotic cells compared to non-treated cells and inducible GFP cells (**Figure S8A-B**). In addition, we measured cell proliferation after 10 to 14 days of Dox treatment and found no significant impact of TALE2-TET1-CD induction on proliferation relative to the control cells, even though this coincides with the period of highest *SERPINB5* induction (**Figure S8C**).

Material and methods

Cell culture. MDA-MB-231 cells (ATCC: HTB-26) were cultured in RPMI 1640 medium (Gibco, Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (FBS) and 50 µg/ml gentamycin sulfate (Gibco, Thermo Fisher Scientific). For DNA methylation inhibition experiments, MDA-MB-231 cells were treated with 5-aza-2'-deoxycytidine (Sigma) at 0.2, 2 or 5µM and harvested for analysis 24h, 48h and 72h after the beginning of the treatment. Cells were cultured in a humidified atmosphere at 37°C supplemented with 5% CO_2. All cell lines have been tested negative for mycoplasma.

Plasmids and cloning. The TALEs (TALE1 binding site: TAGCAGAATGAGCTGCTGCAGT, TALE2 binding site: TTGGAAGCTGTGCAGACAACAG) were designed and cloned by Labomics, and we obtained plasmids coding for TALE1-VP64-T2A-GFP and TALE2-VP64-T2A-GFP from Labomics. The plasmids contain *Bmt*I and *Xba*I restriction sites flanking the VP64 sequence. The coding sequences of the human TET1 catalytic domain (CD) or the human AID protein were synthesised by Integrated DNA

Technologies (IDT) and cloned into the *Bmt*I and *Xba*I sites. The mutant TET1 plasmid was created by directed mutagenesis using the Phusion High Fidelity DNA polymerase (Thermo Fischer Scientific) to introduce two point mutations (H1671Y and D1673A) (Tahiliani et al., 2009). The dCas9-TET1-CD-GFP plasmid was constructed based on the pdCas9-DNMT3A-EGFP plasmid (Addgene #71666). The original plasmid was digested with the restriction enzymes *Bam*HI and *Fse*I to remove the DNMT3A fragment and replace it with the catalytic domain of human TET1 that was amplified by PCR using forward and reverse primers introducing *Bam*HI and *Fse*I restriction sites. Two undesired *Bbs*I restriction site within the catalytic domain of TET1 were removed by site-directed mutagenesis. The TET1 was mutated by directed mutagenesis to introduce the same two point mutations as described above. Three sgRNAs targeting the *SERPINB5* promoter (sgRNA2: TTCCTGCCCGAACATGTTGG, sgRNA3: ATTGTGGACAAGCTGCCAAG, sgRNA4: GTAGGAGAGGAGTGCCGCCG) were designed using http://crispr.mit.edu, synthesised as pairs of oligonucleotides, annealed and cloned into the *Bbs*I site of the dCas9-TET1-CD-GFP plasmids (WT and mut). For the T7 endonuclease assays, the gRNAs were cloned into the *Bbs*I site of the pSpCas9(BB)-2A-GFP plasmid (Addgene #48138). In order to create an inducible MDA-MB-231 cell line expressing TALE2-TET1 under the control of a Doxyxycline-inducible promoter, we subcloned the TALE2-TET1 sequence in the MCS-1 and the EGFP sequence in the MCS-2 of the lentiviral pLVX-TRE3G-IRES plasmid (Tet-On 3G System, Clontech #631362). A plasmid containing only the EGFP sequence cloned in the MCS-2 was used to create the control cell line.

Transfections. MDA-MB-231 cells were transfected with TALE-VP64-T2A-GFP, TALE-TET1-CD-T2A-GFP (WT and mut), TALE-AID-T2A-GFP and dCas9-TET1-CD-T2A-GFP (WT and mut) plasmids using the Trans-IT BrCa transfection reagent (Mirus Bio) according to the manufacturer's instruction. In brief, 60-80% confluent cells in 100 mm dishes were transfected with 15µg of plasmid. As a control, we also transfected cells with a plasmid coding for GFP only (pEGFP-C1, Clontech). GFP-positive and negative cells were sorted 48h after transfection for DNA methylation and gene expression analysis using *a* BD FACS Vantage cell Sorter (BD Biosciences). All transfections were performed in triplicate.

Inducible cell lines and Dox treatment. MDA-MB-231 cells stably expressing TALE2-TET1-CD-IRES-GFP or EGFP alone under the control of a Doxycycline-inducible promoter (PTRE3G) were generated by lentiviral transduction. Production of lentiviral particles in HEK293T cells and transduction of MDA-MB-231 cells were performed at the GIGA-Viral Vectors platform (University of Liège, Belgium). Cells were transduced with the pLVX-Tet3G plasmid and twice with the pLVX-TRE3G-TALE2-TET1-CD-IRES-EGFP or the pLVX-TRE3G-IRES-EGFP (coding for EGFP alone). After transduction, stably infected cells were selected with Puromycin and G418 and kept as a pool. For the induction

experiments, cells were treated with Doxycycline (1 µg/mL, Clontech) for up to 14 days. The medium and Doxycycline were renewed every day. Non-treated cells were cultured in parallel to serve as controls. Doxycycline was then removed from the medium and the cells were cultured up to 30 days after Dox removal. The induction was followed at several time points during and after treatment by monitoring GFP fluorescence by flow cytometry using a FACS Calibur (BD Biosciences).

T7 endonuclease assays. Genomic DNA was extracted from MDA-MB-231 cells transfected with plasmids coding for Cas9 and sgRNA 3 and 4 targeting the *SERPINB5* promoter. The regions flanking the sgRNA target site in the *SERPINB5* promoter were amplified by PCR with the Platinum Taq DNA polymerase (Invitrogen, Thermo Fisher Scientific) from 2 ng of DNA using the following PCR conditions: 20 cycles of 30s at 95°C, 30s at 64–54°C (with a 0.5°C decrease per cycle), 50s at 72°C followed by 35 cycles of 30s at 95°C, 30s at 54°C, 50s at 72°C. 200ng of the purified PCR products were denatured-annealed following the reaction: 95°C 5 min; -2°C/s to 85°C; -0,1°C/s to 25°C; 12°C. The products were digested with 10U of T7 endonuclease I for 20 min at 37°C. The endo I was inactivated with 25mM EDTA and the products were run in a 1,5% agarose gel. The primers used for the T7 endonuclease assays are given in the **Supplementary Table 2**.

COBRA. Genomic DNA was extracted with the AllPrep DNA/RNA Mini Kit (Qiagen). Between 200 and 1000 ng of genomic DNA was converted using the EpiTect Bisulfite Kit (Qiagen) according to the manufacturer's instructions. For COBRA, the *SERPINB5* promoter region was amplified by PCR with the Platinum Taq DNA Polymerase (Invitrogen, Thermo Fisher Scientific) using the following conditions: 20 cycles of 30s at 95°C, 30s at 62–52°C (with a 0.5°C decrease per cycle), 50s at 72°C followed by 35 cycles of 30s at 95°C, 30s at 52°C, 50s at 72°C. 40 ng of PCR amplicon was digested with 10U *Taq*αI restriction enzyme (Thermo Fisher Scientific) for 1h at 65°C, and an equal amount of PCR amplicon was used for the undigested control. The samples were then purified with the QIAquick PCR purification kit (Qiagen) and migrated on a 2.5% agarose gel. The primers used for COBRA are given in the **Supplementary Table 2**.

High throughput sequencing of bisulfite amplicons. Bisulfite converted DNA was amplified with two pairs of primers in the *SERPINB5* promoter using the same PCR conditions as for COBRA. The forward primers were barcoded at the 5' end with 4 nucleotides to allow multiplexing of samples in the sequencing runs. The primers are given in the **Supplementary Table 2**. The amplicons were purified using AMPure XP beads (Beckman Coulter), pooled and sequencing libraries were prepared and sequenced (2 x 125 bp) on a HiSeq 4000 by GATC Biotech. We obtained a sequencing depth between 10,000 and 50,000X for each experimental sample. Reads were trimmed with TrimGalore

(v0.4.2) to remove adapter sequences and low-quality ends with a Phred score below 20, and aligned to the expected amplicon sequences with BSMAP (v2.74). For the mapping, we allowed a maximum of 4 mismatches. We extracted methylation scores as the ratio of the number of Cs over the total number of Cs and Ts and combined CpG methylation ratios from both strands.

RRBS. We performed RRBS in three biological replicates of non-transfected cells and cells transfected with TALE2-TET1-CD-WT or TALE2-TET1-CD-mut. RRBS libraries were prepared as described (Auclair et al., 2014), with some modifications. Because the *SERPINB5* promoter sequence does not contain two *Msp*I sites, we performed a double digestion protocol. Briefly, 50 ng genomic DNA were digested for 5h with *Msp*I and *Mun*I (Thermo Fisher Scientific) followed by end-repair, A-tailing and ligation to methylated adapters with T4 DNA ligase (Thermo Fisher Scientific) in Tango 1X buffer. Small fragments (150 to 400 bp) were selected by electrophoresis on a 3% (w/v) agarose 0.5X TBE gel and purified with the MinElute gel extraction kit (Qiagen). Bisulfite conversion was performed by two rounds of conversion with the EpiTect kit (Qiagen) according to the manufacturer's instructions. RRBS libraries were amplified by 14 cycles of PCR with PfU Turbo Cx hotstart DNA polymerase (Agilent) and purified with AMPure magnetic beads (Beckman Coulter). The libraries were paired-end sequenced (2x75 bp) with a HiSeq4000 by Integragen SA (Evry, France). Sequencing reads were trimmed with TrimGalore (v0.4.2) and aligned to the human hg38 genome with BSMAP (v2.74). For the mapping, we allowed a maximum of two mismatches. We combined CpG methylation ratios from both strands and only retained CpGs with a minimum sequencing depth of 8X. To compare methylation profiles between samples, we generated pairwise plots with methylation values calculated in 400 bp windows containing at least 3 CpGs. The dendrogram was generated based on the Euclidean distances between CpG methylation scores with the dist and hclust functions in R. To identify DMRs, we used the eDMR algorithm from the methylKit R package (Li et al., 2013b). We identified DMRs with a change in methylation >15%, at least 4 differentially methylated CpGs (DMCs), and a p-value <0.01 (t-test) between the triplicates. DMRs were considered promoter proximal if they are located -500 to +500 bp from an annotated RefSeq TSS.

RT-qPCR. Total RNA was extracted together with DNA samples with the AllPrep DNA/RNA Mini Kit (Qiagen). RNA samples were reverse transcribed with the Maxima First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). Quantitative PCR reactions were performed with the KAPA SYBR FAST kit (Clinisciences) on a StepOnePlus Real-Time PCR System (Life Technologies, Thermo Fisher Scientific). Amplification cycling conditions were as follows: 95 °C for 3 min and 40 cycles (95 °C for 20 sec, 60 °C for 30 s). All samples were assayed in triplicate reactions, and we systematically amplified in parallel no-RT controls to exclude the presence of contaminating genomic DNA. For data

representation, the expression of genes (*SERPINB5*, *DAPK1*, *CDH1*) was normalized relative to the expression of three housekeeping genes: *ACTB*, *RPL13A* and *SDHA*. For the expression analysis in inducible cells, we then calculated the fold change of *SERPINB5* expression in dox treated cells relative to non-treated cells at each time point. The primer sequences for all genes are given in the **Supplementary Table 2**.

Western blotting. Total proteins were extracted with the RIPA Lysis Buffer System (Clinisciences) supplemented with Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific) and Halt Phosphatase Inhibitor Cocktail (Thermo Fisher Scientific). Proteins were run on 10% SDS-polyacrylamide gels. The proteins were transferred to a 0,2 µm Nitrocellulose membrane (BioRad) using the Trans-Blot Turbo Blotting System (BioRad). The membranes were washed in Tris-buffered saline solution (TBS 10X, Euromedex) with 0,05% Tween 20 and blocked with 5% nonfat milk in TBS 1X solution with Tween 20 for 1 h. The membranes were incubated with primary antibodies overnight at 4°C (SERPINB5) or 1 h at RT (VINCULIN), then washed and incubated for 1h with horseradish peroxidase-conjugated secondary antibodies followed by chemiluminescence detection using the ECL detection reagent (Amersham, GE Healthcare). SERPINB5 was detected with a mouse monoclonal antibody (BD Biosciences #554292, 1:1000) and a secondary sheep anti-mouse antibody (GE Healthcare Lifescience #RPN4201, 1:30000). Vinculin was used as a loading control and detected with a rabbit anti-vinculin antibody (Abcam #ab129002, 1:10000) and a secondary goat anti-rabbit antibody (Sigma #A0545, 1:50000).

Growth and apoptosis assays. We assessed proliferation and apoptosis in the doxycycline inducible cell lines. To quantitatively study the percentage of cell death (early apoptotic, late apoptotic or necrotic cells) we performed Annexin-V PE and 7-AAD staining using commercial kits (BD Bioscience) according to the manufacturer's instructions. The percentages of apoptotic cells were measured by flow cytometry using a FACSCelesta flow cytometer (BD Biosciences) 7, 10 and 14 days after the beginning of the Dox treatment, and 10 days after Dox removal. A positive control was induced at each time point with campthotecin (Sigma) at 8µM for 4h at 37ºC to determine the flow cytometry settings. To measure cell proliferation, Dox induced cells (9 days after the beginning of the treatment) were seeded in 96-well plates (2,000 cells per well) and cultured for 1 to 5 days. The medium containing Dox was refreshed every day. The cells were incubated with the WST-1 Cell Proliferation Reagent (Roche) for 4h at 37°C, and then the absorption was measured at 440 nm using the Asys UVM 340 Microplate Reader (Biochrom). All experiments were performed in biological triplicate.

Statistical analysis. All experiments were performed at least in three independent biological replicates. Error bars represent standard error of mean (S.E.M.). Statistical significance was tested with a two-tailed, unpaired, unequal variance t-test.

Discussion

It is well known that aberrant promoter CpG methylation of tumor suppressor genes contributes to cancer progression. Previous work has highlighted the ability to use programmable DNA binding domains to perform TET1-mediated targeted DNA demethylation in mammalian cells, suggesting that these methods could have therapeutic potential to achieve specific reactivation of TSGs in cancer. However, some key issues still remain to be investigated in more depth, in particular comparing the efficiency of different tethering strategies, investigating the genome wide specificity of the demethylation, and investigating the durability of the reprogramming. In this study, we provide several important findings that contribute to address these issues.

First, we compared the efficiency of TALE and dCas9-based strategies to perform targeted TET1-mediated demethylation of the *SERPINB5* promoter in breast cancer cells. Although dCas9 is easier to use, our results indicate that TALE-TET1 fusion proteins were more efficient than dCas9-TET1 for inducing demethylation and reactivation of *SERPINB5*. A previous study also reported a lower efficiency of the dCas9 system compared to TALEs for activating endogenous genes in mouse fibroblasts (Gao et al., 2014). Similarly, TALEs were found more efficient than ZFPs and dCas9 when fused to VP64 for targeted activation of *SERPINB5* in human cells (Garcia-Bloj et al., 2016). Together these studies and our work suggest that TALEs are more potent than dCas9 for targeted activation experiments. One probable explanation is that the dCas9 complex has an intrinsic negative impact on gene activation by physically interfering with transcriptional elongation and transcription factor binding (Qi et al., 2013). Another possibility is that dCas9 might not efficiently bind to loci containing dense CpG methylation, as suggested (Cano-Rodriguez et al., 2016; Wu et al., 2014). Our TALE2-TET1 construct achieved robust reactivation of *SERPINB5* expression, however in contrast to a previous report (Beltran et al., 2007) this was not sufficient to induce apoptosis and reduced cell proliferation of MDA-MB-231 cells. Most likely this is because the fold change activation achieved by TALE2-TET1 in inducible cell lines is lower than the one previously obtained with constitutive expression of a synthetic ZFP-VP64 construct (Beltran et al., 2007). In support of this, the previous study showed a positive correlation between the magnitude of transcriptional upregulation of *SERPINB5* achieved by different ZFP-VP64 constructs and the phenotypic impact in the cells (Beltran et al., 2007).

Second, we investigated the genome-wide specificity of demethylation induced by tethering TET1 with TALEs. RRBS showed that TALE2-TET1 induces demethylation of the *SERPINB5* promoter without altering global methylation and with minimal off-target effects. This is the first study demonstrating the genome-wide specificity of targeted demethylation with engineered TALE-TET1 constructs. In comparison, the genome-wide specificity of dCas9-TET1 has never been tested in previous reports. However several groups independently reported high off-target methylation activity with transfection or transduction of dCas9-DNMT3A, even in the presence of sgRNAs (Galonska et al., 2018; Lin et al., 2018; Pflueger et al., 2018). This suggests that dCas9-TET1 could also have substantial off-target demethylation activity, which remains to be tested in the future.

Third, we investigated the durability of TSG reactivation induced by targeted demethylation. In the past years, several studies investigated the durability of targeted methylation deposition by DNMT3A. Some studies showed that targeted DNA methylation can result in sustained silencing (McDonald et al., 2016; Stolzenburg et al., 2015), whereas others found that targeted methylation is lost upon cell culture (Amabile et al., 2016; Galonska et al., 2018; Kungulovski et al., 2015; Vojta et al., 2016). In contrast, very few studies have addressed the stability of epigenetic reactivation by programmable DNA binding proteins. In particular, in all previous work performing targeted activation of *SERPINB5* or targeted demethylation of TSGs in cancer cells, the durability of the activation was not investigated. Here, using a doxycycline inducible system, we showed that TET1-mediated demethylation and reactivation of *SERPINB5* is durable in cultured cells. *SERPINB5* reactivation was detected at the mRNA and protein level for up to 30 days after doxycycline removal. This is in line with another study showing that dCas9-TET1 mediated activation of the *FRM1* gene is stable for two weeks in cultured cells after inhibition of dCas9 activity (Liu et al., 2018). Collectively, these data show that targeted removal of DNA methylation can induce durable gene activation and reinforce the idea that DNA methylation is the major determinant of long-term silencing. In support of this, tethering H3K4me3 to gene promoters leads to sustained gene activation only when the target promoter has no DNA methylation (Cano-Rodriguez et al., 2016).

Although *SERPINB5* reactivation is durable, we observed a slight decrease in expression and increase in DNA methylation over time in culture. This could indicate a slow remethylation process, which could be tested by simultaneously expressing TALE2-TET1 and inhibiting *de novo* methyltransferases. Alternatively, it is plausible that this reflects a competitive advantage of SERPINB5-negative cells. Indeed, although we did not observe a reduced cell proliferation at the population level, previous studies demonstrated that high SERPINB5 expression inhibits proliferation of MDA-MB-231 cells. Therefore, if there is cell to cell heterogeneity in the response to doxycycline, it is likely that the cells that are the most demethylated and express the highest levels of *SERPINB5* in the population grow slower, leading to a progressive apparent increase in methylation. Further work

is also needed to reinforce the efficiency and durability of the demethylation. This could be achieved by combining TET1 with other epigenome editing tools (Amabile et al., 2016), or using the dCas9-SunTag system that allows the recruitment of multiple copies of TET1 to a targeted site (Morita et al., 2016).

Although the durability of methylation reprogramming is probably context-dependent and needs to be tested at other loci and in other cell types, our work highlights the potential of targeted methylation engineering to induce durable reactivation of TSGs in cancer cells with less collateral effects than current drugs inhibiting DNA methylation genome-wide.

Data availability

The RRBS data is available in the Gene Expression Omnibus (GEO) under the accession number GSE119752.

(Reviewer link: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE119752, token: ahmbqucyhrmdhkd)

Funding

This work was supported by the European Research Council (Consolidator grant number 615371 to M.W.), the Fondation ARC (SFI20121205729), the Ligue contre le Cancer (Grand Est), the Institut National du Cancer (INCa), the ITMO Cancer. A.A.-L. was supported by a fellowship from the Fondation pour la Recherche Médicale.

References

Amabile, A., Migliara, A., Capasso, P., Biffi, M., Cittaro, D., Naldini, L., and Lombardo, A. (2016). Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing. Cell *167*, 219-232 e214.

Auclair, G., Guibert, S., Bender, A., and Weber, M. (2014). Ontogeny of CpG island methylation and specificity of DNMT3 methyltransferases during embryonic development in the mouse. Genome biology *15*, 545.

Beltran, A., Parikh, S., Liu, Y., Cuevas, B.D., Johnson, G.L., Futscher, B.W., and Blancafort, P. (2007). Re-activation of a dormant tumor suppressor gene maspin by designed transcription factors. Oncogene *26*, 2791-2798.

Cano-Rodriguez, D., Gjaltema, R.A., Jilderda, L.J., Jellema, P., Dokter-Fokkens, J., Ruiters, M.H., and Rots, M.G. (2016). Writing of H3K4Me3 overcomes epigenetic silencing in a sustained but context-dependent manner. Nature communications *7*, 12284.

Chen, H., Kazemier, H.G., de Groote, M.L., Ruiters, M.H., Xu, G.L., and Rots, M.G. (2014). Induced DNA demethylation by targeting Ten-Eleven Translocation 2 to the human ICAM-1 promoter. Nucleic acids research *42*, 1563-1574.

Cher, M.L., Biliran, H.R., Jr., Bhagat, S., Meng, Y., Che, M., Lockett, J., Abrams, J., Fridman, R., Zachareas, M., and Sheng, S. (2003). Maspin expression inhibits osteolysis, tumor growth, and angiogenesis in a model of prostate cancer bone metastasis. Proceedings of the National Academy of Sciences of the United States of America *100*, 7847-7852.

Choudhury, S.R., Cui, Y., Lubecka, K., Stefanska, B., and Irudayaraj, J. (2016). CRISPR-dCas9 mediated TET1 targeting for selective DNA demethylation at BRCA1 promoter. Oncotarget *7*, 46545-46556.

Domann, F.E., Rice, J.C., Hendrix, M.J., and Futscher, B.W. (2000). Epigenetic silencing of maspin gene expression in human breast cancers. International journal of cancer *85*, 805-810.

Dzinic, S.H., Bernardo, M.M., Li, X., Fernandez-Valdivia, R., Ho, Y.S., Mi, Q.S., Bandyopadhyay, S., Lonardo, F., Vranic, S., Oliveira, D.S.*, et al.* (2017). An Essential Role of Maspin in Embryogenesis and Tumor Suppression. Cancer research *77*, 886-896.

Futscher, B.W., Oshiro, M.M., Wozniak, R.J., Holtan, N., Hanigan, C.L., Duan, H., and Domann, F.E. (2002). Role for DNA methylation in the control of cell type specific maspin expression. Nature genetics *31*, 175-179.

Galonska, C., Charlton, J., Mattei, A.L., Donaghey, J., Clement, K., Gu, H., Mohammad, A.W., Stamenova, E.K., Cacchiarelli, D., Klages, S.*, et al.* (2018). Genome-wide tracking of dCas9-methyltransferase footprints. Nature communications *9*, 597.

Gao, F., Shi, H.Y., Daughty, C., Cella, N., and Zhang, M. (2004). Maspin plays an essential role in early embryonic development. Development *131*, 1479-1489.

Gao, X., Tsang, J.C., Gaba, F., Wu, D., Lu, L., and Liu, P. (2014). Comparison of TALE designer transcription factors and the CRISPR/dCas9 in regulation of gene expression by targeting enhancers. Nucleic acids research *42*, e155.

Garcia-Bloj, B., Moses, C., Sgro, A., Plani-Lam, J., Arooj, M., Duffy, C., Thiruvengadam, S., Sorolla, A., Rashwan, R., Mancera, R.L.*, et al.* (2016). Waking up dormant tumor suppressor genes with zinc fingers, TALEs and the CRISPR/dCas9 system. Oncotarget *7*, 60535-60554.

Huisman, C., van der Wijst, M.G., Schokker, M., Blancafort, P., Terpstra, M.M., Kok, K., van der Zee, A.G., Schuuring, E., Wisman, G.B., and Rots, M.G. (2016). Re-expression of Selected Epigenetically Silenced Candidate Tumor Suppressor Genes in Cervical Cancer by TET2-directed Demethylation. Molecular therapy : the journal of the American Society of Gene Therapy *24*, 536-547.

Kungulovski, G., and Jeltsch, A. (2016). Epigenome Editing: State of the Art, Concepts, and Perspectives. Trends in genetics : TIG *32*, 101-113.

Kungulovski, G., Nunna, S., Thomas, M., Zanger, U.M., Reinhardt, R., and Jeltsch, A. (2015). Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained. Epigenetics & chromatin *8*, 12.

Li, S., Garrett-Bakelman, F.E., Akalin, A., Zumbo, P., Levine, R., To, B.L., Lewis, I.D., Brown, A.L., D'Andrea, R.J., Melnick, A.*, et al.* (2013). An optimized algorithm for detecting and annotating regional differential methylation. BMC bioinformatics *14 Suppl 5*, S10.

Lin, L., Liu, Y., Xu, F., Huang, J., Daugaard, T.F., Petersen, T.S., Hansen, B., Ye, L., Zhou, Q., Fang, F.*, et al.* (2018). Genome-wide determination of on-target and off-target characteristics for RNA-guided DNA methylation by dCas9 methyltransferases. GigaScience *7*, 1-19.

Liu, X.S., Wu, H., Krzisch, M., Wu, X., Graef, J., Muffat, J., Hnisz, D., Li, C.H., Yuan, B., Xu, C.*, et al.* (2018). Rescue of Fragile X Syndrome Neurons by DNA Methylation Editing of the FMR1 Gene. Cell *172*, 979-992 e976.

Maass, N., Biallek, M., Rosel, F., Schem, C., Ohike, N., Zhang, M., Jonat, W., and Nagasaki, K. (2002). Hypermethylation and histone deacetylation lead to silencing of the maspin gene in human breast cancer. Biochemical and biophysical research communications *297*, 125-128.

Maeder, M.L., Angstman, J.F., Richardson, M.E., Linder, S.J., Cascio, V.M., Tsai, S.Q., Ho, Q.H., Sander, J.D., Reyon, D., Bernstein, B.E.*, et al.* (2013). Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. Nature biotechnology *31*, 1137-1142.

McDonald, J.I., Celik, H., Rois, L.E., Fishberger, G., Fowler, T., Rees, R., Kramer, A., Martens, A., Edwards, J.R., and Challen, G.A. (2016). Reprogrammable CRISPR/Cas9-based system for inducing site-specific DNA methylation. Biology open *5*, 866-874.

Morita, S., Noguchi, H., Horii, T., Nakabayashi, K., Kimura, M., Okamura, K., Sakai, A., Nakashima, H., Hata, K., Nakashima, K.*, et al.* (2016). Targeted DNA demethylation in vivo using dCas9-peptide repeat and scFv-TET1 catalytic domain fusions. Nature biotechnology *34*, 1060-1065.

Pflueger, C., Tan, D., Swain, T., Nguyen, T., Pflueger, J., Nefzger, C., Polo, J.M., Ford, E., and Lister, R. (2018). A modular dCas9-SunTag DNMT3A epigenome editing system overcomes pervasive off-target activity of direct fusion dCas9-DNMT3A constructs. Genome research. 28, 1193-1206.

Primeau, M., Gagnon, J., and Momparler, R.L. (2003). Synergistic antineoplastic action of DNA methylation inhibitor 5-AZA-2'-deoxycytidine and histone deacetylase inhibitor depsipeptide on human breast carcinoma cells. International journal of cancer *103*, 177-184.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell *152*, 1173-1183.

Ramiro, A.R., and Barreto, V.M. (2015). Activation-induced cytidine deaminase and active cytidine demethylation. Trends in biochemical sciences *40*, 172-181.

Rivenbark, A.G., Stolzenburg, S., Beltran, A.S., Yuan, X., Rots, M.G., Strahl, B.D., and Blancafort, P. (2012). Epigenetic reprogramming of cancer cells via targeted DNA methylation. Epigenetics : official journal of the DNA Methylation Society *7*, 350-360.

Seftor, R.E., Seftor, E.A., Sheng, S., Pemberton, P.A., Sager, R., and Hendrix, M.J. (1998). maspin suppresses the invasive phenotype of human breast carcinoma. Cancer research *58*, 5681-5685.

Shao, L.J., Shi, H.Y., Ayala, G., Rowley, D., and Zhang, M. (2008). Haploinsufficiency of the maspin tumor suppressor gene leads to hyperplastic lesions in prostate. Cancer research *68*, 5143-5151.

Sheng, S., Carey, J., Seftor, E.A., Dias, L., Hendrix, M.J., and Sager, R. (1996). Maspin acts at the cell membrane to inhibit invasion and motility of mammary and prostatic cancer cells. Proceedings of the National Academy of Sciences of the United States of America *93*, 11669-11674.

Shi, H.Y., Liang, R., Templeton, N.S., and Zhang, M. (2002). Inhibition of breast tumor progression by systemic delivery of the maspin gene in a syngeneic tumor model. Molecular therapy : the journal of the American Society of Gene Therapy *5*, 755-761.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. Nature reviews. Genetics *14*, 204-220.

Stolzenburg, S., Beltran, A.S., Swift-Scanlan, T., Rivenbark, A.G., Rashwan, R., and Blancafort, P. (2015). Stable oncogenic silencing in vivo by programmable and targeted de novo DNA methylation in breast cancer. Oncogene *34*, 5427-5435.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L*., et al.* (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930-935.

Vojta, A., Dobrinic, P., Tadic, V., Bockor, L., Korac, P., Julg, B., Klasic, M., and Zoldos, V. (2016). Repurposing the CRISPR-Cas9 system for targeted DNA methylation. Nucleic acids research *44*, 5615-5628.

Wozniak, R.J., Klimecki, W.T., Lau, S.S., Feinstein, Y., and Futscher, B.W. (2007). 5-Aza-2'-deoxycytidine-mediated reductions in G9A histone methyltransferase and histone H3 K9 di-methylation levels are linked to tumor suppressor gene reactivation. Oncogene *26*, 77-90.

Wu, X., Scott, D.A., Kriz, A.J., Chiu, A.C., Hsu, P.D., Dadon, D.B., Cheng, A.W., Trevino, A.E., Konermann, S., Chen, S*., et al.* (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nature biotechnology *32*, 670-676.

Wu, X., and Zhang, Y. (2017). TET-mediated active DNA demethylation: mechanism, function and beyond. Nature reviews. Genetics *18*, 517-534.

Xu, X., Tao, Y., Gao, X., Zhang, L., Li, X., Zou, W., Ruan, K., Wang, F., Xu, G.L., and Hu, R. (2016). A CRISPR-based approach for targeted DNA demethylation. Cell discovery *2*, 16009.

Zhang, M., Sheng, S., Maass, N., and Sager, R. (1997). mMaspin: the mouse homolog of a human tumor suppressor gene inhibits mammary tumor invasion and motility. Mol Med *3*, 49-59.

Zou, Z., Anisowicz, A., Hendrix, M.J., Thor, A., Neveu, M., Sheng, S., Rafidi, K., Seftor, E., and Sager, R. (1994). Maspin, a serpin with tumor-suppressing activity in human mammary epithelial cells. Science *263*, 526-529.

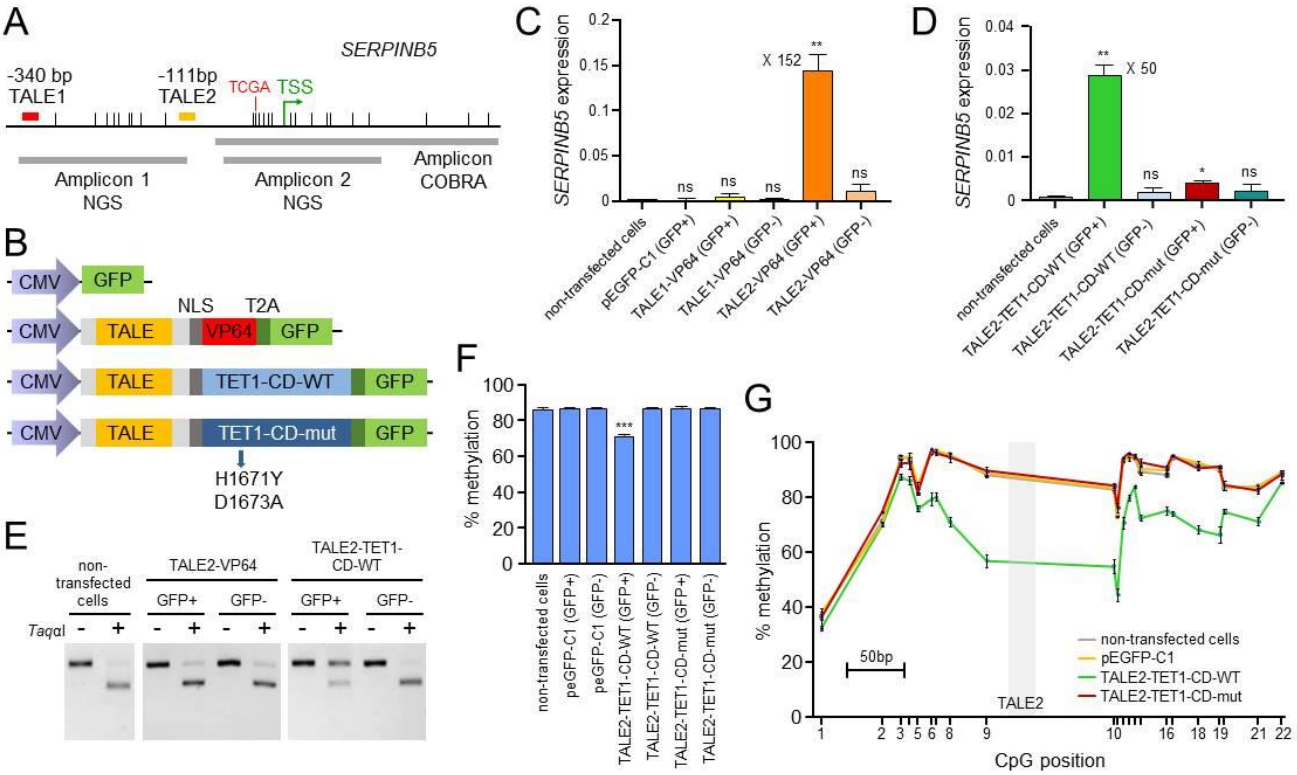


Figure 1. TALE-TET1 mediated demethylation and reactivation of *SERPINB5* in MDA-MB-231 cells.

A. Schematic illustration of the human *SERPINB5* promoter region and the TALE binding sites. The TALE1 and TALE2 binding sites are located -340 bp and -111 bp from the TSS, respectively. The vertical bars represent the position of CpGs. The positions of the bisulfite PCR amplicons for COBRA and next generation sequencing (NGS) are indicated in grey, and the position of the TaqαI restriction site for COBRA is indicated in red. **B.** Schematic illustration of the vectors coding for GFP, TALE-VP64 and TALE-TET1-CD (WT and mutant) used in the experiments. Transcription is driven by a CMV promoter (arrow). NLS= Nuclear Localisation Signal. **C.** RT-qPCR analysis of *SERPINB5* expression in cells transfected with TALE-VP64 vectors, calculated as a ratio relative to the expression of three housekeeping genes (RPL13A, SDHA, ACTB) (mean ± SEM, n= 3 or 4 biological replicates). For cells transfected with TALE2-VP64, the fold change relative to non-transfected cells is indicated. **D.** RT-qPCR analysis of SERPINB5 expression in cells transfected with TALE2-TET1-CD vectors (mean ± SEM, n= 3 biological replicates). For cells transfected with TALE2-TET1-CD-WT, the fold change relative to non-transfected cells is indicated. **E.** Methylation analysis in the *SERPINB5* promoter by COBRA. Hypomethylation is evident is cells transfected with TALE2-TET1-CD-WT. **F.** Methylation analysis in the *SERPINB5* promoter by amplicon-Next Generation Sequencing (NGS). The histogram shows the average methylation of all CpGs contained in both amplicons. The global methylation decreases from 87% in non-transfected cells to 71% in cells transfected with TALE2-TET1-CD-WT. Data are shown as the mean ± SEM (n= 3 or 4 biological replicates). **G.** Graph showing the methylation scores at individual CpGs calculated by amplicon-NGS. Data are shown as the mean ± SEM (n= 3 or 4 biological replicates). The position of the TALE2 binding site is indicated in grey. ns: not significant, *: p<0.05, **: p<0.01, ***: p<0.001 (unpaired t-test relative to non-transfected cells).

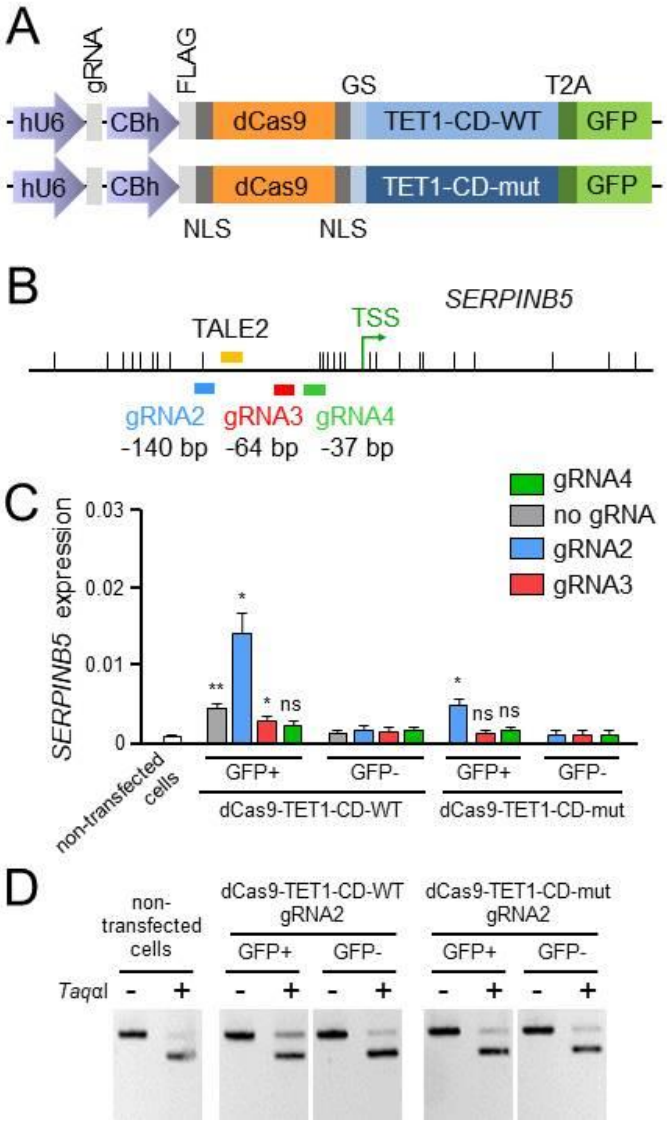


Figure 2. Low efficiency of dCas9-TET1 for reactivating *SERPINB5* in MDA-MB-231 cells.

A. Schematic illustration of the vectors coding for the gRNA and dCas9-TET1-CD (WT and mutant) used in the experiments. The promoters are indicated by arrows. **B.** Location of the gRNA sequences in the human *SERPINB5* promoter. The gRNA2, gRNA3 and gRNA4 are located -140 bp, -64 bp and -37 bp from the TSS, respectively. **C.** RT-qPCR analysis of *SERPINB5* expression in cells transfected with dCas9-TET1-CD vectors (mean ± SEM, n= 3 biological replicates). ns: not significant, *: p<0.05, **: p<0.01, ***: p<0.001 (unpaired t-test relative to non-transfected cells). **D.** Methylation analysis in the *SERPINB5* promoter by COBRA in cells transfected with dCas9-TET1-CD vectors.

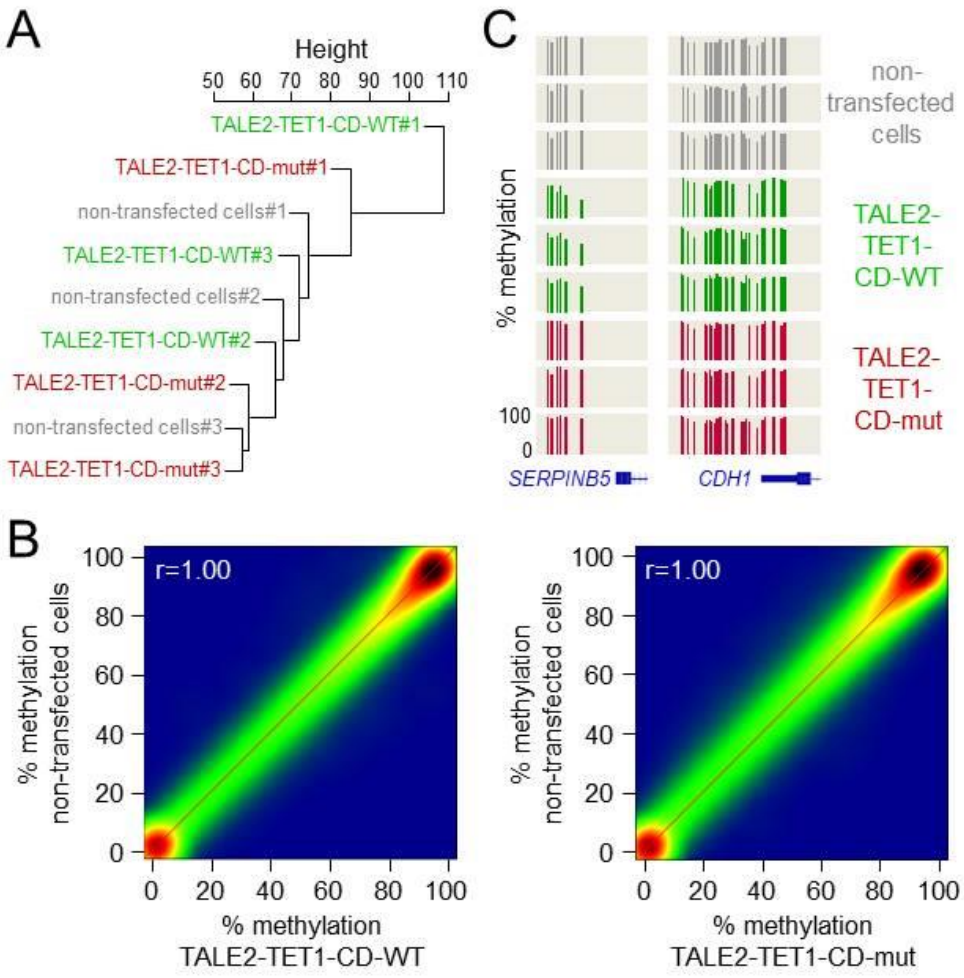


Figure 3. Genome-wide specificity of TALE-mediated demethylation.

A. Dendrogram representing the hierarchical clustering of CpG methylation scores measured in three independent RRBS experiments in cells transfected with TALE2-TET1-CD-WT, TALE2-TET1-CD-mut and non-transfected cells. **B.** Pairwise correlation plot of methylation scores (measured in 400 bp tiles) in non-transfected cells compared to cells expressing TALE2-TET1-CD-WT or TALE2-TET1-CD-mut. The colors ranging from blue to dark red indicate the density of data points. **C.** Single CpG methylation profiles assessed by RRBS in the SERPINB5 and CDH1 promoters in cells transfected with TALE2-TET1-CD-WT (green), TALE2-TET1-CD-mut (red) and non-transfected cells (grey). Each vertical bar represents the methylation of a single CpG from 0 to 100%.

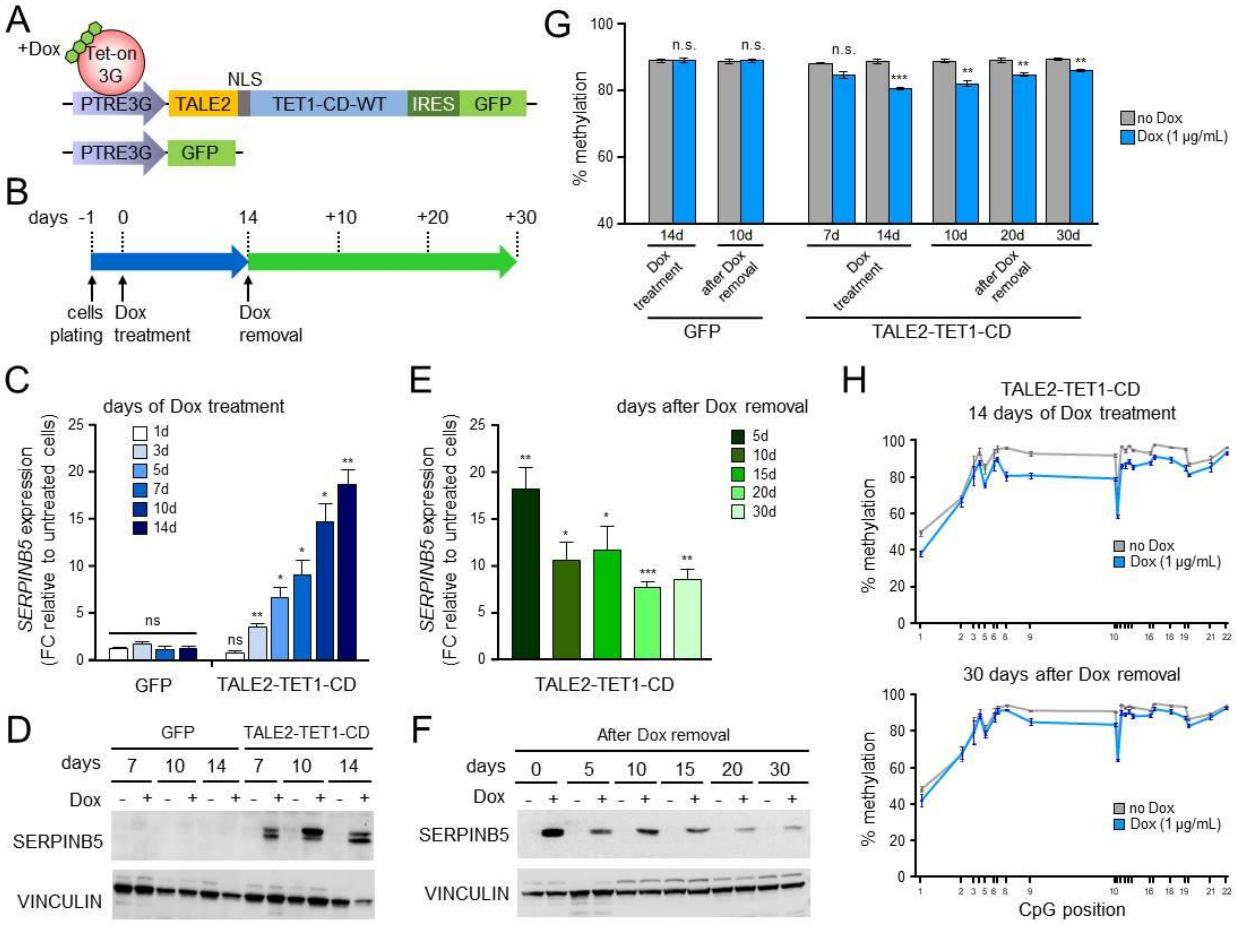
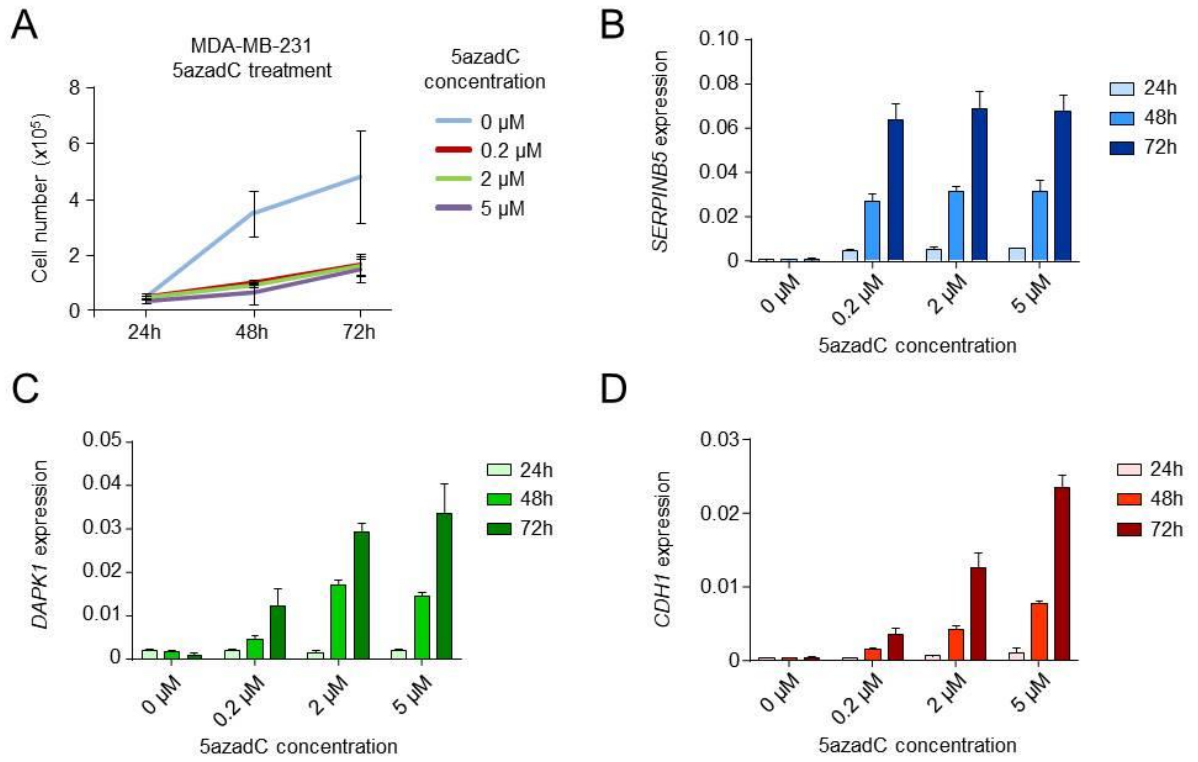


Figure 4. Reactivation of *SERPINB5* by TALE-mediated demethylation is durable.

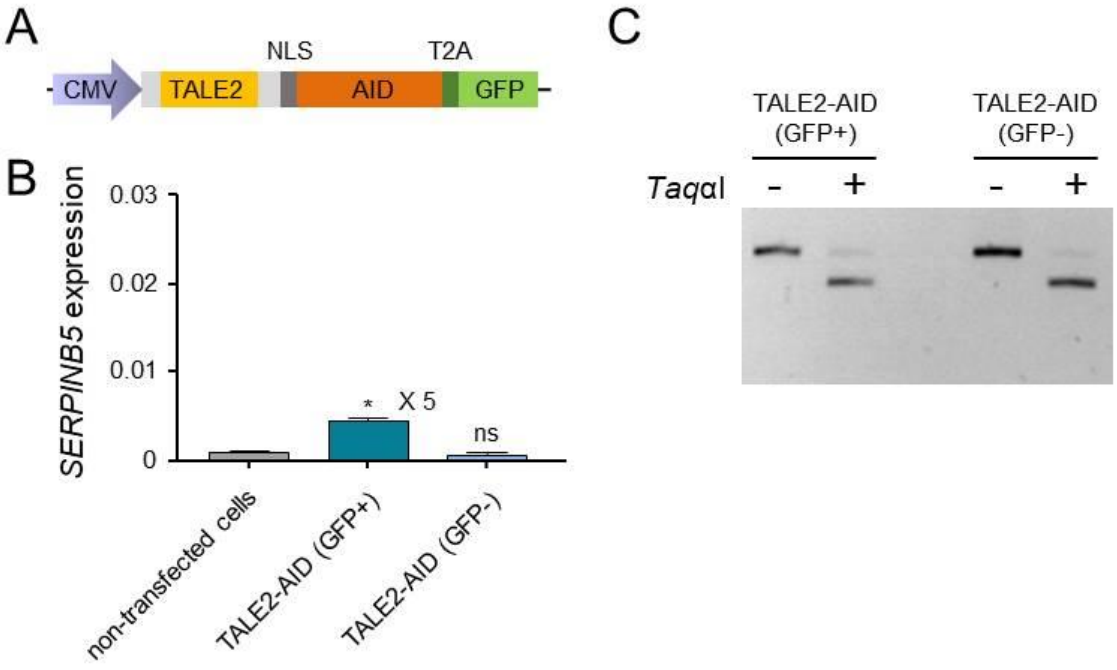
A. Schematic illustration of the vectors used to generate stable inducible cells lines. The expression of TALE2-TET1-CD-WT-IRES-GFP or GFP alone is under the control of a TRE3G promoter that is induced by the Tet-On 3G transactivator protein in the presence of Doxycycline (Dox) **B.** Timeline of Dox induction experiments. Cells were treated with Dox for 14 days and sub-cultured for up to 30 days after Dox removal. **C.** RT-qPCR analysis of *SERPINB5* expression during Dox treatment. At each time point, the expression in Dox treated cells was normalized to the expression in untreated cells cultured in parallel. Data are shown as the mean ± SEM (n= 3 biological replicates). **D.** Western blot analysis of SERPINB5 protein expression during Dox treatment. **E.** RT-qPCR analysis of SERPINB5 expression after Dox removal in TALE2-TET1-CD-WT-IRES-GFP cells. At each time point, the expression in Dox-treated cells was normalized to the expression in untreated cells. Data are shown as the mean ± SEM (n= 3 biological replicates). **F.** Western blot analysis of SERPINB5 protein expression after Dox removal in TALE2-TET1-CD-WT-IRES-GFP cells. **G.** Methylation level of *SERPINB5* promoter measured by amplicon NGS in inducible cells during and after Dox treatment (mean ± SEM, n= 3 biological replicates). **H.** Methylation profiles of single CpGs in TALE2-TET1-CD-WT-IRES-GFP cells after Dox treatment and Dox removal (mean ± SEM, n= 3 biological replicates). ns: not significant, *: p<0.05, **: p<0.01, ***: p<0.001 (unpaired t-test relative to untreated cells).

Supplementary information



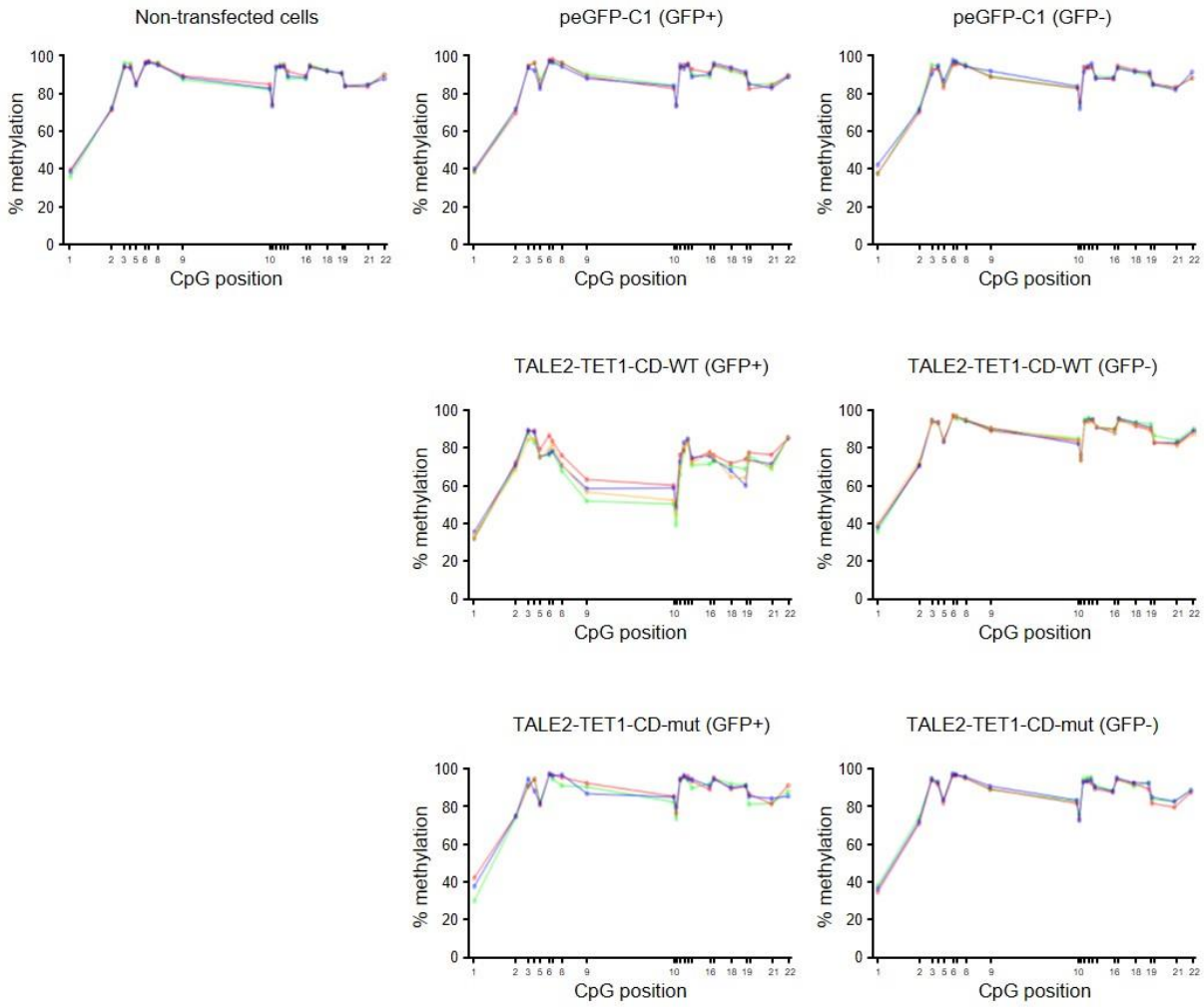
Supplementary Figure 1. *SERPINB5* expression is regulated by DNA methylation in MDA-MB-231 cells.

A. Cell proliferation measured for 72h in MDA-MB-231 cells treated with various concentrations of 5-aza-2'-deoxycytidine (5azadC). **B-D.** Expression measured by RT-qPCR of *SERPINB5*, *DAPK1* and *CDH1* in MDA-MB-231 cells treated with various concentrations of 5azadC (mean ±SEM, n=2 biological replicates). The expression was calculated as a ratio relative to the expression of three housekeeping genes (*RPL13A*, *SDHA*, *ACTB*).



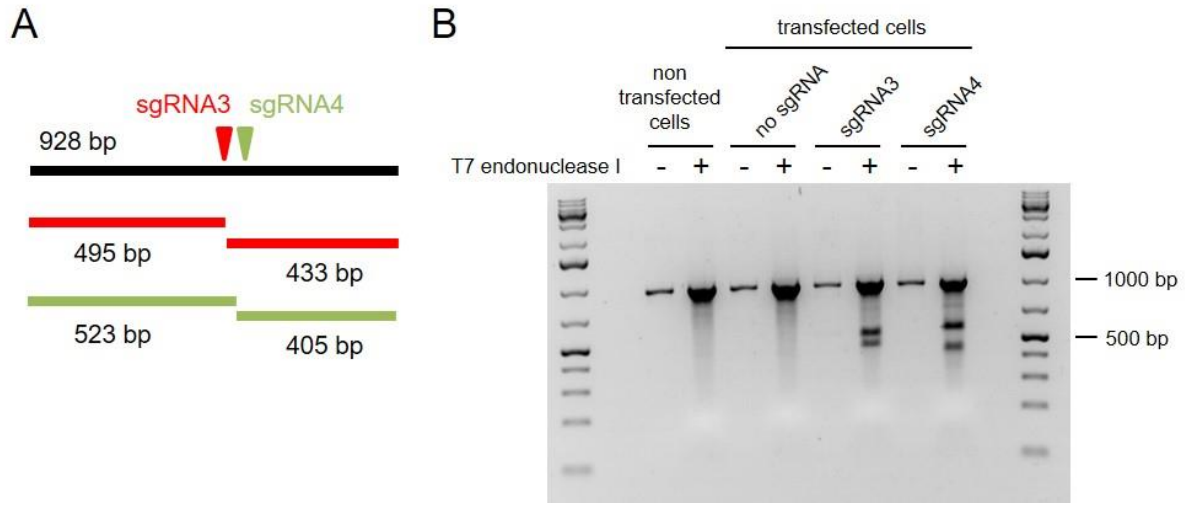
Supplementary Figure 2. TALE-mediated recruitment of AID does not strongly activate *SERPINB5* expression in MDA-MB-231 cells.

A. Schematic illustration of the vector used in the experiment. TALE2 was fused to the full length human AID protein. **B.** RT-qPCR analysis of *SERPINB5* expression in cells transfected with TALE2-hAID (mean ± SEM, n= 2 biological replicates). ns: not significant, *: $p<0.05$ (unpaired t-test relative to non-transfected cells). **C.** DNA methylation analysis of the *SERPINB5* promoter by COBRA in cells transfected with TALE2-hAID.



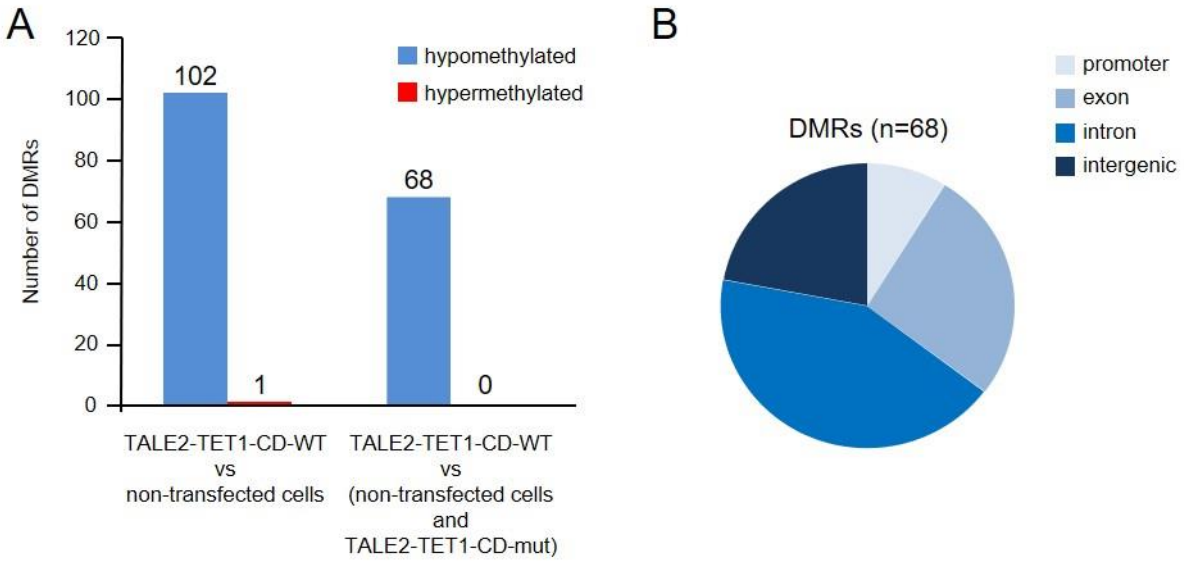
Supplementary Figure 3. CpG methylation profiles in the *SERPINB5* promoter by amplicon-NGS.

Graphs depicting the methylation scores at individual CpGs measured by amplicon NGS in non-transfected cells and cells transfected with GFP, TALE2-TET1-CD-WT or TALE2-TET1-CD-mut. For each condition, the different colors represent biological replicates (n=3 or 4).



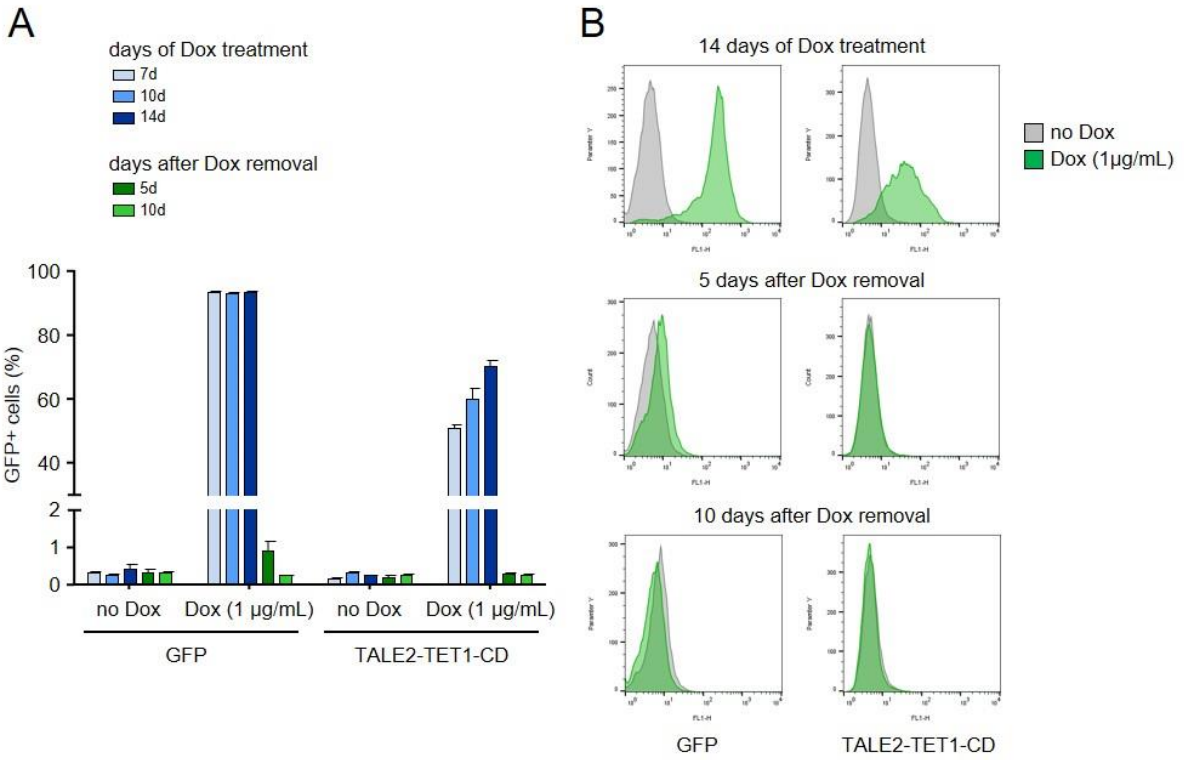
Supplementary Figure 4. Validation of the sgRNAs targeting the *SERPINB5* promoter by T7 endonuclease I assay.

A. PCR amplicon with the position of the sgRNAs and the predicted sizes of the cleavage products. **B.** T7 endonuclease I assay visualised by gel electrophoresis in cells transfected with sgRNA3 and sgRNA4 compared to the control cells. Both sgRNA3 and sgRNA4 lead to T7 endonuclease I digestion products.



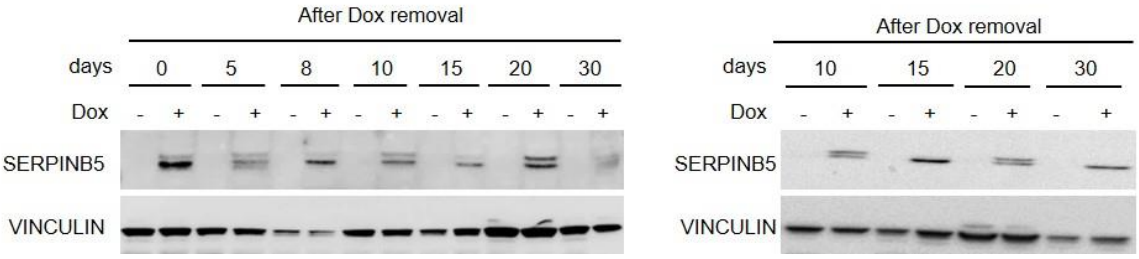
Supplementary Figure 5. Analysis of differentially methylated regions (DMRs) in cells expressing TALE2-TET1-CD.

A. Number of DMRs (>15%, p <0.01, t-test) identified by RRBS in cells expressing TALE2-TET1-CD-WT compared to non-transfected cells, or compared to non-transfected cells and cells expressing TALE-TET1-CD-mut. The numbers of DMRs are indicated on the graph. **B.** Pie chart representing the distribution of DMRs in promoters (-500 to +500 bp relative to TSS), exons, introns and intergenic regions.



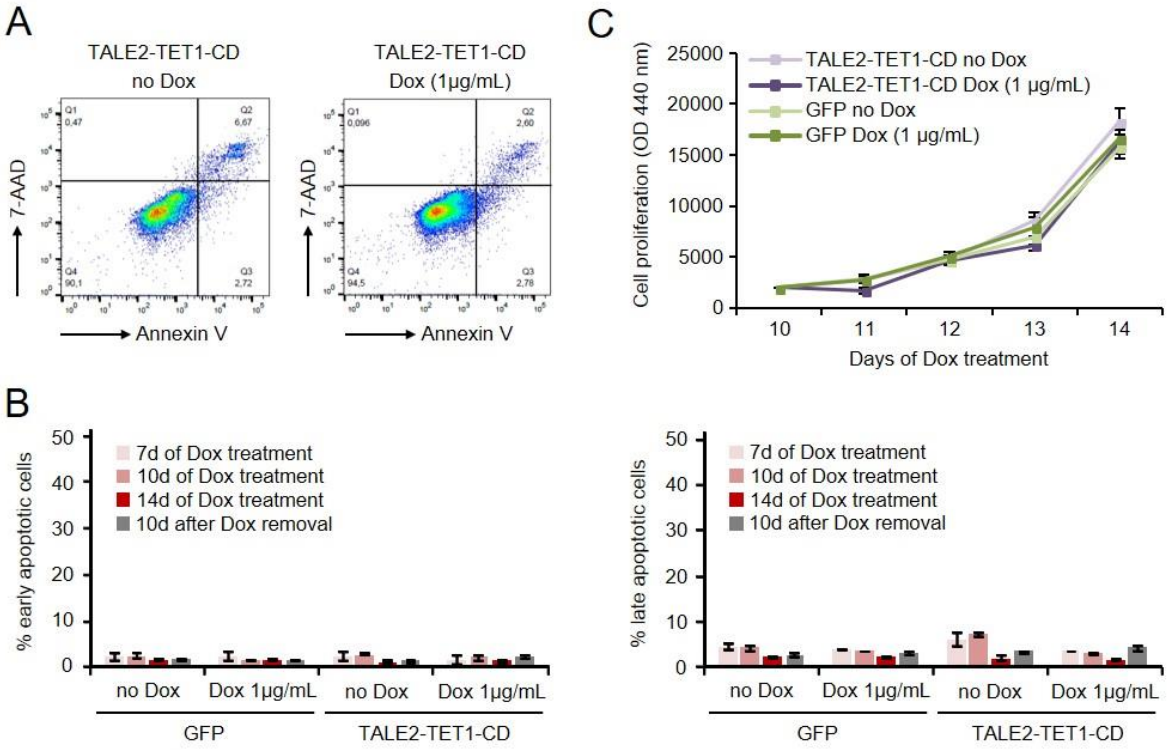
Supplementary Figure 6. Dox induction in MDA-MB-231 inducible cell lines.

A. Graph illustrating the percentage of GFP-positive cells measured by flow cytometry during Dox treatment and after Dox removal in GFP and TALE2-TET1-CD-IRES-GFP inducible cells lines. **B.** Flow cytometry plots of GFP expression after Dox treatment and Dox removal in GFP and TALE2-TET1-CD-IRES-GFP inducible cells lines.



Supplementary Figure 7. Durable expression of SERPINB5 protein after Dox removal in inducible TALE2-TET1-CD cells.

The figure shows western blots of SERPINB5 protein expression after Dox removal in two additional independent induction experiments in TALE2-TET1-CD-WT-IRES-GFP inducible cells. VINCULIN was used as a loading control.



Supplementary Figure 8. Impact of *SERPINB5* re-expression on apoptosis and cell proliferation of MDA-MB-231 cells.

A. Representative example of flow cytometry analysis of apoptosis (assessed by Annexin V and 7-AAD staining) in MDA-MB-231 cells expressing inducible TALE2-TET1-CD. Early apoptotic cells are in the lower right corner (positive for Annexin V only), and late apoptotic cells are in the upper right corner (double positive for Annexin V and 7-AAD). The numbers in quadrants indicate the percentages of cells. **B.** Bar graphs showing the percentage or early and late apoptotic cells in MDA-MB-231 cells expressing inducible TALE2-TET1-CD or GFP, quantified by flow cytometry after 7, 10 and 14 days of Dox treatment, and 10 days after Dox removal. Data are expressed as the mean ± SEM (n=3 independent experiments). **C.** Graph showing the proliferation (assessed by WST-1 Cell Proliferation Reagent) of MDA-MB-231 cells expressing inducible TALE2-TET1-CD or GFP between 10 and 14 days of doxycycline treatment. Data are expressed as the mean ± SEM (n=3 independent experiments).

Supplementary Table 1. List of DMRs.

A1:S57hromosor	start	end	width	PeakID	Ov_Rep	Rep_name	Ov_Exon	GeneName_Exon	Ov_Intron	GeneName_Intron	Ov_CGI	TSSpos	Dist to closest TSS	GeneName_TSS	num.CpGs	num.DMCs	5mC TET1wt-control	5mC TET1wt-TET1mut
chr1	1360122	1360185	64	chr1:1360122-1360185	0		0		1	MXRA8	0	1358794	-1360	MXRA8	6	5	-0,190	-0,177
chr1	223748986	223749294	309	chr1:223748986-223749294	0		1	CAPN2	1	CAPN2	1	223712416	36724	CAPN2	24	17	-0,202	-0,152
chr10	76405	76616	212	chr10:76405-76616	0		0		0		1	49238	-27272	TUBB8	13	10	-0,157	-0,154
chr10	67850283	67850474	192	chr10:67850283-67850474	0		0		0		0	67838179	-12199	DNAJC12	9	6	-0,155	-0,175
chr10	79152556	79152632	77	chr10:79152556-79152632	0		0		1	ZMIZ1	0	79069034	83560	ZMIZ1	6	5	-0,169	-0,161
chr11	792280	792352	73	chr11:792280-792352	0		1	SLC25A22	1	SLC25A22	1	790126	-2190	CEND1	6	4	-0,156	-0,157
chr11	63606681	63606802	122	chr11:63606681-63606802	1	LINE/L2/L2a	0		1	PLA2G16	0	63614437	7695	PLA2G16	11	6	-0,168	-0,186
chr11	66859684	66859789	106	chr11:66859684-66859789	0		1	LRFN4	1	PC	0	66857404	2332	LRFN4	6	6	-0,202	-0,209
chr12	104294258	104294377	120	chr12:104294258-104294377	1	LTR/ERVK/LTR13	0		1	TXNRD1	0	104286681	7637	TXNRD1	7	4	-0,161	-0,152
chr12	104303675	104304205	531	chr12:104303675-104304205	0		1	EID3	1	TXNRD1	1	104303731	209	EID3	44	37	-0,230	-0,226
chr13	20473235	20473273	39	chr13:20473235-20473273	0		0		1	CRYL1	0	20433846	-39408	MIR4499	4	4	-0,179	-0,168
chr13	111254529	111254623	95	chr13:111254529-111254623	0		0		1	ARHGEF7	1	111241144	13432	ARHGEF7	8	8	-0,352	-0,292
chr14	67738962	67739089	128	chr14:67738962-67739089	0		0		0		0	67701885	37141	RDH12	5	4	-0,311	-0,336
chr15	41585250	41585474	225	chr15:41585250-41585474	0		0		0		1	41559971	25391	TYRO3	21	15	-0,161	-0,164
chr15	72197119	72197239	121	chr15:72197119-72197239	0		0		1	GRAMD2	1	72197795	616	GRAMD2	10	8	-0,172	-0,193
chr16	555218	555308	91	chr16:555218-555308	0		0		0		1	560421	-5158	PRR35	14	14	-0,279	-0,268
chr16	1093586	1093939	354	chr16:1093586-1093939	0		1	C1QTNF8	0		1	1096244	2482	C1QTNF8	39	23	-0,159	-0,153
chr16	1413855	1414088	234	chr16:1413855-1414088	0		1	UNKL	1	UNKL	1	1414720	748	UNKL	17	7	-0,157	-0,181
chr16	2095771	2095819	49	chr16:2095771-2095819	1	LINE/L2/L2	0		1	PKD1	0	2091435	4360	LOC105371049	6	6	-0,291	-0,285
chr16	3881746	3881837	92	chr16:3881746-3881837	0		0		0		0	3880120	-1672	CREBBP	9	6	-0,179	-0,183
chr16	30418955	30419063	109	chr16:30418955-30419063	0		0		0		0	30429800	10791	DCTPP1	6	4	-0,209	-0,212
chr16	89910011	89910167	157	chr16:89910011-89910167	0		0		1	TCF25	0	89917878	-7789	MC1R	13	13	-0,181	-0,184
chr17	41810930	41811021	92	chr17:41810930-41811021	0		1	P3H4	0		0	41812199	1223	P3H4	6	5	-0,222	-0,229
chr17	42680196	42680236	41	chr17:42680196-42680236	0		1	CCR10	0		1	42681827	1611	CCR10	9	7	-0,201	-0,251
chr17	44210324	44210726	403	chr17:44210324-44210726	0		1	UBTF	1	UBTF	1	44207839	-2686	MIR6782	39	27	-0,176	-0,171
chr17	49035923	49035997	75	chr17:49035923-49035997	0		0		1	IGF2BP1	1	48997411	38549	IGF2BP1	7	6	-0,236	-0,199
chr17	75589536	75589641	106	chr17:75589536-75589641	0		1	MYO15B	0		1	75588057	1531	MYO15B	9	4	-0,215	-0,244
chr17	82242381	82242482	102	chr17:82242381-82242482	0		0		0		0	82236667	5765	MIR6787	7	6	-0,268	-0,244
chr17	82254904	82255161	258	chr17:82254904-82255161	0		0		1	CSNK1D	1	82236667	18365	MIR6787	16	10	-0,163	-0,162
chr18	63476692	63476760	69	chr18:63476692-63476760	0		0		0		0	63476910	-184	SERPINB5	5	4	-0,193	-0,219
chr19	1009044	1009223	180	chr19:1009044-1009223	0		1	GRIN3B	1	GRIN3B	1	1000418	8716	GRIN3B	16	15	-0,173	-0,170
chr19	1106628	1106698	71	chr19:1106628-1106698	0		1	GPX4	0		1	1104477	2186	GPX4	6	5	-0,207	-0,190
chr19	1245597	1245697	101	chr19:1245597-1245697	0		0		0		0	1248552	-2905	MIDN	8	8	-0,343	-0,343
chr19	3656171	3656597	427	chr19:3656171-3656597	0		1	PIP5K1C	1	PIP5K1C	0	3626815	-29569	CACTIN	21	14	-0,175	-0,161
chr19	4054049	4054558	510	chr19:4054049-4054558	0		1	ZBTB7A	0		1	4065529	11225	ZBTB7A	58	51	-0,177	-0,171
chr19	4369374	4369545	172	chr19:4369374-4369545	0		0		1	SH3GL1	0	4343526	25934	MPND	8	8	-0,208	-0,224
chr19	49793344	49793392	49	chr19:49793344-49793392	0		0		1	AP2A1	0	49791865	1503	MIR6799	5	4	-0,238	-0,227
chr19	55486371	55486705	335	chr19:55486371-55486705	0		1	NAT14	1	NAT14	1	55485189	1349	NAT14	21	13	-0,163	-0,153
chr2	10385354	10385611	258	chr2:10385354-10385611	0		0		1	HPCAL1	0	10368763	16719	HPCAL1	6	5	-0,176	-0,166
chr2	38536582	38536695	114	chr2:38536582-38536695	0		0		0		0	38515740	-20898	LOC101929596	7	5	-0,169	-0,176
chr2	120223014	120223041	28	chr2:120223014-120223041	0		1	TMEM185B	0		1	120223408	380	TMEM185B	9	9	-0,268	-0,273
chr2	202035101	202035345	245	chr2:202035101-202035345	0		1	FZD7	0		1	202034586	637	FZD7	20	16	-0,163	-0,163
chr20	4149758	4149848	91	chr20:4149758-4149848	0		0		1	SMOX	1	4148778	1025	SMOX	4	4	-0,161	-0,176
chr20	45836222	45836445	224	chr20:45836222-45836445	0		0		1	SNX21	0	45833830	2504	SNX21	14	10	-0,285	-0,280
chr20	62518840	62518944	105	chr20:62518840-62518944	0		0		0		0	62550452	-31560	MIR1-1HG	9	7	-0,197	-0,219
chr20	63686269	63686371	103	chr20:63686269-63686371	0		0		1	RTEL1--RTEL1-TNFRSF6B	1	63696650	-10330	TNFRSF6B	10	9	-0,194	-0,185
chr21	45328906	45328983	78	chr21:45328906-45328983	0		0		0		0	45341990	13046	LINC00316	7	6	-0,173	-0,166
chr21	45874547	45874608	62	chr21:45874547-45874608	0		0		1	PCBP3	0	45849960	24618	PCBP3	6	6	-0,288	-0,241
chr22	35447674	35447827	154	chr22:35447674-35447827	0		0		0		0	35400122	47628	MCM5	7	7	-0,255	-0,235
chr22	46102921	46103014	94	chr22:46102921-46103014	0		0		1	MIRLET7BHG	0	46112748	-9780	MIRLET7A3	5	4	-0,327	-0,319
chr22	50273531	50273586	56	chr22:50273531-50273586	1	LTR/ERVK/LTR13	0		0		0	50270393	-3165	MAPK11	8	8	-0,222	-0,221
chr22	50283540	50283684	145	chr22:50283540-50283684	0		1	PLXNB2	1	PLXNB2	1	50270393	-13219	MAPK11	5	5	-0,196	-0,205
chr3	18443714	18444093	380	chr3:18443714-18444093	0		1	SATB1	1	SATB1	1	18444118	214	SATB1	18	14	-0,179	-0,163
chr3	49902683	49903228	546	chr3:49902683-49903228	0		1	MST1R	0		1	49903637	681	MST1R	22	15	-0,218	-0,199
chr4	88457645	88457775	131	chr4:88457645-88457775	0		0		1	HERC5	1	88457116	594	HERC5	17	14	-0,234	-0,230
chr5	669842	670296	455	chr5:669842-670296	0		0		1	TPPP	0	693395	23326	TPPP	17	11	-0,153	-0,158
chr7	836014	836077	64	chr7:836014-836077	0		0		1	SUN1	0	832504	3542	SUN1	6	6	-0,175	-0,189
chr7	1027391	1027463	73	chr7:1027391-1027463	0		0		1	C7orf50	1	1028432	1005	C7orf50	9	7	-0,223	-0,210
chr7	2644808	2644828	21	chr7:2644808-2644828	0		0		1	TTYH3	0	2631968	12850	TTYH3	4	4	-0,349	-0,350
chr7	151795317	151795400	84	chr7:151795317-151795400	0		0		1	PRKAG2	0	151814871	19513	PRKAG2	4	4	-0,245	-0,231
chr8	22768830	22768875	46	chr8:22768830-22768875	0		0		1	PEBP4	0	22693302	-75550	EGR3	5	4	-0,211	-0,207
chr8	53882530	53882613	84	chr8:53882530-53882613	0		0		1	RGS20	1	53880871	1701	RGS20	8	8	-0,244	-0,237
chr9	75891106	75891313	208	chr9:75891106-75891313	0		1	PCSK5	0		1	75890643	567	PCSK5	21	16	-0,213	-0,214
chr9	107487177	107487334	158	chr9:107487177-107487334	0		1	KLF4	1	KLF4	0	107489720	2464	KLF4	9	8	-0,223	-0,216
chr9	107487929	107488216	288	chr9:107487929-107488216	0		1	KLF4	0		1	107489720	1648	KLF4	30	26	-0,215	-0,209
chr9	136577592	136577643	52	chr9:136577592-136577643	0		0		0		0	136547841	29777	NALT1	6	6	-0,312	-0,324
chr9	137514361	137514426	66	chr9:137514361-137514426	0		0		1	PNPLA7	0	137550569	36175	PNPLA7	4	4	-0,211	-0,193
chrX	154410095	154410219	125	chrX:154410095-154410219	0		0		1	DNASE1L1	0	154409278	-879	DNASE1L1	6	4	-0,234	-0,255

Primers for COBRA	
SERPINB5 Forward	TTTTGGAAGTTGTGTAGATAATAGTAATTTT
SERPINB5 Reverse	CCCTCTACAATATTCCTAATACTCAAAAA
Primers for bisulfite amplicon sequencing	
SERPINB5 amplicon1 Forward	NNNNTGTTGTAGTTTATATAAAAAGAATGGAGA (NNNN = tag)
SERPINB5 amplicon1 Reverse	ATCTACACAACTTCCAAAAAACCTC
SERPINB5 amplicon2 Forward	NNNNAATTGTGGATAAGTTGTTAAGAGGTTTG (NNNN = tag)
SERPINB5 amplicon2 Reverse	CTACCCCACCTTACTTACCTAAAATCAC
Primers for T7 endonuclease assay	
Primer gSERPINB5 F	CAGGCCTGAGTAATCCTAATCACAG
Primer gSERPINB5 R	GAACACTGCAGAGGGAGTAAAATC
Primers for RT-qPCR	
SERPINB5 primer F	GCCACGTTCTGTATGGGAAA
SERPINB5 primer R	TGGACTCATCCTCCACATCC
SDHA primer F	GCTCATGCATGTGTCCATGT
SDHA primer R	CAGCCACTAGGTGCCAATCT
RPL13A primer F	CCGAGAAGAACGTGGAGAAG
RPL13A primer R	GGCAACGCATGAGGAATTA
ACTB primer F	AGGTGTGGTGCCAGATTTTC
ACTB primer R	GGCATGGGTCAGAAGGATTC
DAPK1 primer F	GGCGAGCTGTTTGACTTCTT
DAPK1 primer R	GATTTGAAGGGAGTGCAGGT
CDH1 primer F	CTCGACACCCGATTCAAAGT
CDH1 primer R	GGCGTAGACCAAGAAATGGA
Primers for directed mutagenesis of hTET1-CD (H1672Y, D1674A)	
hTET1 Directed mutagenesis Forward	GACTTCTGTGCTCATCCCTACAGGGCCATTCACAACATGAATAA
hTET1 Directed mutagenesis Reverse	TTATTCATGTTGTGAATGGCCCTGTAGGGATGAGCACAGAAGTC
Primers for directed mutagenesis of *Bbs*I sites in hTET1-CD	
BbsI site 1 - Forward	GTCCCCGAAAACTGCTTCAGCCACACCAGCTCCACTGAA
BbsI site 1 - Reverse	GCTGAAGCAGTTTTCGGGGACCAGGAGAAGCCTGGAGATG
BbsI site 2 - Forward	CTTCGGGAAGGCTCAGTGGTGCCAATGCAGCTGCTGCTGAT
BbsI site 2 - Reverse	CACCACTGAGCCTTCCCGAAGGCATCGTACAGTGGGGAGTG

Supplementary Table 2. Primers used in the experiments.

Annex 1: Figures

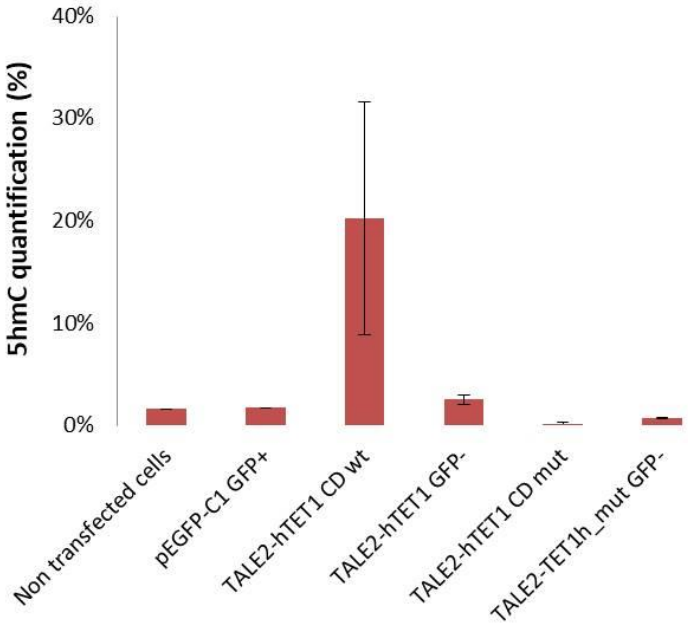


Figure Annex Manuscript 1. 5-hydoxymethylcytosine (5hmC) quantification in MDA-MB-231 cells.

5hmC quantification was performed with a EpiMark® 5-hmC and 5-mC Analysis Kit, consisting on glucosylation of DNA and restriction digestion with enzymes MspI and HpaII. The glucosylation step transfers glucose only onto 5-hydroxymethylcytosine (generating glucosylated 5-hydroxymethylcytosine [5-ghmC]) and protecting it from digestion. Both enzymes, MspI and HpaII, cut at CCGG site. MspI is not sensitive to methylation and cuts any time except in 5-ghmC sites. MspI cus 5mC or unmodified cytosine. In contrast, HpaII is sensitive to methylation, which means that is blocked by any of these modifications. HpaII only cuts in unmodified cytosine. We quantify the 5hmC % in *SERPINB5* promoter by qPCR analysis. We detect an increment in the 5hmC (around 20%) only in the cells transfected with the plasmid TALE2-hTET1 wt (GFP+).

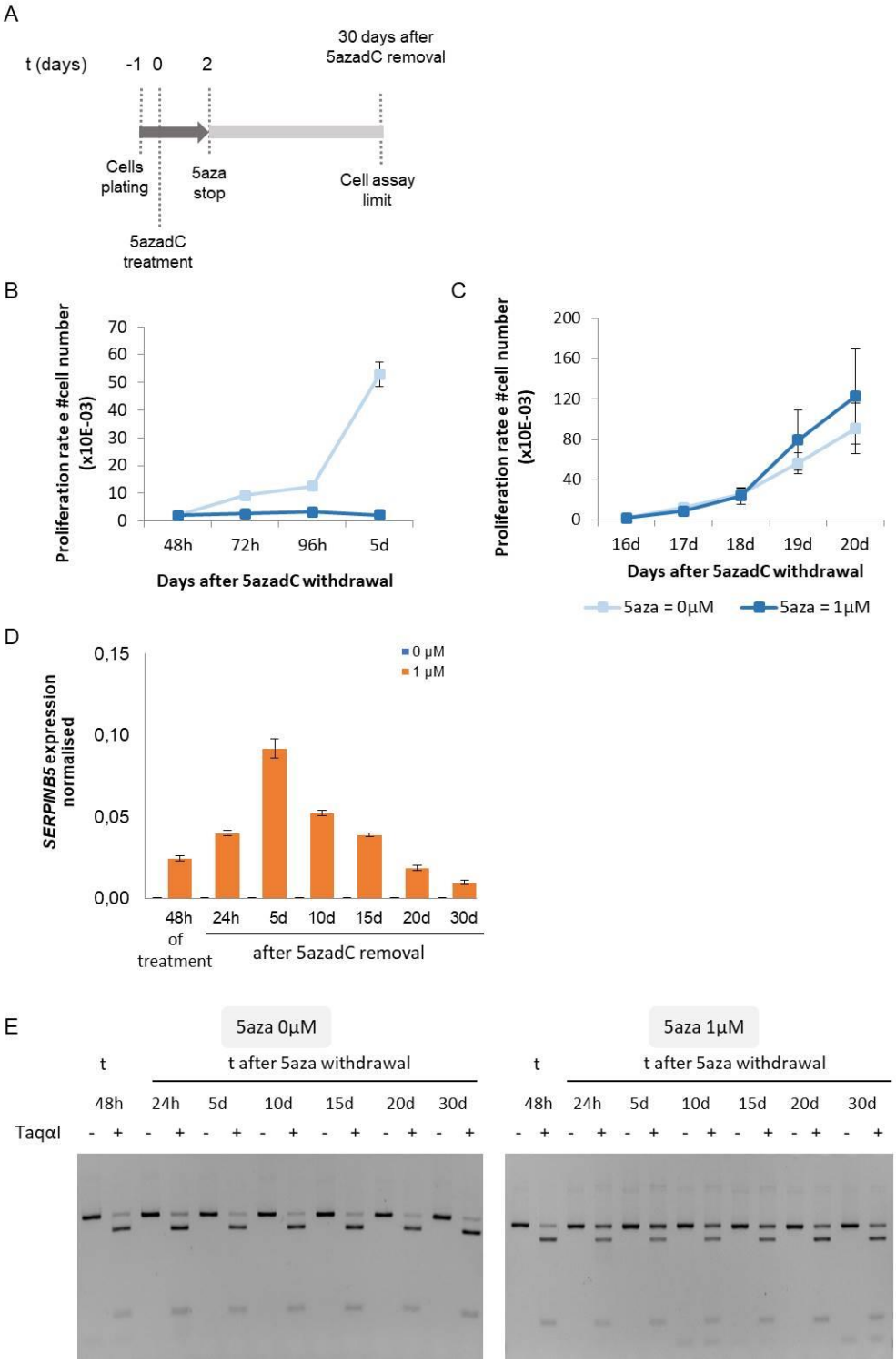


Figure Annex Manuscript 2. 5azadC traitement kinetics on MDA-MB-231.

A. Timeline of 5azadC treatment experiments. Cells were treated with 5azadC for 48h and sub-cultured for up to 30 days. **B.** Cell proliferation measured between 48h and 5d after 5azadC withdrawal. **C.** Cell proliferation measured between 16d and 20d after 5azadC withdrawal. **D.** RT-qPCR analysis of SERPINB5 expression during 5azadC treatment and after 5azadC removal. At each time point, the expression is normalised to three housekeeping genes (ACTB, RPL13A, SDHA). **E.** Methylation analysis in the SERPINB5 promoter by COBRA in control cells and cells treated with 5azadC. Data are shown as the mean ± SEM (n= 3 biological replicates).

RESULTS (2)

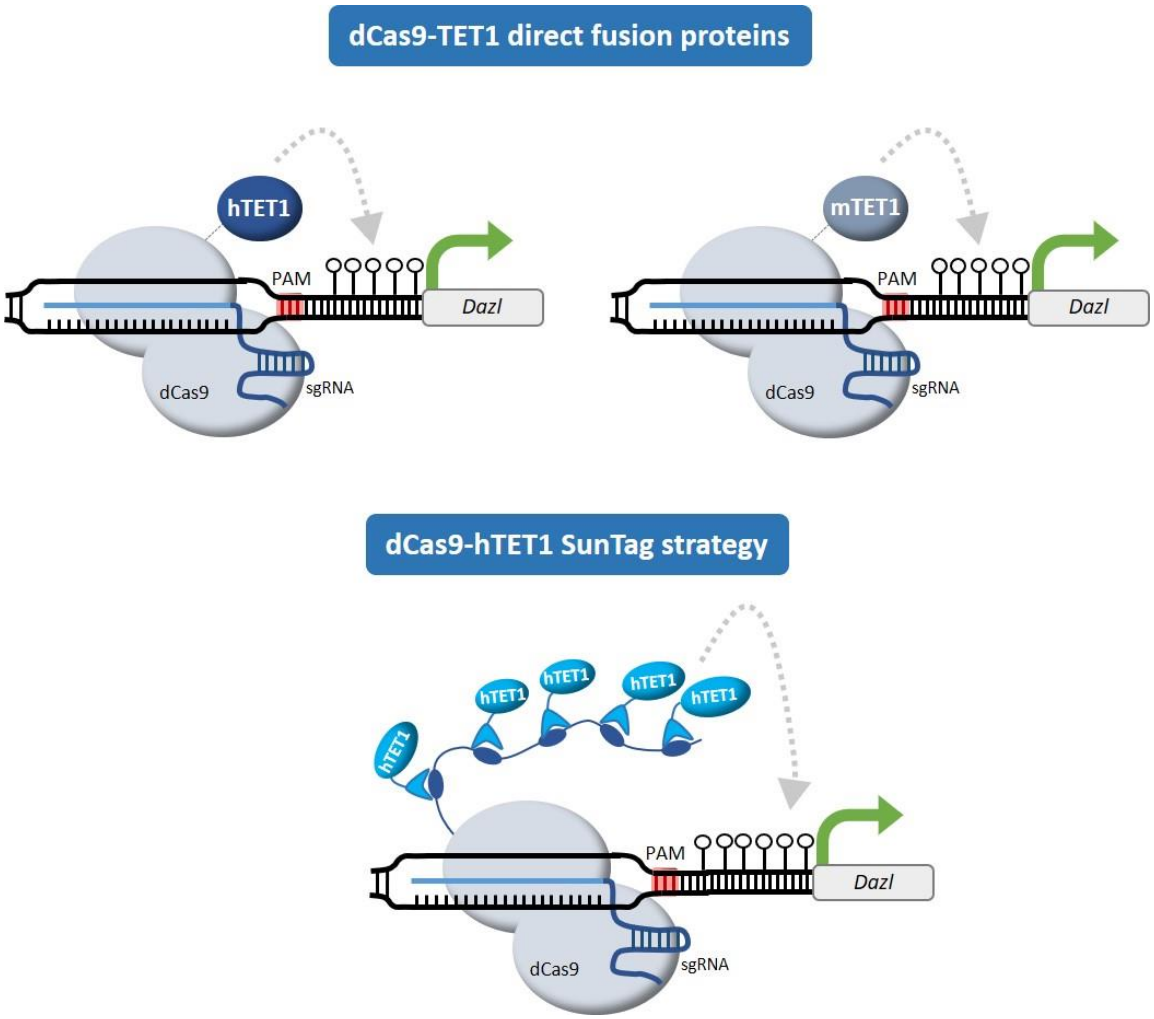
1 Objectives

During the second part of my thesis I have worked to optimize the dCas9 tool in order to improve the efficiency of the demethylation. To do so, we decided to compare different parameters and strategies on the same target genes in the same cellular system. The objective was to determine the best dCas9 strategy to perform efficient targeted methylation editing that can be implemented in routine in mouse and human cells to validate hypothesis concerning the functional role of DNA methylation in the genome.

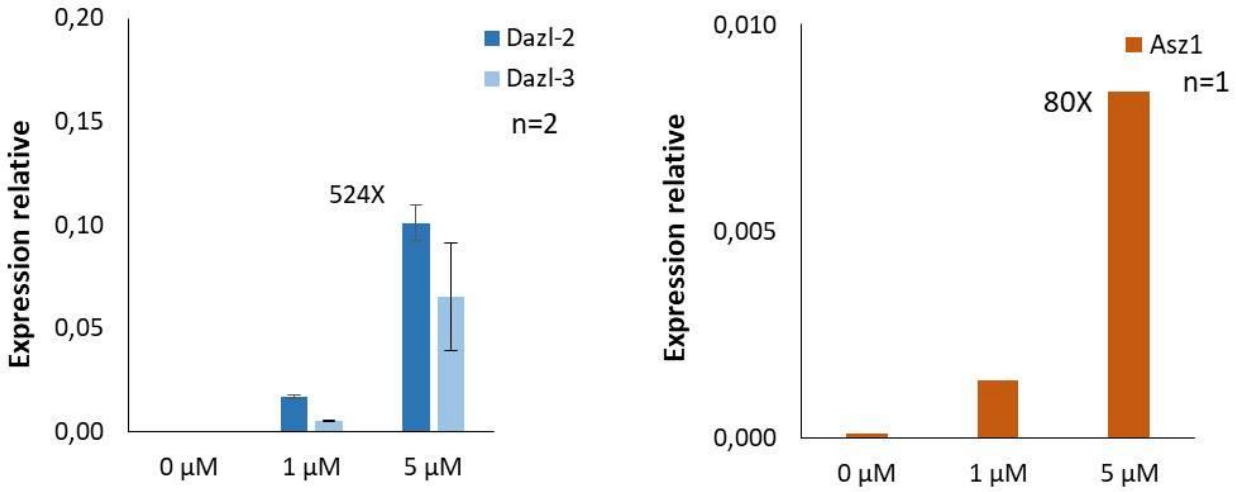
For this part I focused my work on mouse cells because this is the system most used in the laboratory to study the role of DNA methylation in genome regulation. We have chosen as models the gametic genes *Dazl* and *Asz1*. These genes are well known to be silenced by mechanisms implying promoter DNA methylation. Moreover, data from our laboratory demonstrated that *Dazl* and *Asz1* are amongst the genes most strongly de-repressed in mouse embryos (E8.5) KO for DNMT1 or KO for DNMT3A/3B that show a global loss of methylation, therefore they constitute good models to evaluate the efficiency of methylation editing.

After validating the sensibility of both genes to 5aza, I have constructed plasmids coding for direct fusion proteins dCas9-hTET1 and also the version with the catalytic domain of mouse TET1 (mTET1) dCas9-mTET1 (**Results Figure 1**). We have compared the efficiency of both tools and different gRNAs targeting *Dazl* or *Aszl* in Mouse Embryonic Fibroblasts.

Because we obtained poor efficiencies with direct dCas9-TET1 fusion proteins, we then wished to test a dCas9-hTET1 strategy based on the Suntag system. SunTag has been shown to improve the efficiency and the specificity of CRISPR-dCas9-based epigenome editing tools because it allows the recruitment of several TET1 molecules at the targeted site.

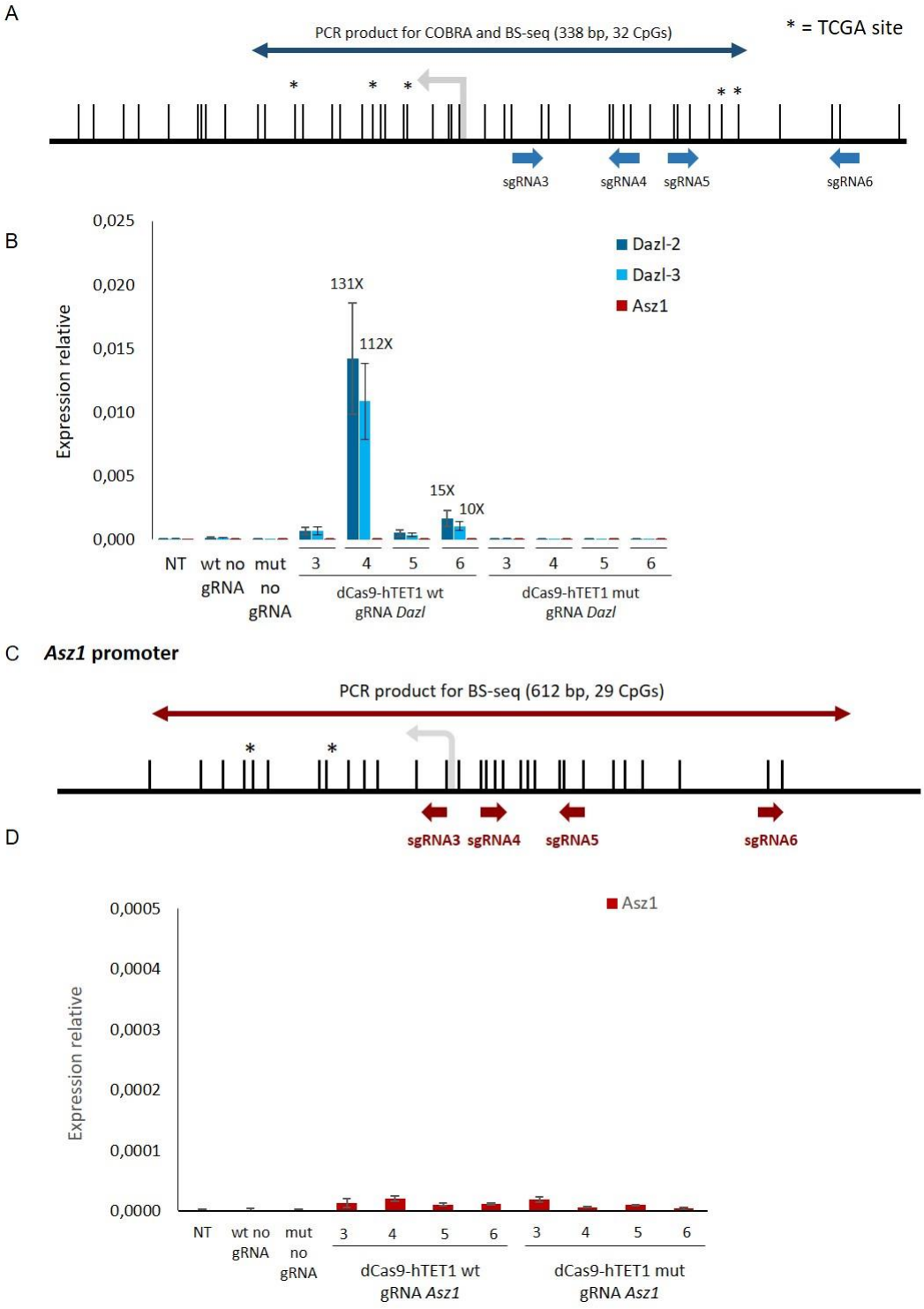


Results Figure 1. Strategies tested for efficient dCas9-based targeted demethylation in mouse cells.



Results Figure 2. *Dazl* and *Asz1* expression are regulated by DNA methylation in MEFs.

Dazl and *Asz1* mRNA expression was analysed by RT-qPCR in MEF cells treated for 96h with various concentrations of 5aza (1 or 5 µM). Expression of *Dazl* was measured with two pairs of primers in exon 11 (last exon, Dazl-2) and exon 2 (Dazl-3). The expression was normalized with three housekeeping genes for Dazl (*Hprt*, *Mrpl32*, *Rpl13a*) (mean±SEM, n=2) and two for *Asz1* (*Rpl13a*, *Gus*) (n=1).



Results Figure 3. dCas9-hTET1 upregulates *Dazl* but no *Asz1*.

A. Schematic illustration of the *Dazl* promoter region and the sgRNAs binding sites. The vertical bars represent the CpGs. The region amplified by PCR for COBRA and Bisulfite sequencing (BS-Seq) is represented in dark blue. The TCGA motifs for use in COBRA (*TaqαI* restriction site) are signalled with an asterisk (*).
B. *Dazl* and *Asz1* mRNA expression in MEF cells transfected with dCas9-hTET1 targeting *Dazl* promoter (sgRNAs 3-6). Analysis by RT-qPCR was performed 48h after puromycine selection (72h after transfection). The expression was normalized with three housekeeping genes (*Rpl13a*, *Gus* and *B2m*) (mean ±SEM, n=3). For cells transfected with dCas9-hTET1 wt gRNA-4 and 6, the fold change relative to cells transfected with dCas9-hTET1 wt without gRNA is indicated.
C. Schematic illustration of the *Asz1* promoter region and the sgRNAs binding sites. The vertical bars represent the CpGs.
D. *Asz1* mRNA expression in MEF cells transfected with dCas9-hTET1 targeting *Asz1* promoter (sgRNAs 3-6). Analysis by RT-qPCR was performed as described in B.

2 Results

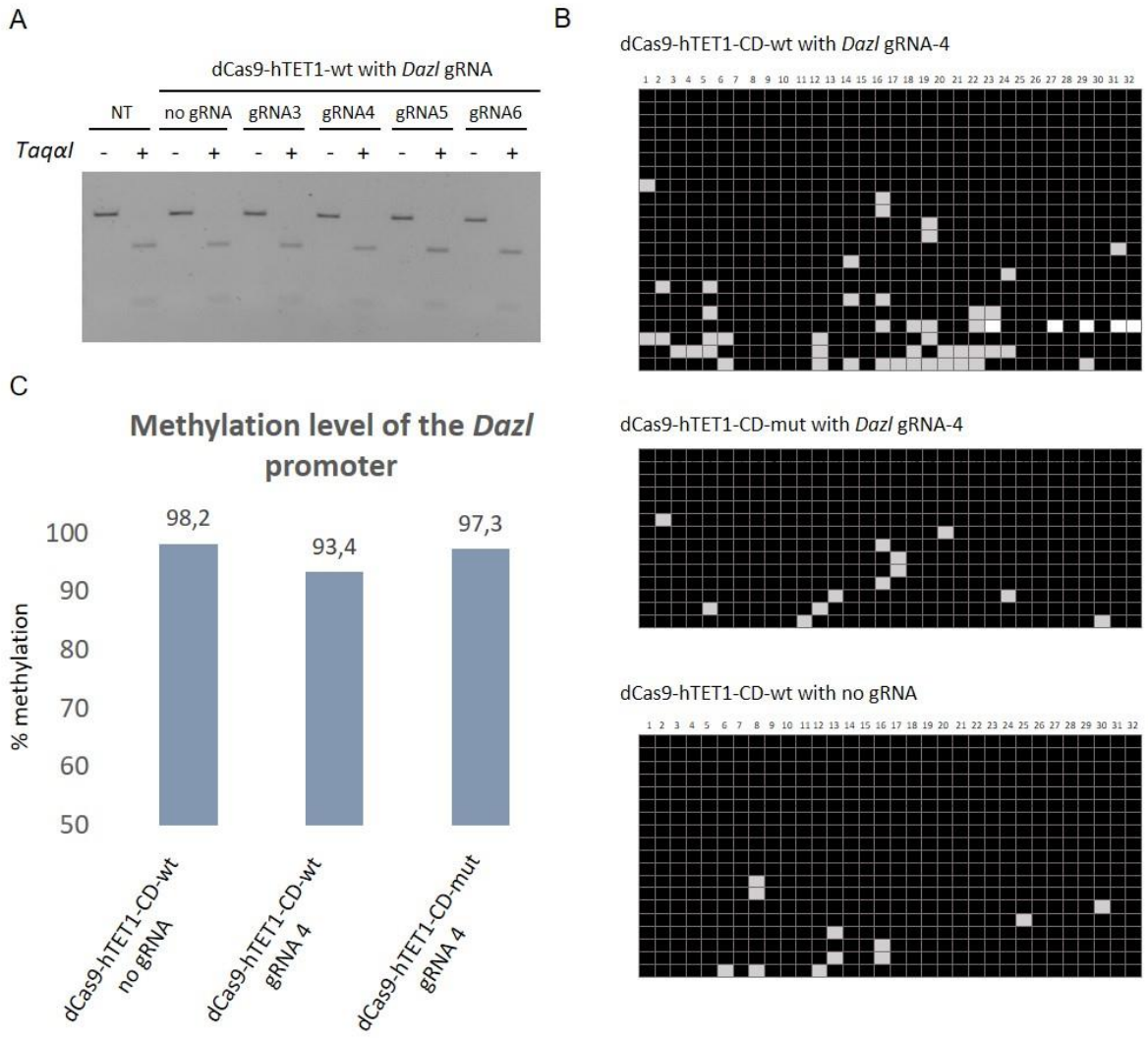
Dazl and Asz1 have methylation-sensitive promoters

Dazl and *Asz1* are gametic genes that are well known to be silenced by DNA methylation. To confirm their methylation-dependent expression in our MEF system, MEFs were treated with the DNA methyltransferase inhibitor 5-aza-2'-deoxycytidine (5azadC) at 1 or 5µM for 96 hours. I found that *Dazl* and *Asz1* expressions are strongly upregulated in the cells treated with 5azadC compared to non-treated cells, this upregulation exceeding 500-fold for *Dazl* and 80-fold for *Asz1* in cells treated with 5µM (**Results Figure 2**). That confirms that both genes have a methylation-sensitive promoter and thus are good candidates for our studies.

Efficiency of dCas9-hTET1 on Dazl and Asz1 expression

I tested the ability of CRISPR-dCas9 fused to the human TET1 catalytic domain to induce epigenetic reactivation of *Dazl*. I first transfected MEF cells with the plasmid coding for the catalytically inactive Cas9 (dCas9) fused to the human catalytic domain of TET1 (dCas9-hTET1-CD-wt) or the catalytically dead mutant (dCas9-hTET1-CD-mut) together with a sgRNA targeting the *Dazl* promoter. We tested four gRNAs targeting *Dazl* promoter independently (**Results Figure 3A**) as well as a condition without sgRNA as control. Cells were selected for puromycin resistance during 48h and harvested to measure *Dazl* expression by RT-qPCR. Our results indicate that, remarkably, only one sgRNA (sgRNA4) was able to induce *Dazl* expression up to 131 times (using primer Dazl-2) and 112 times (primer Dazl-3) compared to the control cells. In contrast, the gRNAs 3, 5 and 6 either had a weak effect or did not induce a significant derepression of *Dazl* expression (**Results Figure 3B**). To analyse if the effect of dCas9-hTET1-CD can be extrapolated to other genes, we targeted dCas9-hTET1-CD to the *Asz1* promoter with 4 different gRNAs (**Results Figure 3C**). Unfortunately, we observed that none of the gRNA was able to induce a strong expression of *Asz1* (**Results Figure 3D**).

Taken together, we conclude that the dCas9-hTET1 effect is largely dependent on the gRNA and the susceptibility of a gene to be epigenetically reprogrammed, and that the efficiency of the dCas9-hTET1 system is generally low. This supports our previous conclusions obtained with dCas9-hTET1 on the *SERPINB5* gene in MDA-MB-231 cells (see part 1 of the thesis).



Results Figure 4. dCas9-hTET1 induced demethylation of *Dazl* promoter.

A. Methylation analysis in *Dazl* promoter by COBRA.
B. Methylation analysis in *Dazl* promoter by bisulfite sequencing. Amplicon used for BS-seq contains 32 CpGs. Each row constitutes the sequence of an individual clone and is formed by squares representing methylated cytosines (black) and unmethylated (grey).
C. Global methylation in *Dazl* promoter calculated from BS-seq data.

dCas9-hTET1 induced demethylation of Dazl promoter is not efficient

We then evaluated the demethylation achieved by dCas9-hTET1 on *Dazl* promoter with the sgRNA4. First, we performed Combined Bisulfite Restriction Analysis (COBRA) to evaluate cytosine methylation in TCGA sites located close to the TSS of *Dazl* (**Results Figure 3A**). The COBRA results suggest that loss of DNA methylation induced by any of the gRNA, even gRNA4, is minimal (**Results Figure 4A**). However, the results are limited to the TCGA sites and could be biased by the position of the restriction sites, therefore we also performed bisulfite sequencing (BS-seq) on the same amplified region. We analysed 19 clones for dCas9-TET1-CD-wt no gRNA, 22 clones for dCas9-hTET1-CD-wt gRNA4 and 14 clones for dCas9-hTET1-CD-mut gRNA4. As shown in **Results Figure 4B and 4C**, the demethylation induced by the gRNA4 is detectable but minimal. Compared with the cells transfected with the plasmid without sgRNA, the loss of methylation is around 5% (**Results Figure 4C**). Interestingly, the targeted region of gRNA4 is covering the CpGs 23-26 and the strongest levels of demethylation are observed for CpGs just downstream of the CpG 23-26 (**Results Figure 4B**). Further experiments based on NGS amplicon sequencing should be carried in order to more accurately quantify the level of demethylation.

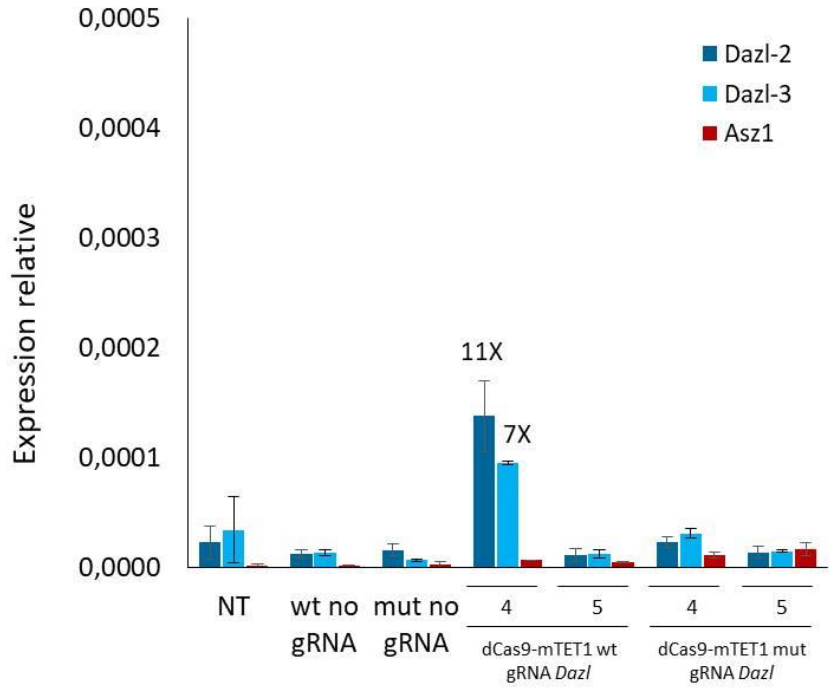
Comparison of dCas9-hTET1 and dCas9-mTET1 on Dazl expression

Because we obtained low demethylation and reactivation with dCas9 fused to the human TET1, I tested if fusing dCas9 to the catalytic domain of mouse TET1 (mTET1-CD) improves the efficiency of the system. I cloned novel plasmids by replacing hTET1-CD with mTET1-CD and transfected MEFs with the plasmids dCas9-mTET1-CD (wt or mutant) containing the gRNA4 targeting the *Dazl* promoter, that efficiently reactivated *Dazl* expression when combined with dCas9-hTET1, or the gRNA5 targeting the *Dazl* promoter. Surprisingly, the effect achieved with dCas9-mTET1 and gRNA4 was much weaker compared to the dCas9-hTET1 (11 fold-change with primer Dazl-2) (**Results Figure 5**). We conclude that mouse TET1 is less efficient than human TET1 for dCas9-based targeted demethylation, even in mouse cells.

Comparison of dCas9-hTET1 and dCas9-Suntag systems

Having observed low efficiency of direct dCas9-hTET1 fusion systems, the next step was to test the dCas9-suntag system on *Dazl* and *Asz1* in MEFs. For this, I obtained an all-in-one plasmid coding for dCas9-multiGCN4-2A-scFv-GFP-hTET1-CD (dCas9-Suntag) and cloned the gRNAs targeting *Dazl* and *Asz1* in this plasmid.

First of all, we tested the gRNA-4 targeting *Dazl* promoter because it was the most efficient gRNA in the experiments using direct fusion proteins. We used the cells transfected with the plasmid



Results Figure 5. Mouse TET1 is less efficient than human TET1 for dCas9-based targeted demethylation.

Dazl and *Asz1* mRNA expression on MEF cells transfected with dCas9-mTET1 targeting *Dazl* promoter. Analysis by RT-qPCR 48h after puromycine selection (72h after transfection). The expression has been normalized with three housekeeping genes (*Rpl13a*, *Gus* and *B2m*) (mean ±SEM, n=3). For cells transfected with dCas9-mTET1 wt gRNA-4, the fold change relative to cells transfected with dCas9-mTET1 wt without gRNA is indicated.

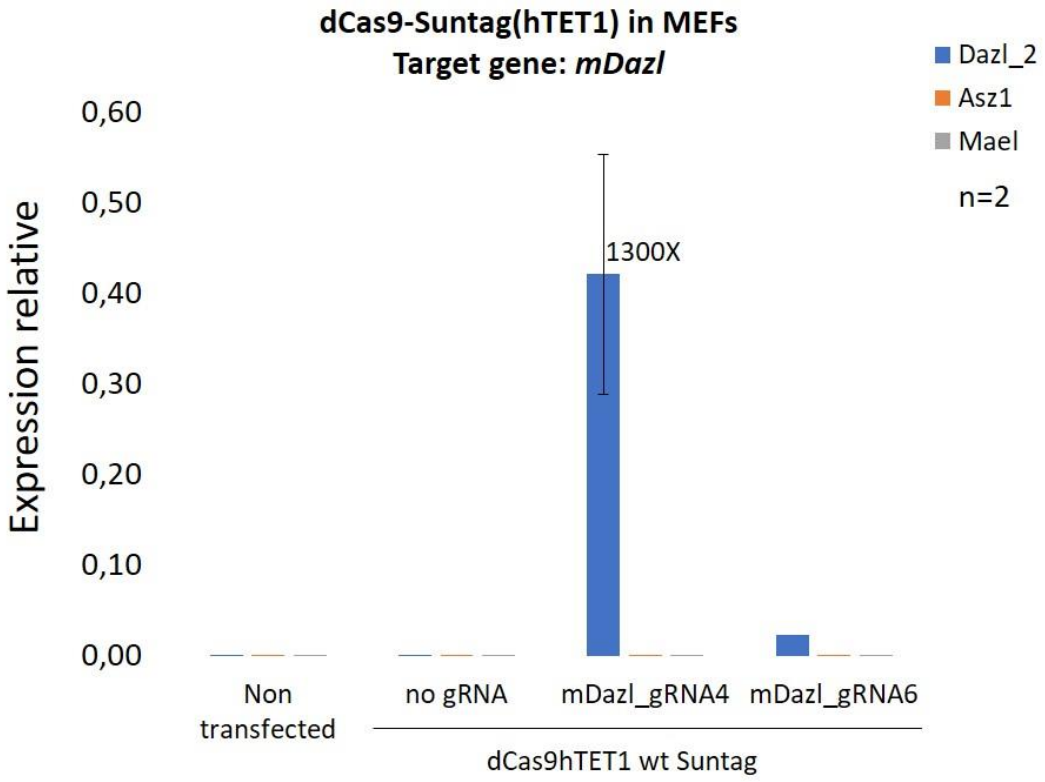
dCas9-Suntag without gRNA as a control. Cells were sorted by FACS 72h after transfection (GFP+ and GFP- cell populations). We measured *Dazl*, *Asz1* and *Mael* expression by RT-qPCR. As shown in **Results Figure 6**, dCas9-Suntag induced a remarkable reexpression of *Dazl* compared to the direct fusion protein dCas9-hTET1 (see compilation of relative expressions of *Dazl* in **Results Table 1**). This induction was increased by 1300 times if compared to the cells transfected with the plasmid without gRNA (**Results Figure 6**). The other studied genes expression is not affected, which suggests a specific effect of the Suntag-hTET1 construction. We analysed the demethylation achieved on *Dazl* promoter by COBRA (**Results Figure 7A**) and Bisulfite sequencing (**Results Figure 7B and C**). The results suggest that the promoter of *Dazl* is successfully demethylated after transfection with dCas9-Suntag *Dazl* gRNA-4, whereas the promoter of other methylated gene such as *Asz1* is not affected (**Results Figure 7B and C**). The methylation level on *Dazl* promoter goes from 98,6% in the non-transfected cells to 36% in the cells transfected with dCas9-Suntag *Dazl* gRNA-4, which means that *Dazl* promoter undergoes a loss of methylation of around 63%. As these are preliminary results, we are actually repeating the experiments and testing other target genes such as *Asz1* and *Mael*. The high reexpression and demethylation obtained are promising results to validate the dCas9-Suntag as an efficient tool to perform epigenome editing in MEF cells.

3 Conclusion and perspectives

This work has validated the usage of dCas9-hTET1-based tools to induce a targeted demethylation in mouse cells. Neverthelss, further studies are necessary to optimize the tool and conclude which experimental conditions are the most efficient. For this reason, a systematic study of various parameters in the same cellular and gene context is necessary.

As we have measured the demethylation rate at 48h or 72h post transitory transfection, we could test if harvesting the cells at later time points would increase the efficiency of demethylation. We are also currently testing the dCas9-Suntag system to compare its efficacy with dCas9-TET1.

To follow on this work, it will be also interesting to compare the genome wide specificity of the dCas9-TET1 and dCas9-Suntag systems by RRBS, as well as test if generating stable cell lines improves the efficacy compared to transitory transfections.



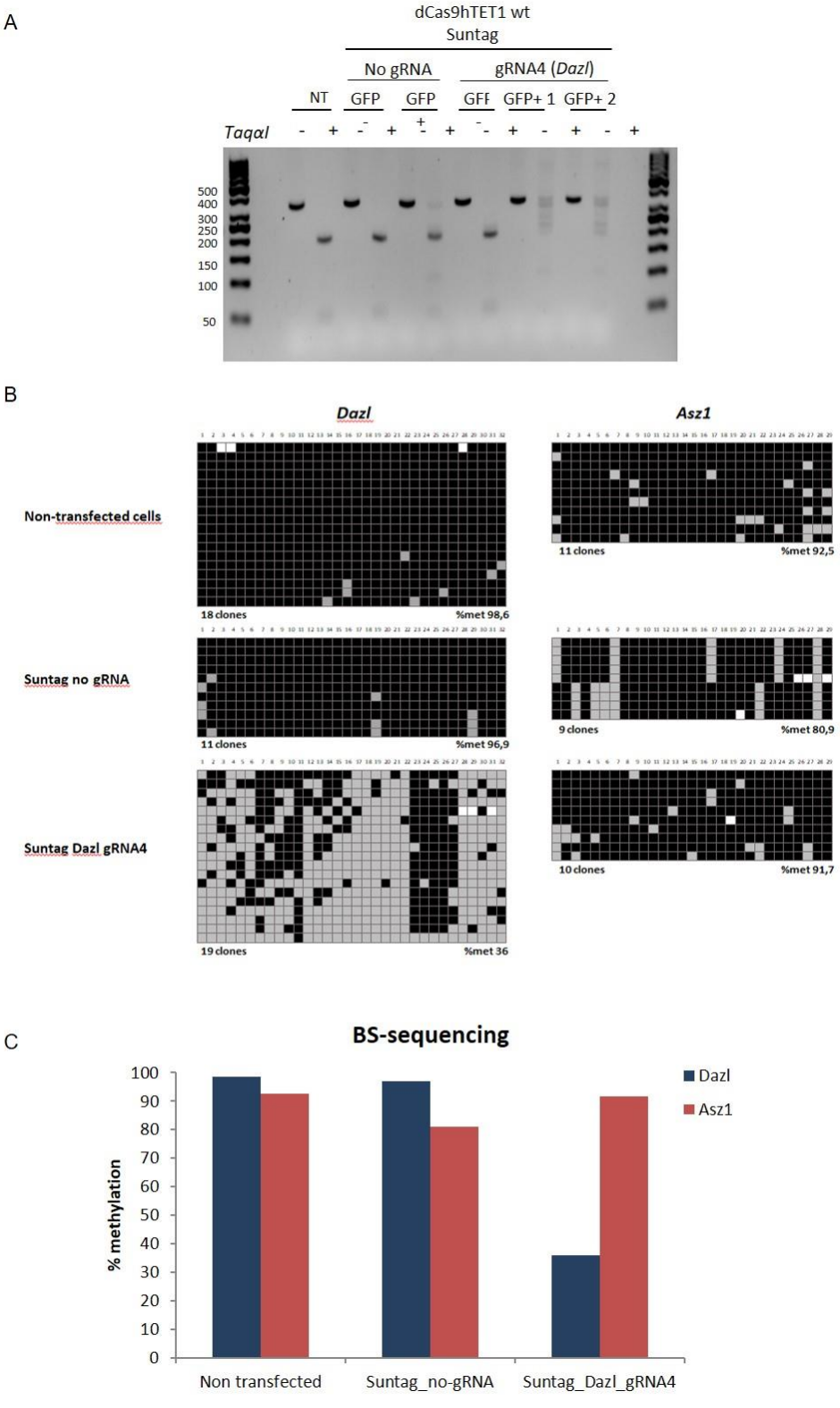
Results Figure 6. dCas9-Suntag (hTET1) induces *Dazl* expression more efficiently compared to the direct fusion protein dCas9-hTET1.

Dazl, *Asz1* and *Mael* mRNA expression on MEF cells transfected with with pPlatTETgRNA2 (Suntag hTET1) plasmid, without gRNA or with gNRA number 4 or 6 targeting *Dazl* promoter. Analysis by RT-qPCR 72h after GFP FACS sorting. The expression has been normalized with three housekeeping genes (*Rpl13a*, *Gus* and *B2m*) (mean ±SEM n=2 for gRNA4, n=1 for gRNA6). For cells transfected with Suntag hTET1 *Dazl* gRNA-4, the fold change relative to cells transfected with Suntag hTET1 without gRNA is indicated.

	Relative expression *Dazl*
5azadC	0,101
dCas9-hTET1	0,0142
dCas9-mTET1	0,0001
dCas9-Suntag (hTET1)	0,4209

Results Table 1. Relative expression of *Dazl* after 5azadC treatment and transfection with methylation editing tools.

Dazl relative expression in different experiments compared to housekeeping genes.



Results Figure 7. dCas9-Suntag (hTET1) induces *Dazl* demethylation more efficiently compared to the direct fusion protein dCas9-hTET1.

A. Methylation analysis in *Dazl* promoter by COBRA
B. Bisulfite sequencing. Methylation status of *Dazl* and *Asz1* promoters (control) in MEF cells non transfected and transfected with pPlatTETgRNA2 (Suntag hTET1) plasmid, without gRNA or with gNRA number 4 targeting *Dazl* promoter. Each line represents an individual clone and each column a CpG in the amplicon (black = methylated cytosine, grey = unmethylated cytosine).
B. Quantification of the methylation obtained by Bisulfite sequencing.

DISCUSSION

DISCUSSION

Impact of the research

In this Thesis, I have made several contributions to the field of epigenetic editing.

In the first part of the project we have compared two epigenome editing tools (TALE and CRISPR-dCas9) coupled to hTET1 and we have evaluated their efficiency to demethylate a model tumor suppressor gene, *SERPINB5*, which is methylated and silenced in aggressive tumors. We have corroborated the correlation between promoter methylation and gene repression in this gene. Furthermore, we have developed an inducible TALE-hTET1 system allowing controlled *SERPINB5* promoter demethylation in breast cancer cells. This technique allowed us to elucidate the temporal kinetics of promoter DNA demethylation and investigate the durability of the DNA methylation reprogramming and its long-term impact on gene expression at the *SERPINB5* locus. Thanks to this system, we report for the first time a long-lasting reprogramming of gene expression of a tumor suppressor gene through DNA promoter demethylation in an inducible system.

I also tested for the first time the possible use of AID in DNA demethylation through targeted epigenome editing. We were not able to demonstrate any strong effect of tethering AID on demethylation nor expression of *SERPINB5* in our cellular model. Indeed, the role of AID in the cytosine demethylation pathway is still controversial, and our results do not support the idea that AID can initiate DNA demethylation in MDA-MB-231 cells.

It is well known that tumor suppressor genes present an aberrant promoter CpG methylation that leads to the repression of gene expression in cancer cells. Until now, different groups have evaluated the power of re-expression of tumor suppressor genes and its phenotypical consequences (Beltran et al., 2007; Choudhury et al., 2016; Huisman et al., 2016). We wanted to go further by demonstrating that it is possible to use epigenetic editing to stably reprogram the expression of a tumor suppressor gene in cancer cells. As systems based in genome targeting are rapidly being developed in therapeutics, our work is a proof-of-concept to show that it is possible to redesign single gene expression in a durable way. These results will pave the way towards the development of novel innovative therapeutic approaches for the treatment of cancers.

Comparison of TALE and CRISPR-dCas9 tools

Over the last years, we have witnessed a fast progress in the development of epigenome editing tools first based on programmable DNA binding proteins (Zinc-Fingers and TALEs) and more recently based on the sgRNA-based CRISPR-dCas9. All these systems have been coupled to different

	ADVANTAGES	DISADVANTAGES
ZFPs	* Succesfully used in combination with multiple effector domains * Less immunogenic because existing in eukaryotic cells: important for clinical applications * Smaller size: efficient delivery into cells * Efficient in some regions were dCas9 does not show any effect (e.g. methylated regions) * Approved for use in clinical trials	* Protein-based: laborious synthesis, time-consuming * Not flexible design * Off-targets
TALEs	* Succesfully used in combination with multiple effector domains * Highly efficient and specific * Compared to ZFP: one module recognises one single base instead of three, giving more flexibility to sequence design and increased specificity * Approved for use in clinical trials	* Protein-based: laborious synthesis, time-consuming * Tandem repeats: amplification and cloning laborious, susceptibility to DNA rearrangements * Not flexible design * Plasmids of big size: complicated delivery into some types of cells * Limitations when choosing target sequence: presence of a T in the 5' end before the target sequence * Can not bind to methylated sequences * Immunogenicity (engineered from bacterial factors)
CRISPR -dCas9	* Succesfully used in combination of multiple effector domains * Flexibility, multiplexing, target design simplicity * Cost and time effective * Possibility of engineering different parts of the tool (sgRNAs, dCas9) to improve the efficiency and specificity (e.g. Suntag, aptamers)	* Nucleotide-based system: Off-target effects, they can be overcome with improved tools * Limitations when choosing target sequence: necessity PAM sequence * Plasmids of big size: complicated delivery into some types of cells or *in vivo* delivery * Immunogenicity (engineered from bacterial factors)

Discussion Table 1. Advantages and disadvantages of the different DNA binding domains

effector domains in order to efficiently perform epigenome editing at various levels: histones acetylation, histones methylation, DNA methylation and demethylation (**See Introductory Table 6**). Even if all strategies have been shown to efficiently induce epigenome changes, CRISPR-dCas9 is simpler, more flexible and less expensive. On the other hand, the choice of the target sequence is limited because of the mandatory presence of a PAM motif just downstream of the target sequence. TALE proteins give a larger choice in target sequences. Nevertheless, we have to take into consideration that they require time-consuming cloning and that they can be sensitive to DNA methylation (Bultmann et al., 2012; Deng et al., 2012b) unless they are built according to codes that allow to overcome this disadvantage (Valton et al., 2012).

In our project, we designed TALE and CRISPR sgRNAs targeting the same region in the CpG island of the *SERPINB5* promoter and observed a higher efficiency of demethylation and reactivation of the gene using the protein-based TALE system. We tested 2 different TALE proteins and 3 sgRNAs targeting the *SERPINB5* promoter. In our tests of TALE-VP64 fusion proteins, TALE2 was the only one able to reactivate *SERPINB5*. A possible explanation is that the TALE1 binding site is located too far away from the TSS. Concerning CRISPR-dCas9, we found that it was less powerful than TALE-based system for all gRNA tested. The two gRNAs not showing any effect on *SERPINB5* re-expression (sgRNAs 3 and 4) still efficiently induced DNA breaks in the promoter with the Cas9, as shown by the T7 endo I assay. TALE2-hTET1 induced reactivation is less potent than TALE2-VP64. This was expected because VP64 directly recruits the transcriptional activator complex. This advantage of the TALE system compared to CRISPR-dCas9 when coupled to transcription activators or repressors had already been reported by others (Gao et al., 2014b; Garcia-Bloj et al., 2016; Zhang et al., 2015b). However, in another report, dCas9 fused to epigenetic activators like p300 was shown to achieve a more robust activation than TALE (Hilton et al., 2015). Concerning the demethylation effectors, only one study compared both tools in the same gene, concluding that CRISPR-dCas9 fused to hTET1 promoted a higher demethylation but in a narrower region than TALE-hTET1 (Liu et al., 2016b). Interestingly, it has been shown that dCas9 binding efficacy is compromised when targeting hypermethylated sequences (Cano-Rodriguez et al., 2016), which could explain that the dCas9 system has a lower efficiency in demethylation editing. Some other comparisons among the systems can be found in **Discussion Table 1**. In summary, our data and the published literature suggest that TALEs can be more efficient than dCas9 for targeted epigenetic reprogramming. This is probably dependent on the chromatin context of each loci, the epigenetic effector and the cellular context.

Altogether, these results show a chromatin and gene dependent efficiency. As there are a very few examples of direct comparison of different targeting tools in the literature, further experiments are necessary to clarify this issue. For that, a project comparing all the available DBDs (TALE, dCas9, SunTag…) fused to the different TET domains and targeting the same region in the same cellular

context is required. Only this way we would be able to compare the efficiency and efficacy of the tools.

Comparison of hTET1 and mTET1 for targeted demethylation

In the second part of the project we compared the effects of dCas9-hTET1 and dCas9-mTET1 on methylation and expression of the mouse *Dazl* promoter. We found that mouse TET1 was not efficient compared to human TET1. We have no explanation to the lack of efficiency of mouse TET1 protein in triggering targeted demethylation, but this finding seems to be corroborated by the literature. Indeed the use of human TET1-CD is widespread in literature and its efficiency has been demonstrated even in plants (Gallego-Bartolomé et al., 2018). In contrast TET proteins from mice have been used only in some articles. mTET1 has been proved to partially demethylate its targets resulting in a weak reactivation of the expression (Okada et al., 2017). The Rots team also showed a limited activity of mTET1, although mTET2 could be more efficient compared to mTET1 (Chen et al., 2014; Huisman et al., 2016).

Specificity of the TALE-hTET1 tool

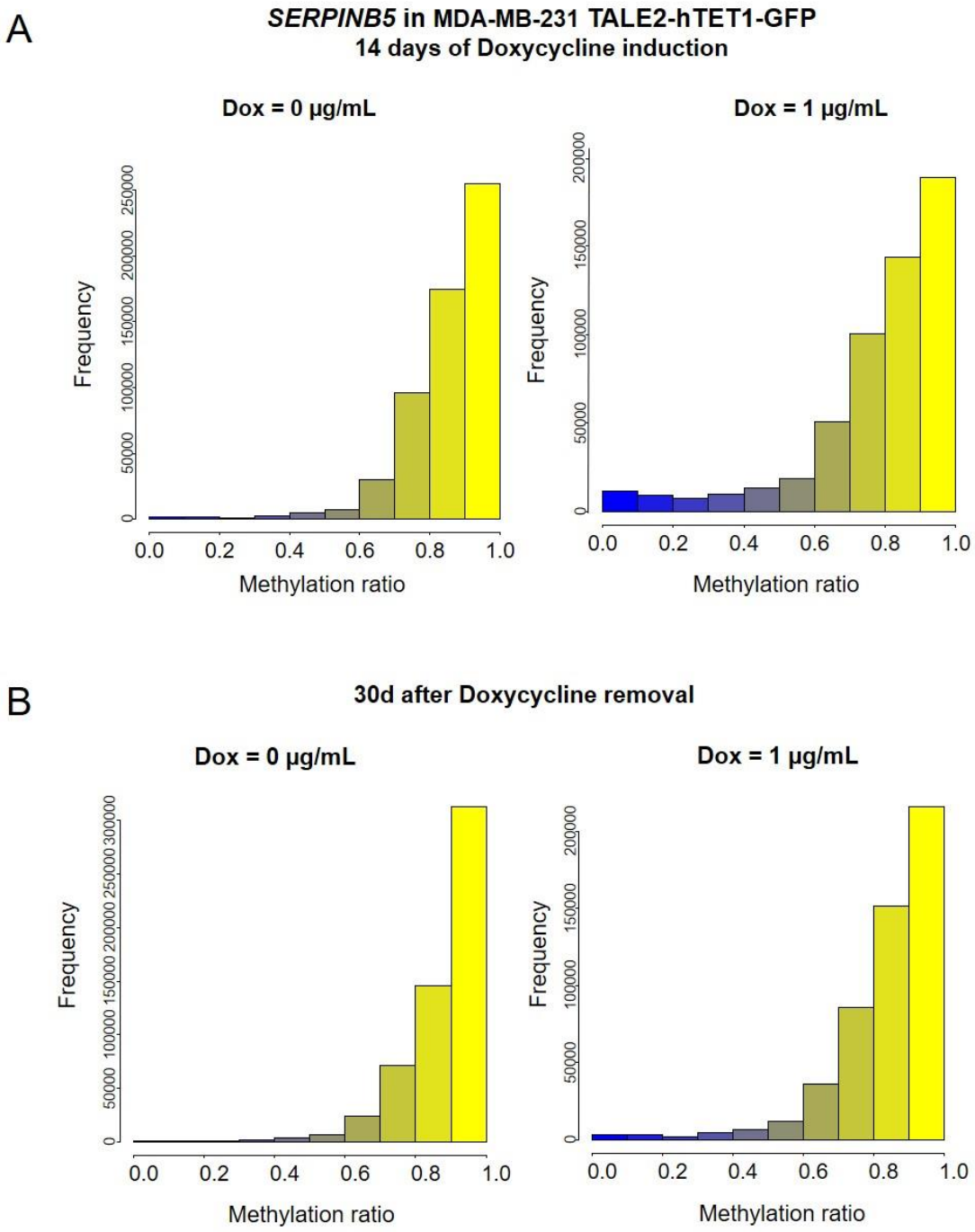
RRBS did not reveal a global demethylation effect after transfection with the TALE-hTET1 fusion protein, suggesting a highly specific effect on the *SERPIN5* promoter methylation. Besides *SERPINB5*, some other DMRs are found, but we cannot conclude if these are off-target effects or secondary effects due to *SERPINB5* reactivation. It would be interesting to analyse if the DMRs contain sequences that closely match the TALE binding sites with some mismatches. We did not analyse the genome wide specificity of dCas9-hTET1 but, according to published data, the dCas9 system promotes major off-target effects when fused to DNMT3A (Galonska et al., 2018; Lin et al., 2018; Pflueger et al., 2018). To overcome this problem, we are currently working in the optimisation of the dCas9-SunTag strategy for targeted TET1 demethylation. In the case of dCas9-DNMT3A, this system strongly improves the specificity of targeted methylation compared to direct dCas9-DNMT3A fusions (Lin et al., 2018; Pflueger et al., 2018).

Durability of the reprogramming

In order to study the stability of the reprogramming, we created a stable MDA-MB-231 cell line expressing the transgene TALE2-hTET1 under the control of a doxycycline inducible promoter. The induction system worked well because the level of the transgene expression measured by EGFP

signal positively correlated with the expression level of *SERPINB5* and the demethylation effect. Moreover, we found that the reactivation and demethylation of *SERPINB5* was durable and did not require the continued expression of the effector domain. Indeed, we were able to detect *SERPINB5* expression even 30 days after doxycycline removal from the media. The expression of *SERPINB5* in the absence of the effector domain seems to be maintained at equivalent levels at the different points of the kinetics. Even though the methylation does not reach original levels even 30 days after dox removal, it seems to return to the promoter in a slow manner. We can think of two hypotheses to explain why DNA methylation comes back to *SERPINB5* promoter. First, it is possible that the chromatin landscape has not been entirely remodelled into accessible chromatin. In this context, *de novo* methylation could be deposited in the promoter through time. Experiments analysing the remodelling of histones marks after the induced demethylation would be informative. Moreover, experiments knocking out or down the DNMTs would help us to validate this hypothesis. Second, we cannot exclude that cells re-expressing *SERPINB5* undergo a negative selection. Previous studies showed that high expression of *SERPINB5* inhibits cell proliferation in MDA-MB-231 cells (Beltran et al., 2007; Garcia-Bloj et al., 2016). Apoptosis and proliferation assays did not show differences between treated and untreated cells at the population level in our hands. Nevertheless, we cannot exclude that some cells that are the most demethylated and express the highest levels of *SERPINB5* grow slower. In fact, when we analyse the methylation on *SERPINB5* promoter at single cell level 14 days after induction, we observe a strong heterogeneity, with some cells presenting a strongly demethylated promoter. This proportion decreases at 30 days after doxycycline removal. **Discussion Figure 1** shows the distribution of single allele methylation in the amplicons of *SERPINB5* promoter at these two points of the kinetics, 0 meaning that the allele is fully unmethylated and 1 meaning that it is fully methylated. In the non-treated cells, most of the values are close to 1, which means that the majority of alleles are fully methylated. However, cells treated with doxycycline and expressing the TALE2-TET1 construct present a different pattern. At 14 days of induction, we observe an increased proportion of alleles with a methylation close to 0, meaning that when demethylation takes place it occurs on the majority of the CpGs on the promoter (**Discussion Figure 1A**). Thus, there is a small proportion of cells showing almost complete loss of methylation in the *SERPINB5* promoter. This could lead to strong *SERPINB5* expression and inhibition of cell growth in only a minority of cells that remain undetectable in our phenotypical studies. This cell to cell heterogeneity might be explained by the fact that the inducible cell line generated by viral transduction was kept as a pool and we did not select clones.

Until now, only a few of the studies working on epigenome editing have evaluated the stability of the effect in time, and the results are conflicting. Concerning the epigenetic silencing of genes with



Discussion Figure 1. Single-allele methylation analysis in the *SERPINB5* promoter.

The graphs show the distribution of methylation scores in single alleles in MDA-MB-231 cells 14 days after doxycycline induction (**A**) and 30 days after doxycycline removal (**B**). The X-axis represents the methylation level (from 0 to 1), 0 meaning that all the CpGs of the allele are demethylated (blue), and 1 meaning that all the CpGs on the allele are methylated (yellow).

DBD coupled to DNMT3A, transitory transfections did not lead to a permanent silencing of the targeted genes (Kungulovski et al., 2015), whereas the use of stable doxycycline inducible cell lines showed that targeted DNA methylation is stable once the DBD-DNMT3A construct is not expressed anymore (McDonald et al., 2016; Rivenbark et al., 2012; Stolzenburg et al., 2015). One critic of these approaches is that minimal leaks in the doxycycline inducible system could lead to continued expression of the transgene at undetectable levels sufficient to maintain the effect. In another study, the transitory transfection of DNMT3A, DNMT3L and KRAB lead to permanent silencing of the *B2M* gene for up to 50 days. This combination had an impact not only on *de novo* methylation but also on histone marks, suggesting that chromatin environment should be modulated to achieve a durable repression (Amabile et al., 2016). Strikingly, the VEGF-A gene was stably repressed in K-562 cells by using the triple combination (Amabile et al., 2016) but it was not in SKOV3 cells when treated only with ZF-DNMT3A (Kungulovski et al., 2015). This suggests that cell-dependent context could have an impact on the durability of epigenetic silencing. Up to now there are no published studies using a stable doxycycline inducible system to evaluate the durability of targeted TET1-dependent demethylation of a gene. In the published literature, one report showed that TALE-hTET1 induced demethylation and re-expression of the *HBB* gene in K562 cells is lost upon cell culture (Maeder et al., 2013b). This could be explained by the fact that they performed transitory transfections, as it was observed in the DNMT3A epigenetic editing studies. A different approach by Liu *et al* showed that dCas9-hTET1-mediated demethylation of the *FMR1* gene in Fragile X iPSCs was stable for at least 2 weeks after inhibition of dCas9 via AcrIIA4 expression (Liu et al., 2018). Interestingly, demethylation was sufficient to induce an active chromatin state of the *FMR1* promoter characterised by increased PolII, H3K4me3 and H3K27ac and decreased H3K9me3. Cano-Rodríguez et al demonstrated that the induction of both H3K4me3 (via PRDM9) and H3K79me3 (via DOT1L) and the absence of methylation were critical to maintain a durable re-expression of epigenetically silenced genes (Cano-Rodriguez et al., 2016). Both articles suggest again that a remodelling of chromatin marks in addition to DNA methylation editing is necessary to achieve stable reprogramming of genes.

Perspectives to improve the demethylation effect on targeted regions

We can propose several strategies to improve the efficiency of the targeted demethylation.

In our experiments with the inducible MDA-MB-231 cells, we worked with the global population and did not select individual clones. We could use flow cytometry to isolate the cells with the highest GFP signal and create individual clones to test if they show a higher demethylation efficiency. Nevertheless, this approach requires to be careful because side effects could result from the random insertion site and so we should use several clones to validate the results.

Several authors report the synergistic effect of the transfection of several TALE-effector domains in combination (Hu et al., 2014; Maeder et al., 2013c, 2013b; Perez-Pinera et al., 2013a). CRISPR-dCas9 is an easier tool for multiplexing. Some reports show a better reprogramming when using a pool of gRNAs targeted to a specific locus (Maeder et al., 2013a; Perez-Pinera et al., 2013b) while others do not (Xu et al., 2016). Co-transfection of different targeting platforms combined to multiple effector domains maximised the activation in highly methylated promoters (Garcia-Bloj et al., 2016). Following the idea that multiple effectors act synergistically in one locus, the dCas9-SunTag system has been shown to highly enrich the demethylation rate when fused to hTET1 (Morita et al., 2016). SunTag system or sgRNA aptamers are interesting alternatives to avoid steric hindrance issued from the targeting of multiple dCas9 to the same region.

An interesting experiment to optimise targeted DNA demethylation would be an accurate comparison of different effector domains (such as TET1, TET2 and TET3) targeted to the same gene in the same cellular context. To my knowledge, there is only one article comparing the efficiency of the three TET proteins (from mouse origin) for targeted demethylation (Chen et al., 2014). In their study system (ovary cancer cell line), TET3 had no effect, but we cannot exclude a tissue-dependent effect. Thus, reproducing similar comparative studies in different cell types will help to determine the best tool depending on the context. Demethylases from other organisms such as plants should also be tested. As human TET1 has been shown to work in plants (Gallego-Bartolomé et al., 2018), it would be interesting to study species crosstalk.

Concerning the demethylation efficiency, it has been recently shown that vitamin C (ascorbic acid) can act as a cofactor of TET activity, leading to enhanced active demethylation (Blaschke et al., 2013; Shenoy et al., 2017; Yin et al., 2013). Specifically, Shenoy et al showed that treatment of lymphoma cells with vitamin C reactivated tumor suppressor genes and increased the sensitivity to chemotherapy drugs. All these data suggest that, even if the effect is global, its use could be interesting in certain experiments to strengthen the effect of the demethylation.

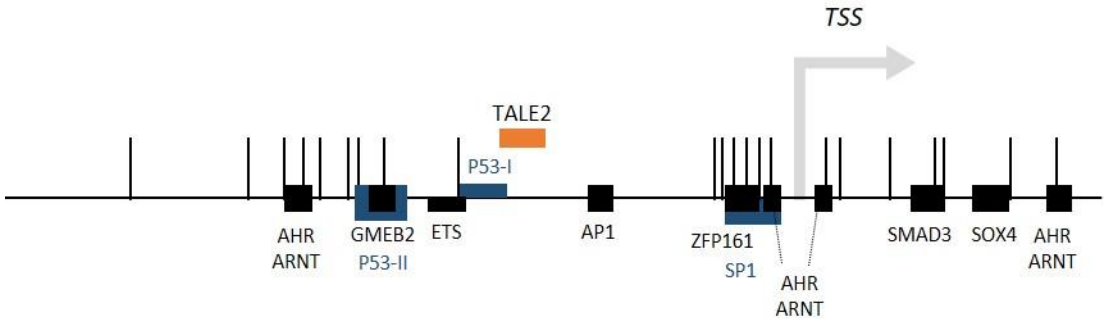
Epigenetic regulation of SERPINB5 expression

The demethylation achieved on *SERPINB5* promoter is stronger at the boundaries of the TALE binding site (CpGs 9 and 10) and spreads in a region of approximately 500 bp, being more robust in the CpGs downstream from the TALE binding site (CpGs 10-22). We hypothesise that the hydroxymethylation and demethylation of the *SERPINB5* promoter make key CpGs involved in the epigenetic control of gene expression accessible. That might allow to the binding of endogenous methylation-sensitive transcription factors. *SERPINB5* expression is known to be controlled by p53 (Oshiro et al., 2003; Zou et al., 1994) or other factors such as TGFβ and Smad that enhance a p53-

dependent expression (Wang et al., 2007). Since MDA-MB-231 have not a functional p53 (Muller and Vousden, 2014; Oshiro et al., 2003), other factors could be involved in *SERPINB5* re-expression. When p53 is absent in cancer cells, p63 can substitute for p53 in transcriptional activation of *SERPINB5* (Kim et al., 2004; Spiesbach et al., 2005). After verification of the position of the p53 binding site in the *SERPINB5* promoter, we confirmed that it is targeted by demethylation in our cells, therefore TAP63, which recognises the same consensus site as p53, could access to the binding site (**Discussion Figure 2**). Ets and AP1 have also been reported to induce *SERPINB5* expression (Zhang et al., 1997). While Ets seems to be a methylation-sensitive TF (Gaston and Fried, 1995), AP-1 sensitivity is controversial (Fujimoto et al., 2005; Kong et al., 2012a). A ZFP161 binding site is located in the *SERPINB5* promoter just downstream of the TALE binding sequence. This ZFP acts as both an activator and a repressor that binds to CpG-rich sequences (Numoto et al., 1999; Obata et al., 1999). ZFP161 has been proposed to be methylation-sensitive (Fang et al., 2006), so it could be a good candidate for mediating methylation-dependent expression of *SERPINB5*. Further experiments knocking out or down the candidates are required to confirm these hypotheses. This way, optimisation of epigenome editing strategies could provide experimental tools to test and prove the methylation-sensitivity of transcription factors.

Function of SERPINB5 in cancer cells

SERPINB5 has been reported to have antitumoral properties in breast cancer cells *in vitro* and *in vivo* by increasing apoptosis and diminishing the invasive potential (Latha et al., 2005; Lee et al., 2012b; Sheng et al., 1996; Shi et al., 2007; Zou et al., 1994). In contrast to what other groups observed (Beltran et al., 2007; Garcia-Bloj et al., 2016), re-expression of *SERPINB5* did not have any effect neither in proliferation nor apoptosis in our MDA-MB-231 cells. Both previous articles (Beltran et al., 2007; Garcia-Bloj et al., 2016) used targeted artificial transcription factors and correlated a higher upregulation level of *SERPINB5* with a higher phenotypical impact. We can hypothesise that the re-expression level achieved with our inducible system is maybe insufficient to induce apoptosis or have a strong impact on cell proliferation. The tumor suppressor potential of *SERPINB5* has recently been called into question (Teoh et al., 2014). In this work, the authors were not able to reproduce previously published experiments and did not observe any phenotypical changes when ectopically re-expressing *SERPINB5* in MDA-MB-231 cells. They argued that they had not performed a clonal selection in contrast of the majority of published reports. Clonal selection could lead to the observation of clonal-derived artefacts instead of consequences derived from the *SERPINB5* activation. In our work, we also did not select clonal populations in order to avoid secondary effects caused by the TALE2-hTET1 incorporation sites.



Discussion Figure 2. Transcription factor binding sites in the *SERPINB5* promoter.

Perspectives

When planning epigenome editing experiments, three parameters should be taken into account: the DNA binding domain module, the effector domain and the objective of the experiment. ZFP, TALEs and CRISPR have all been shown to be efficient, but their benefits and disadvantages should be taken into consideration in the context of each experiment. For example, CRISPR-dCas9 has been shown to not efficiently target hypermethylated regions whereas ZF overcome this problem. However, since lots of efforts are being made to improve the CRISPR-dCas9 system and because it is much easier to use, it is almost certain that it will be the DBD of choice in future experiments. Its flexibility and multiplexing ability make it also a better platform for performing multi-target or screening experiments; it is straight forward to design libraries of gRNAs targeting all genes or regulatory regions such as distal enhancers. Another interesting way to simultaneously target multiple regions is by fusing the ED to a protein domain known to recognise specific regions in the genome. For example, CXXC and MBD domains can be useful to target unmethylated or methylated CpG regions respectively. Indeed, one group has fused TET1-CD to the MBD domain and shown that this can lead to simultaneous demethylation and activation of multiple repressed TSGs in cancer cells (Mizuguchi et al., 2016). Another interesting development of epigenetic editing tools would be to fuse ED to transcription factors. This way, we could expect to modulate specific sets of genes.

The choice of the best targeting strategy and the effector domain is dependent on the choice of the target locus. Results published in literature do not provide a conclusive answer as to what is the best approach for epigenetic editing. As every article uses different target locus, in different cell models and with different transfection strategies, it is very complicated to establish accurate comparisons. Clearly, the effects seem to depend on the chromatin landscape that makes some genes more susceptible to changes than others. For these reasons, a systematic comparison of all available strategies at multiple genes, representing different chromatin contexts in the same cellular system, would be highly informative. This way it would be possible to evaluate and compare the real advantages and drawbacks of each targeting method without the confounding effect of using different cell lines.

The application of epigenome editing strategies in therapeutics is appealing but further work is needed to overcome major imitations such as off-target effects, delivery and durability of the effects. In cancer, epigenome editing could be useful to sensitise cells to other therapies. Combination of epi-drugs inhibiting DNA methylation or histone acetylation with chemo or immunotherapies improves the efficiency of the treatments. Indeed, epi-drugs administration (e.g. 5aza) prior to conventional

therapy is supposed to sensitise the cells to the therapy. The main problem is that these drugs induce side effects because of their genome-wide effects. Studying the methylation status of key genes in patient tumors could be an asset when applying personalised medicine and to predict drug response. This way, we could elaborate personalised strategies combining targeted epigenome editing and conventional therapies. For example, targeted demethylation and re-expression of the oestrogen receptor would sensitise breast cancer cells that are resistant to hormone therapy. As we saw in the introduction chapter, cancer-testis antigens are good candidates to be used as targets for immunotherapy because they are silenced in somatic cells but often re-expressed in cancer cells. We can think about strategies aiming to epigenetically activate combinations of CTAs or other cell surface receptors in a determined tumor in order to target it with antibodies. Another strategy could be to use epigenome editing to induce the re-expression of combinations of tumor suppressor genes expecting to induce cell death in a particular tumor, weakening the tumor mass and making it more susceptible to conventional therapies. When using epigenetic editing with the objective to sensitise cells to other therapies, the effect of reprogramming should be stable the time of the treatment, therefore studying the durability of epigenetic editing is of particular importance.

The use of epigenetic editing in the treatment of other pathologies is more challenging. For example, if we want to use epigenome editing to correct methylation-dependent imprinting disorders, there are two main limitations. First, as of today there has been no demonstration that epigenetic editing can induce a life-long reprogramming of epigenetic marks. Second, epigenetic editing with DNA binding modules do not differentiate between alleles. One solution could be to choose targeted sequences containing specific polymorphisms in patients. Moreover, it should be studied if epigenetic correction could be performed directly in the germline (*in vitro*). In this line, other germline epimutations linked to predisposition to familial cancer (e.g. *BRCA1*, *hMLH1*, *hMSH2*) or other diseases could be corrected by direct epigenetic editing in the oocyte, sperm or early embryos *in vitro*. CRISPR-Cas9-based genetic corrections have been successfully achieved in mouse spermatogonia or embryos (Huai et al., 2017; Wu et al., 2015) and even in human embryos (Ma et al., 2017), paving the way to epigenome editing studies.

Delivery constitutes another main challenge in epigenome editing, particularly delivery to solid tumors. First, it is necessary to choose in which form to deliver the effectors (plasmid DNA, RNA or protein). Second, another challenge is to target the delivery to the tumor cells. This could be done by detecting tumor-specific cell surface receptors and using the appropriate antibodies. For example, epigenome editing effectors could be delivered by nanoparticles coated with antibodies targeting a specific tumor. In addition, *ex vivo* strategies are interesting for the treatment of some cancers, in particular leukaemias. Currently, there are clinical trials using Cas9 genome editing. Genetic

corrections are performed *ex vivo* and the cells are reinfused in the patients. One can envision that durable epigenetic reprogramming and reinfusion of cells in patients could be tested for future therapies. Finally, Adeno-virus-delivered gene therapy has been used in clinical trials for a long time, but the first FDA approved gene-therapy with AAV dates only from 2017. Using viral vectors for clinical applications requires careful study of their safety in patients.

An important point concerning epigenome editing are the ethical implications from its utilisation in humans. Where is the balance of risks and benefits? Gene therapy treatment on the fatal disease X-linked severe combined immunodeficiency disorder (X-SCID) (also known as the "bubble boys" disease) was shown to be extremely efficient. However, several patients developed leukaemia after the treatment, raising a debate on the safety of these treatments. For these reasons, the major challenges that need to be faced before clinical applications are the development of fine-tuned vectors and strict evaluation of safety and potential side effects in clinical trials. Moreover, strict laws should be developed in order to promote a conscious and non-dangerous use of genome and epigenome editing tools.

Conclusion

We have presented the proof-of-concept that a repressed tumor suppressor gene can be specifically and durably reprogrammed via DNA demethylation in cancer cells. Targeted demethylation could have potential use in therapeutic applications and represents a good alternative to prevent global effects of current epigenetic drugs such as 5aza. In addition, targeted demethylation could be used to reprogram gene expression in other methylation-associated diseases like imprinting-related syndromes. Future work should address if combination of different epigenome editing strategies or effector domains can lead to an increased efficiency and stability of the effect with minimal off-targets. Reaching efficient targeting with minimal off-target effects and appropriate delivery systems *in vivo* are now the ultimate challenges to translate epigenome editing in clinics. In conclusion, my work contributed to show that targeted epigenome editing efficiently works in mammalian cells even if it is still at the beginning of its development.

REFERENCES

REFERENCES

Aapola, U., Kawasaki, K., Scott, H.S., Ollila, J., Vihinen, M., Heino, M., Shintani, A., Kawasaki, K., Minoshima, S., Krohn, K., et al. (2000). Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family. Genomics *65*, 293–298.

Abdouni, H., King, J.J., Suliman, M., Quinlan, M., Fifield, H., and Larijani, M. (2013). Zebrafish AID is capable of deaminating methylated deoxycytidines. Nucleic Acids Res. *41*, 5457–5468.

Abi Habib, W., Azzi, S., Brioude, F., Steunou, V., Thibaud, N., Neves, C.D., Le Jule, M., Chantot-Bastaraud, S., Keren, B., Lyonnet, S., et al. (2014). Extensive investigation of the IGF2/H19 imprinting control region reveals novel OCT4/SOX2 binding site defects associated with specific methylation patterns in Beckwith-Wiedemann syndrome. Hum. Mol. Genet. *23*, 5763–5773.

Abraham, S., Zhang, W., Greenberg, N., and Zhang, M. (2003). Maspin Functions as Tumor Suppressor by Increasing Cell Adhesion to Extracellular Matrix in Prostate Tumor Cells. J. Urol. *169*, 1157–1161.

Abu-Amero, S., Wakeling, E.L., Preece, M., Whittaker, J., Stanier, P., and Moore, G.E. (2010). Epigenetic signatures of Silver-Russell syndrome. J. Med. Genet. *47*, 150–154.

Abudayyeh, O.O., Gootenberg, J.S., Essletzbichler, P., Han, S., Joung, J., Belanto, J.J., Verdine, V., Cox, D.B.T., Kellner, M.J., Regev, A., et al. (2017). RNA targeting with CRISPR–Cas13. Nature *550*, 280–284.

Agodi, A., Barchitta, M., Quattrocchi, A., Maugeri, A., and Vinciguerra, M. (2015). DAPK1 Promoter Methylation and Cervical Cancer Risk: A Systematic Review and a Meta-Analysis. PLOS ONE *10*, e0135078.

Agoston, A.T., Argani, P., Yegnasubramanian, S., De Marzo, A.M., Ansari-Lari, M.A., Hicks, J.L., Davidson, N.E., and Nelson, W.G. (2005). Increased Protein Stability Causes DNA Methyltransferase 1 Dysregulation in Breast Cancer. J. Biol. Chem. *280*, 18302–18310.

Akiyama, Y., Watkins, N., Suzuki, H., Jair, K.-W., van Engeland, M., Esteller, M., Sakai, H., Ren, C.-Y., Yuasa, Y., Herman, J.G., et al. (2003). GATA-4 and GATA-5 Transcription Factor Genes and Potential Downstream Antitumor Target Genes Are Epigenetically Silenced in Colorectal and Gastric Cancer. Mol. Cell. Biol. *23*, 8429–8439.

Alewine, C., Hassan, R., and Pastan, I. (2015). Advances in Anticancer Immunotoxin Therapy. The Oncologist *20*, 176–185.

Al-Kharashi, L.A., Al-Mohanna, F.H., Tulbah, A., and Aboussekhra, A. (2017). The DNA methyl-transferase protein DNMT1 enhances tumor-promoting properties of breast stromal fibroblasts. Oncotarget *9*, 2329–2343.

Allahyari, H., Heidari, S., Ghamgosha, M., Saffarian, P., and Amani, J. (2017). Immunotoxin: A new tool for cancer therapy. Tumor Biol. *39*, 1010428317692226.

Amabile, A., Migliara, A., Capasso, P., Biffi, M., Cittaro, D., Naldini, L., and Lombardo, A. (2016). Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing. Cell *167*, 219-232.e14.

Amary, M.F., Bacsi, K., Maggiani, F., Damato, S., Halai, D., Berisha, F., Pollock, R., O'Donnell, P., Grigoriadis, A., Diss, T., et al. (2011). IDH1 and IDH2 mutations are frequent events in central chondrosarcoma and central and periosteal chondromas but not in other mesenchymal tumours. J. Pathol. *224*, 334–343.

Amato, R.J., Stephenson, J., Hotte, S., Nemunaitis, J., Bélanger, K., Reid, G., and Martell, R.E. (2012). MG98, a Second-Generation DNMT1 Inhibitor, in the Treatment of Advanced Renal Cell Carcinoma. Cancer Invest. *30*, 415–421.

Amir, R.E., Van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U., and Zoghbi, H.Y. (1999). Rett syndrome is caused by mutations in X-linked *MECP2*, encoding methyl-CpG-binding protein 2. Nat. Genet. *23*, 185–188.

Ao, X., Sa, R., Wang, J., Dao, R., Wang, H., and Yu, H. (2016). Activation-induced cytidine deaminase selectively catalyzed active DNA demethylation in pluripotency gene and improved cell reprogramming in bovine SCNT embryo. Cytotechnology *68*, 2637–2648.

Aporntewan, C., Phokaew, C., Piriyapongsa, J., Ngamphiw, C., Ittiwut, C., Tongsima, S., and Mutirangura, A. (2011). Hypomethylation of Intragenic LINE-1 Represses Transcription in Cancer Cells through AGO2. PLoS ONE *6*, e17934.

Arand, J., Spieler, D., Karius, T., Branco, M.R., Meilinger, D., Meissner, A., Jenuwein, T., Xu, G., Leonhardt, H., Wolf, V., et al. (2012). In Vivo Control of CpG and Non-CpG DNA Methylation by DNA Methyltransferases. PLOS Genet. *8*, e1002750.

Aravin, A.A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K.F., Bestor, T., and Hannon, G.J. (2008). A piRNA Pathway Primed by Individual Transposons Is Linked to De Novo DNA Methylation in Mice. Mol. Cell *31*, 785–799.

Arnaud, P. (2010). Genomic imprinting in germ cells: imprints are under control. Reproduction *140*, 411–423.

Asmar, F., Punj, V., Christensen, J., Pedersen, M.T., Pedersen, A., Nielsen, A.B., Hother, C., Ralfkiaer, U., Brown, P., Ralfkiaer, E., et al. (2013). Genome-wide profiling identifies a DNA methylation signature that associates with TET2 mutations in diffuse large B-cell lymphoma. Haematologica *98*, 1912–1920.

Auclair, G., Guibert, S., Bender, A., and Weber, M. (2014). Ontogeny of CpG island methylation and specificity of DNMT3 methyltransferases during embryonic development in the mouse. Genome Biol. *15*.

Baars, R. van, Marel, J. van der, Snijders, P.J.F., Rodriquez-Manfredi, A., Harmsel, B. ter, Munckhof, H.A.M. van den, Ordi, J., Pino, M. del, Sandt, M.M. van de, Wentzensen, N., et al. (2016). CADM1 and MAL methylation status in cervical scrapes is representative of the most severe underlying lesion in women with multiple cervical biopsies. Int. J. Cancer *138*, 463–471.

Bachman, M., Uribe-Lewis, S., Yang, X., Williams, M., Murrell, A., and Balasubramanian, S. (2014). 5-Hydroxymethylcytosine is a predominantly stable DNA modification. Nat. Chem. *6*, 1049–1055.

Bai, J., Choi, S.-H., Ponciano, G., Leung, H., and Leach, J.E. (2000). *Xanthomonas oryzae* pv. *Oryzae* Avirulence Genes Contribute Differently and Specifically to Pathogen Aggressiveness. Mol. Plant. Microbe Interact. *13*, 1322–1329.

Bakulski, K.M., Dolinoy, D.C., Sartor, M.A., Paulson, H.L., Konen, J.R., Lieberman, A.P., Albin, R.L., Hu, H., and Rozek, L.S. (2012). Genome-Wide DNA Methylation Differences Between Late-Onset Alzheimer's Disease and Cognitively Normal Controls in Human Frontal Cortex. J. Alzheimers Dis. *29*, 571–588.

Balss, J., Meyer, J., Mueller, W., Korshunov, A., Hartmann, C., and von Deimling, A. (2008). Analysis of the IDH1 codon 132 mutation in brain tumors. Acta Neuropathol. (Berl.) *116*, 597–602.

Barakat, T.S., and Gribnau, J. (2012). X chromosome inactivation in the cycle of life. Development *139*, 2085–2089.

Barau, J., Teissandier, A., Zamudio, N., Roy, S., Nalesso, V., Hérault, Y., Guillou, F., and Bourc'his, D. (2016). The DNA methyltransferase DNMT3C protects male germ cells from transposon activity. Science *354*, 909–912.

Barbarotta, L., and Hurley, K. (2015). Romidepsin for the Treatment of Peripheral T-Cell Lymphoma. J. Adv. Pract. Oncol. *6*, 22–36.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. Science *315*, 1709–1712.

Bashtrykov, P., Jankevicius, G., Jurkowska, R.Z., Ragozin, S., and Jeltsch, A. (2014). The UHRF1 Protein Stimulates the Activity and Specificity of the Maintenance DNA Methyltransferase DNMT1 by an Allosteric Mechanism. J. Biol. Chem. *289*, 4106–4115.

Baubec, T., Colombo, D.F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A.R., Akalin, A., and Schübeler, D. (2015). Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. Nature *520*, 243–247.

Baylin, S.B., and Jones, P.A. (2011). A decade of exploring the cancer epigenome — biological and translational implications. Nat. Rev. Cancer *11*, 726–734.

Beerli, R.R., Segal, D.J., Dreier, B., and Barbas, C.F. (1998). Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. Proc. Natl. Acad. Sci. U. S. A. *95*, 14628–14633.

Bell, A.C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature *405*, 482–485.

Beltran, A., Parikh, S., Liu, Y., Cuevas, B.D., Johnson, G.L., Futscher, B.W., and Blancafort, P. (2007). Re-activation of a dormant tumor suppressor gene maspin by designed transcription factors. Oncogene *26*, 2791–2798.

Berardi, R., Morgese, F., Onofri, A., Mazzanti, P., Pistelli, M., Ballatore, Z., Savini, A., De Lisa, M., Caramanti, M., Rinaldi, S., et al. (2013). Role of maspin in cancer. Clin Transl Med *2*, 8.

Berman, B.P., Weisenberger, D.J., Aman, J.F., Hinoue, T., Ramjan, Z., Liu, Y., Noushmehr, H., Lange, C.P.E., van Dijk, C.M., Tollenaar, R.A.E.M., et al. (2011). Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. Nat. Genet. *44*, 40–46.

Bernstein, D.L., Le Lay, J.E., Ruano, E.G., and Kaestner, K.H. (2015). TALE-mediated epigenetic suppression of CDKN2A increases replication in human fibroblasts. J. Clin. Invest. *125*, 1998–2006.

Bestor, T.H., and Ingram, V.M. (1983). Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. Proc. Natl. Acad. Sci. U. S. A. *80*, 5559–5563.

Bestor, T., Laudano, A., Mattaliano, R., and Ingram, V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. J. Mol. Biol. *203*, 971–983.

Bhakta, M., and Segal, D.J. (2010). The generation of zinc finger proteins by modular assembly. Methods Mol. Biol. Clifton NJ *649*, 3–30.

Bhogal, B., Arnaudo, A., Dymkowski, A., Best, A., and Davis, T.L. (2004). Methylation at mouse Cdkn1c is acquired during postimplantation development and functions to maintain imprinted expression. Genomics *84*, 961–970.

Bhutani, N., Brady, J.J., Damian, M., Sacco, A., Corbel, S.Y., and Blau, H.M. (2010). Reprogramming towards pluripotency requires AID-dependent DNA demethylation. Nature *463*, 1042–1047.

Blaschke, K., Ebata, K.T., Karimi, M.M., Zepeda-Martínez, J.A., Goyal, P., Mahapatra, S., Tam, A., Laird, D.J., Hirst, M., Rao, A., et al. (2013). Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. Nature *500*, 222–226.

Boch, J., and Bonas, U. (2010). Xanthomonas AvrBs3 Family-Type III Effectors: Discovery and Function. Annu. Rev. Phytopathol. *48*, 419–436.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009). Breaking the code of DNA binding specificity of TAL-type III effectors. Science *326*, 1509–1512.

Bonasio, R., Li, Q., Lian, J., Mutti, N.S., Jin, L., Zhao, H., Zhang, P., Wen, P., Xiang, H., Ding, Y., et al. (2012). Genome-wide and Caste-Specific DNA Methylomes of the Ants Camponotus floridanus and Harpegnathos saltator. Curr. Biol. *22*, 1755–1764.

Borgel, J., Guibert, S., Li, Y., Chiba, H., Schübeler, D., Sasaki, H., Forné, T., and Weber, M. (2010). Targets and dynamics of promoter DNA methylation during early mouse development. Nat. Genet. *42*, 1093–1100.

Bostick, M., Kim, J.K., Estève, P.-O., Clark, A., Pradhan, S., and Jacobsen, S.E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science *317*, 1760–1764.

Boulard, M., Edwards, J.R., and Bestor, T.H. (2015). FBXL10 protects Polycomb-bound genes from hypermethylation. Nat. Genet. *47*, 479–485.

Bourc'his, D. (2001). Dnmt3L and the Establishment of Maternal Genomic Imprints. Science *294*, 2536–2539.

Bourc'his, D., and Bestor, T.H. (2004). Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature *431*, 96–99.

Brait, M., Banerjee, M., Maldonado, L., Ooki, A., Loyo, M., Guida, E., Izumchenko, E., Mangold, L., Humphreys, E., Rosenbaum, E., et al. (2017). Promoter methylation of MCAM, ERα and ERβ in serum of early stage prostate cancer patients. Oncotarget *8*, 15431–15440.

Brandeis, M., Frank, D., Keshet, I., Siegfried, Z., Mendelsohn, M., Nemes, A., Temper, V., Razin, A., and Cedar, H. (1994). Sp1 elements protect a CpG island from de novo methylation. Nature *371*, 435–438.

Bransteitter, R., Pham, P., Scharff, M.D., and Goodman, M.F. (2003). Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. Proc. Natl. Acad. Sci. U. S. A. *100*, 4102–4107.

Brock, M.V., Hooker, C.M., Ota-Machida, E., Han, Y., Guo, M., Ames, S., Glöckner, S., Piantadosi, S., Gabrielson, E., Pridham, G., et al. (2008). DNA methylation markers and early recurrence in stage I lung cancer. N. Engl. J. Med. *358*, 1118–1128.

Brockdorff, N., McCabe, V.M., Norris, D.P., Cooper, P.J., Swift, S., Rastan', S., and Kay, F. (1992). The Product of the Mouse Xist Gene Is a 15 kb Inactive X-Specific Transcript Containing No Consewed ORF and Located in the Nucleus. Cell *71*, 515–526.

Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, R., and Willard, H.F. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature *349*, 38–44.

Brown, C.J., Hendrich, B.D., Rupert, J.L., Xing, Y., Lawrence, J., and Willard't, H.F. (1992). The Human X/ST Gene: Analysis of a 17 kb Inactive X-Specific RNA That Contains Conserved Repeats and Is Highly Localized within the Nucleus. *71*, 527–542.

Brueckner, B., Garcia Boy, R., Siedlecki, P., Musch, T., Kliem, H.C., Zielenkiewicz, P., Suhai, S., Wiessler, M., and Lyko, F. (2005). Epigenetic Reactivation of Tumor Suppressor Genes by a Novel Small-Molecule Inhibitor of Human DNA Methyltransferases. Cancer Res. *65*, 6305–6311.

Brueckner, B., Rius, M., Markelova, M.R., Fichtner, I., Hals, P.-A., Sandvold, M.L., and Lyko, F. (2010). Delivery of 5-Azacytidine to Human Cancer Cells by Elaidic Acid Esterification Increases Therapeutic Drug Efficacy. Mol. Cancer Ther. *9*, 1256–1264.

Buiting, K., Groß, S., Lich, C., Gillessen-Kaesbach, G., El-Maarri, O., and Horsthemke, B. (2003). Epimutations in Prader-Willi and Angelman Syndromes: A Molecular Study of 136 Patients with an Imprinting Defect. Am. J. Hum. Genet. *72*, 571–577.

Bultmann, S., Morbitzer, R., Schmidt, C.S., Thanisch, K., Spada, F., Elsaesser, J., Lahaye, T., and Leonhardt, H. (2012). Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. Nucleic Acids Res. *40*, 5368–5377.

Busque, L., Patel, J.P., Figueroa, M., Vasanthakumar, A., Provost, S., Hamilou, Z., Mollica, L., Li, J., Viale, A., Heguy, A., et al. (2012). Recurrent Somatic TET2 Mutations in Normal Elderly Individuals With Clonal Hematopoiesis. Nat. Genet. *44*, 1179–1181.

Bustos, B. de U., Estal, R.M., Simó, G.P., Farinos, J.S., Marco, C.P., Mira, M.N., Miquel, V.A. de, Sánchez, R.B., Marco, V.S., Ros, M.L., et al. (2018). Aberrant DNA methylation is associated with aggressive clinicopathological features and poor survival in cutaneous melanoma. Br. J. Dermatol. *179*, 394–404.

Butcher, D.T., and Rodenhiser, D.I. (2007). Epigenetic inactivation of BRCA1 is associated with aberrant expression of CTCF and DNA methyltransferase (DNMT3B) in some sporadic breast tumours. Eur. J. Cancer *43*, 210–219.

Caceres, I.I. de, Cortes-Sempere, M., Moratilla, C., Machado-Pinilla, R., Rodriguez-Fanjul, V., Manguán-García, C., Cejas, P., López-Ríos, F., Paz-Ares, L., CastroCarpeño, J. de, et al. (2010). *IGFBP-3* hypermethylation-derived deficiency mediates cisplatin resistance in non-small-cell lung cancer. Oncogene *29*, 1681–1690.

Caldeira, J.R.F., Prando, É.C., Quevedo, F.C., Neto, F.A.M., Rainho, C.A., and Rogatto, S.R. (2006). CDH1 promoter hypermethylation and E-cadherin protein expression in infiltrating breast cancer. BMC Cancer *6*, 48.

Candelaria, M., Herrera, A., Labardini, J., González-Fierro, A., Trejo-Becerril, C., Taja-Chayeb, L., Pérez-Cárdenas, E., de la Cruz-Hernández, E., Arias-Bofill, D., Vidal, S., et al. (2011). Hydralazine and magnesium valproate as epigenetic treatment for myelodysplastic syndrome. Preliminary results of a phase-II trial. Ann. Hematol. *90*, 379–387.

Cano-Rodriguez, D., and Rots, M.G. (2016). Epigenetic Editing: On the Verge of Reprogramming Gene Expression at Will. Curr. Genet. Med. Rep. *4*, 170–179.

Cano-Rodriguez, D., Gjaltema, R.A.F., Jilderda, L.J., Jellema, P., Dokter-Fokkens, J., Ruiters, M.H.J., and Rots, M.G. (2016). Writing of H3K4Me3 overcomes epigenetic silencing in a sustained but context-dependent manner. Nat. Commun. *7*, 12284.

Carmona, F.J., Villanueva, A., Vidal, A., Muñoz, C., Puertas, S., Penin, R.M., Gomà, M., Lujambio, A., Piulats, J.M., Mesía, R., et al. (2012). Epigenetic disruption of cadherin-11 in human cancer metastasis. J. Pathol. *228*, 230–240.

Cebrian-Serrano, A., and Davies, B. (2017). CRISPR-Cas orthologues and variants: optimizing the repertoire, specificity and delivery of genome engineering tools. Mamm. Genome *28*, 247–261.

Ceccaldi, A., Rajavelu, A., Champion, C., Rampon, C., Jurkowska, R., Jankevicius, G., Sénamaud-Beaufort, C., Ponger, L., Gagey, N., Dali Ali, H., et al. (2011). C5-DNA Methyltransferase Inhibitors: From Screening to Effects on Zebrafish Embryo Development. ChemBioChem *12*, 1337–1345.

Ceccaldi, A., Rajavelu, A., Ragozin, S., Sénamaud-Beaufort, C., Bashtrykov, P., Testa, N., Dali-Ali, H., Maulay-Bailly, C., Amand, S., Guianvarc'h, D., et al. (2013). Identification of Novel Inhibitors of DNA Methylation by Screening of a Chemical Library. ACS Chem. Biol. *8*, 543–548.

Chadwick, B.P., and Willard, H.F. (2001). Histone H2A variants and the inactive X chromosome: identification of a second macroH2A variant. Hum. Mol. Genet. *10*, 1101–1113.

Champion, C., Guianvarc'h, D., Sénamaud-Beaufort, C., Jurkowska, R.Z., Jeltsch, A., Ponger, L., Arimondo, P.B., and Guieysse-Peugeot, A.-L. (2010). Mechanistic Insights on the Inhibition of C5 DNA Methyltransferases by Zebularine. PLoS ONE *5*, e12388.

Chan, S.M., and Majeti, R. (2013). Role of DNMT3A, TET2, and IDH1/2 Mutations in Pre-Leukemic Stem Cells in Acute Myeloid Leukemia. Int. J. Hematol. *98*, 648–657.

Chan, T.A., Glockner, S., Yi, J.M., Chen, W., Van Neste, L., Cope, L., Herman, J.G., Velculescu, V., Schuebel, K.E., Ahuja, N., et al. (2008). Convergence of Mutation and Epigenetic Alterations Identifies Common Genes in Cancer That Predict for Poor Prognosis. PLoS Med. *5*, e114.

Chang, P.-Y., Liao, Y.-P., Wang, H.-C., Chen, Y.-C., Huang, R.-L., Wang, Y.-C., Yuan, C.-C., and Lai, H.-C. (2017). An epigenetic signature of adhesion molecules predicts poor prognosis of ovarian cancer patients. Oncotarget *8*, 53432–53449.

Chavez, A., Scheiman, J., Vora, S., Pruitt, B.W., Tuttle, M., Iyer, E., Lin, S., Kiani, S., Guzman, C.D., Wiegand, D.J., et al. (2015). Highly-efficient Cas9-mediated transcriptional programming. Nat. Methods *12*, 326–328.

Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G.-W., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S., et al. (2013a). Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System. Cell *155*, 1479–1491.

Chen, H., Kazemier, H.G., de Groote, M.L., Ruiters, M.H.J., Xu, G.-L., and Rots, M.G. (2014). Induced DNA demethylation by targeting Ten-Eleven Translocation 2 to the human ICAM-1 promoter. Nucleic Acids Res. *42*, 1563–1574.

Chen, J., Guo, L., Zhang, L., Wu, H., Yang, J., Liu, H., Wang, X., Hu, X., Gu, T., Zhou, Z., et al. (2013b). Vitamin C modulates TET1 function during somatic cell reprogramming. Nat. Genet. *45*, 1504–1509.

Chen, K., Wang, G., Peng, L., Liu, S., Fu, X., Zhou, Y., Yu, H., Li, A., Li, J., Zhang, S., et al. (2011). CADM1/TSLC1 inactivation by promoter hypermethylation is a frequent event in colorectal carcinogenesis and correlates with late stages of the disease. Int. J. Cancer *128*, 266–273.

Chen, K., Wu, K., Jiao, X., Wang, L., Ju, X., Wang, M., Di Sante, G., Xu, S., Wang, Q., Li, K., et al. (2015a). The Endogenous Cell-Fate Factor Dachshund restrains Prostate Epithelial Cell Migration via Repression of Cytokine Secretion via a CXCL Signaling Module. Cancer Res. *75*, 1992–2004.

Chen, L., Chen, K., Lavery, L.A., Baker, S.A., Shaw, C.A., Li, W., and Zoghbi, H.Y. (2015b). MeCP2 binds to non-CG methylated DNA as neurons mature, influencing transcription and the timing of onset for Rett syndrome. Proc. Natl. Acad. Sci. U. S. A. *112*, 5509–5514.

Chen, S., Sanjana, N.E., Zheng, K., Shalem, O., Lee, K., Shi, X., Scott, D.A., Song, J., Pan, J.Q., Weissleder, R., et al. (2015c). Genome-wide CRISPR Screen in a Mouse Model of Tumor Growth and Metastasis. Cell *160*, 1246–1260.

Chen, T., Ueda, Y., Dodge, J.E., Wang, Z., and Li, E. (2003). Establishment and Maintenance of Genomic Methylation Patterns in Mouse Embryonic Stem Cells by Dnmt3a and Dnmt3b. Mol. Cell. Biol. *23*, 5594–5605.

Chen, Y., Peng, Y., Fan, S., Li, Y., Xiao, Z.-X., and Li, C. (2018). A double dealing tale of p63: an oncogene or a tumor suppressor. Cell. Mol. Life Sci. *75*, 965–973.

Cheng, A.W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T.W., Rangarajan, S., Shivalila, C.S., Dadon, D.B., and Jaenisch, R. (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. Cell Res. *23*, 1163–1171.

Cheng, A.W., Jillette, N., Lee, P., Plaskon, D., Fujiwara, Y., Wang, W., Taghbalout, A., and Wang, H. (2016). Casilio: a versatile CRISPR-Cas9-Pumilio hybrid for gene regulation and genomic labeling. Cell Res. *26*, 254–257.

Cheng, J.C., Weisenberger, D.J., Gonzales, F.A., Liang, G., Xu, G.-L., Hu, Y.-G., Marquez, V.E., and Jones, P.A. (2004). Continuous Zebularine Treatment Effectively Sustains Demethylation in Human Bladder Cancer Cells. Mol. Cell. Biol. *24*, 1270–1278.

Cher, M.L., Biliran, H.R., Bhagat, S., Meng, Y., Che, M., Lockett, J., Abrams, J., Fridman, R., Zachareas, M., and Sheng, S. (2003). Maspin expression inhibits osteolysis, tumor growth, and angiogenesis in a model of prostate cancer bone metastasis. Proc. Natl. Acad. Sci. *100*, 7847–7852.

Cho, H.-S., Kang, J.G., Lee, J.-H., Lee, J.-J., Jeon, S.K., Ko, J.-H., Kim, D.-S., Park, K.-H., Kim, Y.-S., and Kim, N.-S. (2015). Direct regulation of E-cadherin by targeted histone methylation of TALE-SET fusion protein in cancer cells. Oncotarget *6*, 23837–23844.

Cho, S.W., Kim, S., Kim, J.M., and Kim, J.-S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat. Biotechnol. *31*, 230–232.

Chomyk, A.M., Volsko, C., Tripathi, A., Deckard, S.A., Trapp, B.D., Fox, R.J., and Dutta, R. (2017). DNA methylation in demyelinated multiple sclerosis hippocampus. Sci. Rep. *7:8696*.

Choo, Y., Sánchez-García, I., and Klug, A. (1994). In vivo repression by a site-specific DNA-binding protein designed against an oncogenic sequence. Nature *372*, 642–645.

Choudhury, S.R., Cui, Y., Lubecka, K., Stefanska, B., and Irudayaraj, J. (2016). CRISPR-dCas9 mediated TET1 targeting for selective DNA demethylation at BRCA1 promoter. Oncotarget *7*, 46545–46556.

Chouliaras, L., Mastroeni, D., Delvaux, E., Grover, A., Kenis, G., Hof, P.R., Steinbusch, H.W.M., Coleman, P.D., Rutten, B.P.F., and van den Hove, D.L.A. (2013). Consistent decrease in global DNA methylation and hydroxymethylation in the hippocampus of Alzheimer's disease patients. Neurobiol. Aging *34*, 2091–2099.

Chuang, J.C. (2005). Comparison of biological effects of non-nucleoside DNA methylation inhibitors versus 5-aza-2'-deoxycytidine. Mol. Cancer Ther. *4*, 1515–1520.

Cimmino, L., Dawlaty, M.M., Ndiaye-Lobry, D., Yap, Y.S., Bakogianni, S., Yu, Y., Bhattacharyya, S., Shaknovich, R., Geng, H., Lobry, C., et al. (2015). Tet1 is a tumor suppressor of hematopoietic malignancy. Nat. Immunol. *16*, 653–662.

Clouaire, T., Webb, S., Skene, P., Illingworth, R., Kerr, A., Andrews, R., Lee, J.-H., Skalnik, D., and Bird, A. (2012). Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. Genes Dev. *26*, 1714–1728.

Clozel, T., Yang, S., Elstrom, R.L., Tam, W., Martin, P., Kormaksson, M., Banerjee, S., Vasanthakumar, A., Culjkovic, B., Scott, D.W., et al. (2013). Mechanism-Based Epigenetic Chemosensitization Therapy of Diffuse Large B Cell Lymphoma. Cancer Discov. *3*, 1002–1019.

Coiffier, B., Pro, B., Prince, H.M., Foss, F., Sokol, L., Greenwood, M., Caballero, D., Borchmann, P., Morschhauser, F., Wilhelm, M., et al. (2012). Results from a pivotal, open-label, phase II study of romidepsin in relapsed or refractory peripheral T-cell lymphoma after prior systemic therapy. J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol. *30*, 631–636.

Compere, S.J., and Palmiter, R.D. (1981). DNA Methylation Controls the Inducibility of the Mouse Metallothionein-I Gene in Lymphoid Cells. Cell *25*, 233–240.

Cong, L., Zhou, R., Kuo, Y., Cunniff, M., and Zhang, F. (2012). Comprehensive Interrogation of Natural TALE DNA Binding Modules and Transcriptional Repressor Domains. Nat. Commun. *3*, 968.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., et al. (2013). Multiplex Genome Engineering Using CRISPR/Cas Systems. Science *339*, 819–823.

Conticello, S.G., Thomas, C.J.F., Petersen-Mahrt, S.K., and Neuberger, M.S. (2005). Evolution of the AID/APOBEC Family of Polynucleotide (Deoxy)cytidine Deaminases. Mol. Biol. Evol. *22*, 367–377.

Cordaux, R., and Batzer, M.A. (2009). The impact of retrotransposons on human genome evolution. Nat. Rev. Genet. *10*, 691–703.

Coronel, J., Cetina, L., Pacheco, I., Trejo-Becerril, C., González-Fierro, A., de la Cruz-Hernandez, E., Perez-Cardenas, E., Taja-Chayeb, L., Arias-Bofill, D., Candelaria, M., et al. (2011). A double-blind, placebo-controlled, randomized phase III trial of chemotherapy plus epigenetic therapy with hydralazine valproate for advanced cervical cancer. Preliminary results. Med. Oncol. Northwood Lond. Engl. *28 Suppl 1*, S540-546.

Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D., et al. (2011). Thymine DNA Glycosylase Is Essential for Active DNA Demethylation by Linked Deamination-Base Excision Repair. Cell *146*, 67–79.

Couronné, L., Bastard, C., and Bernard, O.A. (2012). *TET2* and *DNMT3A* Mutations in Human T-Cell Lymphoma. N. Engl. J. Med. *366*, 95–96.

Cox, D.B.T., Gootenberg, J.S., Abudayyeh, O.O., Franklin, B., Kellner, M.J., Joung, J., and Zhang, F. (2017). RNA editing with CRISPR-Cas13. Science *358*, 1019–1027.

Creusot, F., Acs, G., and Christman, J.K. (1982). Inhibition of DNA methyltransferase and induction of Friend erythroleukemia cell differentiation by 5-azacytidine and 5-aza-2'-deoxycytidine. J. Biol. Chem. *257*, 2041–2048.

Csankovszki, G., Nagy, A., and Jaenisch, R. (2001). Synergism of Xist Rna, DNA Methylation, and Histone Hypoacetylation in Maintaining X Chromosome Inactivation. J. Cell Biol. *153*, 773–784.

Cuculis, L., Abil, Z., Zhao, H., and Schroeder, C.M. (2016). TALE proteins search DNA using a rotationally decoupled mechanism. Nat. Chem. Biol. *12*, 831–837.

Cui, H. (2007). Loss of Imprinting of IGF2 as an Epigenetic Marker for the Risk of Human Cancer. Dis. Markers *23*, 105–112.

Cui, S.-Y., Wang, R., and Chen, L.-B. (2014). MicroRNA-145: a potent tumour suppressor that regulates multiple cellular pathways. J. Cell. Mol. Med. *18*, 1913–1926.

Cui, Y., Li, T., Yang, D., Li, S., and Le, W. (2016). miR-29 regulates Tet1 expression and contributes to early differentiation of mouse ESCs. Oncotarget *7*, 64932–64941.

Cunnac, S., Occhialini, A., Barberis, P., Boucher, C., and Genin, S. Inventory and functional analysis of the large Hrp regulon in Ralstonia solanacearum: identification of novel effector proteins translocated to plant host cells through the type III secretion system. Mol. Microbiol. *53*, 115–128.

Dai, H.-Q., Wang, B.-A., Yang, L., Chen, J.-J., Zhu, G.-C., Sun, M.-L., Ge, H., Wang, R., Chapman, D.L., Tang, F., et al. (2016). TET-mediated DNA demethylation controls gastrulation by regulating Lefty–Nodal signalling. Nature *538*, 528–532.

Daniūnaitė, K., Berezniakovas, A., Jankevičius, F., Laurinavičius, A., Lazutka, J.R., and Jarmalaitė, S. (2011). Frequent methylation of RASSF1 and RARB in urine sediments from patients with early stage prostate cancer. Med. Kaunas Lith. *47*, 147–153.

Das, P.M., Ramachandran, K., VanWert, J., Ferdinand, L., Gopisetty, G., Reis, I.M., and Singal, R. (2006). Methylation mediated silencing of TMS1/ASC gene in prostate cancer. Mol. Cancer *5*, 28.

Datta, J., Ghoshal, K., Denny, W.A., Gamage, S.A., Brooke, D.G., Phiasivongsa, P., Redkar, S., and Jacob, S.T. (2009). A New Class of Quinoline-Based DNA Hypomethylating Agents Reactivates Tumor Suppressor Genes by Blocking DNA Methyltransferase 1 Activity and Inducing Its Degradation. Cancer Res. *69*, 4277–4285.

Davies, S.M. (2001). Therapy-related leukemia associated with alkylating agents. Med. Pediatr. Oncol. *36*, 536–540.

Dawlaty, M.M., Ganz, K., Powell, B.E., Hu, Y.-C., Markoulaki, S., Cheng, A.W., Gao, Q., Kim, J., Choi, S.-W., Page, D.C., et al. (2011). Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. Cell Stem Cell *9*, 166–175.

Dawlaty, M.M., Breiling, A., Le, T., Raddatz, G., Barrasa, M.I., Cheng, A.W., Gao, Q., Powell, B.E., Li, Z., Xu, M., et al. (2013). Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. Dev. Cell *24*, 310–323.

de Greef, J.C., Wang, J., Balog, J., den Dunnen, J.T., Frants, R.R., Straasheijm, K.R., Aytekin, C., van der Burg, M., Duprez, L., Ferster, A., et al. (2011). Mutations in ZBTB24 Are Associated with Immunodeficiency, Centromeric Instability, and Facial Anomalies Syndrome Type 2. Am. J. Hum. Genet. *88*, 796–804.

Delhommeau, F., Dupont, S., Della Valle, V., James, C., Trannoy, S., Massé, A., Kosmider, O., Le Couedic, J.-P., Robert, F., Alberdi, A., et al. (2009). Mutation in TET2 in myeloid cancers. N. Engl. J. Med. *360*, 2289–2301.

Deng, C.-X. (2009). SIRT1, Is It a Tumor Promoter or Tumor Suppressor? Int. J. Biol. Sci. *5*, 147–152.

Deng, D., Yan, C., Pan, X., Mahfouz, M., Wang, J., Zhu, J.-K., Shi, Y., and Yan, N. (2012a). Structural Basis for Sequence-Specific Recognition of DNA by TAL Effectors. Science *335*, 720–723.

Deng, D., Yin, P., Yan, C., Pan, X., Gong, X., Qi, S., Xie, T., Mahfouz, M., Zhu, J.-K., Yan, N., et al. (2012b). Recognition of methylated DNA by TAL effectors. Cell Res. *22*, 1502–1504.

Deng, Q.-K., Lei, Y.-G., Lin, Y.-L., Ma, J.-G., and Li, W.-P. (2016). Prognostic Value of Protocadherin10 (PCDH10) Methylation in Serum of Prostate Cancer Patients. Med. Sci. Monit. Int. Med. J. Exp. Clin. Res. *22*, 516–521.

Deng, W., Rupon, J.W., Krivega, I., Breda, L., Motta, I., Jahn, K.S., Reik, A., Gregory, P.D., Rivella, S., Dean, A., et al. (2014). Reactivation of developmentally silenced globin genes by forced chromatin looping. Cell *158*, 849–860.

Desai, P., Mencia-Trinchant, N., Savenkov, O., Simon, M.S., Cheang, G., Lee, S., Samuel, M., Ritchie, E.K., Guzman, M.L., Ballman, K.V., et al. (2018). Somatic mutations precede acute myeloid leukemia years before diagnosis. Nat. Med. *24*, 1015–1023.

Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R.Z., Ragozin, S., and Jeltsch, A. (2010). The Dnmt3a PWWP Domain Reads Histone 3 Lysine 36 Trimethylation and Guides DNA Methylation. J. Biol. Chem. *285*, 26114–26120.

Diaz, L.A., and Bardelli, A. (2014). Liquid Biopsies: Genotyping Circulating Tumor DNA. J. Clin. Oncol. *32*, 579–586.

Diaz-Lagares, A., Mendez-Gonzalez, J., Hervas, D., Saigi, M., Pajares, M.J., Garcia, D., Crujerias, A.B., Pio, R., Montuenga, L.M., Zulueta, J., et al. (2016). A Novel Epigenetic Signature for Early Diagnosis in Lung Cancer. Clin. Cancer Res. *22*, 3361–3371.

Diehl, F., Schmidt, K., Choti, M.A., Romans, K., Goodman, S., Li, M., Thornton, K., Agrawal, N., Sokoll, L., Szabo, S.A., et al. (2008). Circulating mutant DNA to assess tumor dynamics. Nat. Med. *14*, 985–990.

Dolnik, A., Engelmann, J.C., Scharfenberger-Schmeer, M., Mauch, J., Kelkenberg-Schade, S., Haldemann, B., Fries, T., Kronke, J., Kuhn, M.W.M., Paschka, P., et al. (2012). Commonly altered genomic regions in acute myeloid leukemia are enriched for somatic mutations involved in chromatin remodeling and splicing. Blood *120*, e83–e92.

Domann, F.E., Rice, J.C., Hendrix, M.J., and Futscher, B.W. (2000). Epigenetic silencing of maspin gene expression in human breast cancers. Int. J. Cancer *85*, 805–810.

Domcke, S., Bardet, A.F., Adrian Ginno, P., Hartl, D., Burger, L., and Schübeler, D. (2015). Competition between DNA methylation and transcription factors determines binding of NRF1. Nature *528*, 575–579.

Dong, A. (2001). Structure of human DNMT2, an enigmatic DNA methyltransferase homolog that displays denaturant-resistant binding to DNA. Nucleic Acids Res. *29*, 439–448.

Doskocil, J., and Sorm, F. (1962). Distribution of 5-methylcytosine in pyrimidine sequences of deoxyribonucleic acids. Biochim. Biophys. Acta *55*, 953–959.

Doudna, J.A., and Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. Science *346*, 1258096–1258096.

Duan, J., Lu, G., Xie, Z., Lou, M., Luo, J., Guo, L., and Zhang, Y. (2014). Genome-wide identification of CRISPR/Cas9 off-targets in human genome. Cell Res. *24*, 1009–1012.

Duan, Y.P., Castañeda, A., Zhao, G., Erdos, G., and Gabriel, D.W. (1999). Expression of a Single, Host-Specific, Bacterial Pathogenicity Gene in Plant Cells Elicits Division, Enlargement, and Cell Death. Mol. Plant. Microbe Interact. *12*, 556–560.

Duvic, M., and Talpur, R. (2008). Optimizing denileukin diftitox (Ontak®) therapy. Future Oncol. *4*, 457–469.

Dzinic, S.H., Bernardo, M.M., Li, X., Fernandez-Valdivia, R., Ho, Y.-S., Mi, Q.-S., Bandyopadhyay, S., Lonardo, F., Vranic, S., Oliveira, D.S.M., et al. (2017). An Essential Role of *Maspin* in Embryogenesis and Tumor Suppression. Cancer Res. *77*, 886–896.

Ehrlich, M. (2003). The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease. Clin. Immunol. *109*, 17–28.

Emperle, M., Rajavelu, A., Kunert, S., Arimondo, P.B., Reinhardt, R., Jurkowska, R.Z., and Jeltsch, A. (2018). The DNMT3A R882H mutant displays altered flanking sequence preferences. Nucleic Acids Res. *46*, 3130–3139.

Erdmann, A., Menon, Y., Gros, C., Molinier, N., Novosad, N., Samson, A., Gregoire, J.-M., Long, C., Ausseil, F., Halby, L., et al. (2015). Design and synthesis of new non nucleoside inhibitors of DNMT3A. Bioorg. Med. Chem. *23*, 5946–5953.

Esteban, M.A., Wang, T., Qin, B., Yang, J., Qin, D., Cai, J., Li, W., Weng, Z., Chen, J., Ni, S., et al. (2010). Vitamin C Enhances the Generation of Mouse and Human Induced Pluripotent Stem Cells. Cell Stem Cell *6*, 71–79.

Esteller, M. (2007). Cancer epigenomics: DNA methylomes and histone-modification maps. Nat. Rev. Genet. *8*, 286–298.

Esteller, M., Silva, J.M., Dominguez, G., Bonilla, F., Matias-Guiu, X., Lerma, E., Bussaglia, E., Prat, J., Harkes, I.C., Repasky, E.A., et al. (2000a). Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. J. Natl. Cancer Inst. *92*, 564–569.

Esteller, M., Garcia-Foncillas, J., Andion, E., Goodman, S.N., Hidalgo, O.F., Vanaclocha, V., Baylin, S.B., and Herman, J.G. (2000b). Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. N. Engl. J. Med. *343*, 1350–1354.

Fairall, L., Schwabe, J.W., Chapman, L., Finch, J.T., and Rhodes, D. (1993). The crystal structure of a two zinc-finger peptide reveals an extension to the rules for zinc-finger/DNA recognition. Nature *366*, 483–487.

Fang, F., Fan, S., Zhang, X., and Zhang, M.Q. (2006). Predicting methylation status of CpG islands in the human brain. Bioinformatics *22*, 2204–2209.

Fang, F., Turcan, S., Rimner, A., Kaufman, A., Giri, D., Morris, L.G.T., Shen, R., Seshan, V., Mo, Q., Heguy, A., et al. (2011). Breast Cancer Methylomes Establish an Epigenomic Foundation for Metastasis. Sci. Transl. Med. *3*, 75ra25.

Fang, M.Z., Wang, Y., Ai, N., Hou, Z., Sun, Y., Lu, H., Welsh, W., and Yang, C.S. (2003). Tea Polyphenol (−)-Epigallocatechin-3-Gallate Inhibits DNA Methyltransferase and Reactivates Methylation-Silenced Genes in Cancer Cell Lines. Cancer Res. *63*, 7563–7570.

Farzadfard, F., Perli, S.D., and Lu, T.K. (2013). Tunable and Multifunctional Eukaryotic Transcription Factors Based on CRISPR/Cas. ACS Synth. Biol. *2*, 604–613.

Fatemi, M., Hermann, A., Gowher, H., and Jeltsch, A. (2002). Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA. Eur. J. Biochem. *269*, 4981–4984.

Feinberg, A.P., and Vogelstein, B. (1983). Hypomethylation distinguishes genes of some human cancers from their normal counterparts. Nature *301*, 89–92.

Feldmann, A., Ivanek, R., Murr, R., Gaidatzis, D., Burger, L., and Schübeler, D. (2013). Transcription Factor Occupancy Can Mediate Active Turnover of DNA Methylation at Regulatory Regions. PLoS Genet. *9*, e1003994.

Feng, S., Cokus, S.J., Zhang, X., Chen, P.-Y., Bostick, M., Goll, M.G., Hetzel, J., Jain, J., Strauss, S.H., Halpern, M.E., et al. (2010). Conservation and divergence of methylation patterning in plants and animals. Proc. Natl. Acad. Sci. U. S. A. *107*, 8689–8694.

Ficz, G., Branco, M.R., Seisenberger, S., Santos, F., Krueger, F., Hore, T.A., Marques, C.J., Andrews, S., and Reik, W. (2011). Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature *473*, 398–402.

Figueroa, M.E., Lugthart, S., Li, Y., Erpelinck-Verschueren, C., Deng, X., Christos, P.J., Schifano, E., Booth, J., van Putten, W., Skrabanek, L., et al. (2010). DNA Methylation Signatures Identify Biologically Distinct Subtypes in Acute Myeloid Leukemia. Cancer Cell *17*, 13–27.

Finnegan, E.J., and Dennis, E.S. (1993). Isolation and identification by sequence homology of a putative cytosine methyltransferase from Arabidopsis thaliana. Nucleic Acids Res. *21*, 2383–2388.

Frauer, C., Rottach, A., Meilinger, D., Bultmann, S., Fellinger, K., Hasenöder, S., Wang, M., Qin, W., Söding, J., Spada, F., et al. (2011). Different Binding Properties and Function of CXXC Zinc Finger Domains in Dnmt1 and Tet1. PLoS ONE *6*, e16627.

Friedman, S. (1981). The Inhibition of DNA(Cytosine-5)Methylases by 5-Azacytidine: The Effect of Azacytosine-Containing DNA. Mol. Pharmacol. *19*, 314–320.

Friedrich, M.G., Weisenberger, D.J., Cheng, J.C., Chandrasoma, S., Siegmund, K.D., Gonzalgo, M.L., Toma, M.I., Huland, H., Yoo, C., Tsai, Y.C., et al. (2004). Detection of Methylated Apoptosis-Associated Genes in Urine Sediments of Bladder Cancer Patients. Clin. Cancer Res. *10*, 7457–7465.

Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L., and Paul, C.L. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc. Natl. Acad. Sci. U. S. A. *89*, 1827–1831.

Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., and Joung, J.K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat. Biotechnol. *32*, 279–284.

Fu, Y., Rocha, P.P., Luo, V.M., Raviram, R., Deng, Y., Mazzoni, E.O., and Skok, J.A. (2016). CRISPR-dCas9 and sgRNA scaffolds enable dual-colour live imaging of satellite sequences and repeat-enriched individual loci. Nat. Commun. *7*, 11707.

Fujimoto, M., Kitazawa, R., Maeda, S., and Kitazawa, S. (2005). Methylation adjacent to negatively regulating AP-1 site reactivates TrkA gene expression during cancer progression. Oncogene *24*, 5108–5118.

Füller, M., Füller, M., Klein, M., Klein, M., Schmidt, E., Schmidt, E., Rohde, C., Rohde, C., Göllner, S., Göllner, S., et al. (2015). 5-Azacytidine enhances efficacy of multiple chemotherapy drugs in AML and lung cancer with modulation of CpG methylation. Int. J. Oncol. *46*, 1192–1204.

Futscher, B.W., O'Meara, M.M., Kim, C.J., Rennels, M.A., Lu, D., Gruman, L.M., Seftor, R.E.B., Hendrix, M.J.C., and Domann, F.E. (2004). Aberrant Methylation of the Maspin Promoter Is an Early Event in Human Breast Cancer. Neoplasia *6*, 380–389.

Galamb, O., Kalmár, A., Péterfia, B., Csabai, I., Bodor, A., Ribli, D., Krenács, T., Patai, Á.V., Wichmann, B., Barták, B.K., et al. (2016). Aberrant DNA methylation of WNT pathway genes in the development and progression of CIMP-negative colorectal cancer. Epigenetics *11*, 588–602.

Gallego-Bartolomé, J., Gardiner, J., Liu, W., Papikian, A., Ghoshal, B., Kuo, H.Y., Zhao, J.M.-C., Segal, D.J., and Jacobsen, S.E. (2018). Targeted DNA demethylation of the *Arabidopsis* genome using the human TET1 catalytic domain. Proc. Natl. Acad. Sci. *115*, E2125–E2134.

Galonska, C., Charlton, J., Mattei, A.L., Donaghey, J., Clement, K., Gu, H., Mohammad, A.W., Stamenova, E.K., Cacchiarelli, D., Klages, S., et al. (2018). Genome-wide tracking of dCas9-methyltransferase footprints. Nat. Commun. *9*, 597.

Gama-Sosa, M.A., Midgett, R.M., Slagel, V.A., Githens, S., Kuo, K.C., Gehrke, C.W., and Ehrlich, M. (1983). Tissue-specific differences in DNA methylation in various mammals. Biochim. Biophys. Acta *740*, 212–219.

Gao, F. (2004). Maspin plays an essential role in early embryonic development. Development *131*, 1479–1489.

Gao, F., Ji, G., Gao, Z., Han, X., Ye, M., Yuan, Z., Luo, H., Huang, X., Natarajan, K., Wang, J., et al. (2014a). Direct ChIP-bisulfite sequencing reveals a role of H3K27me3 mediating aberrant hypermethylation of promoter CpG islands in cancer cells. Genomics *103*, 204–210.

Gao, X., Yang, J., Tsang, J.C.H., Ooi, J., Wu, D., and Liu, P. (2013). Reprogramming to Pluripotency Using Designer TALE Transcription Factors Targeting Enhancers. Stem Cell Rep. *1*, 183–197.

Gao, X., Tsang, J.C.H., Gaba, F., Wu, D., Lu, L., and Liu, P. (2014b). Comparison of TALE designer transcription factors and the CRISPR/dCas9 in regulation of gene expression by targeting enhancers. Nucleic Acids Res. *42*, e155.

Gao, X., Tao, Y., Lamas, V., Huang, M., Yeh, W.-H., Pan, B., Hu, Y.-J., Hu, J.H., Thompson, D.B., Shu, Y., et al. (2017). Treatment of autosomal dominant hearing loss by in vivo delivery of genome editing agents. Nature *553*, 217–221.

Garcia-Bloj, B., Moses, C., Sgro, A., Plani-Lam, J., Arooj, M., Duffy, C., Thiruvengadam, S., Sorolla, A., Rashwan, R., Mancera, R.L., et al. (2016). Waking up dormant tumor suppressor genes with zinc fingers, TALEs and the CRISPR/dCas9 system. Oncotarget *7*, 60535–60554.

García-Domínguez, P., Dell'Aversana, C., Alvarez, R., Altucci, L., and de Lera, Á.R. (2013). Synthetic approaches to DNMT inhibitor SGI-1027 and effects on the U937 leukemia cell line. Bioorg. Med. Chem. Lett. *23*, 1631–1635.

Gaston, K., and Fried, M. (1995). CpG methylation has differential effects on the binding of YY1 and ETS proteins to the bi-directional promoter of the Surf-1 and Surf-2 genes. Nucleic Acids Res. *23*, 901–909.

Gatto, S., Gagliardi, M., Franzese, M., Leppert, S., Papa, M., Cammisa, M., Grillo, G., Velasco, G., Francastel, C., Toubiana, S., et al. (2017). ICF-specific DNMT3B dysfunction interferes with intragenic regulation of mRNA transcription and alternative splicing. Nucleic Acids Res. *45*, 5739–5756.

Gaudet, F. (2003). Induction of Tumors in Mice by Genomic Hypomethylation. Science *300*, 489–492.

Gebhard, C., Benner, C., Ehrich, M., Schwarzfischer, L., Schilling, E., Klug, M., Dietmaier, W., Thiede, C., Holler, E., Andreesen, R., et al. (2010). General Transcription Factor Binding at CpG Islands in Normal Cells Correlates with Resistance to De novo DNA Methylation in Cancer Cells. Cancer Res. *70*, 1398–1407.

Gendrel, A.-V., Apedaile, A., Coker, H., Termanis, A., Zvetkova, I., Godwin, J., Tang, Y.A., Huntley, D., Montana, G., Taylor, S., et al. (2012). Smchd1-Dependent and -Independent Pathways Determine Developmental Dynamics of CpG Island Methylation on the Inactive X Chromosome. Dev. Cell *23*, 265–279.

Genovese, G., Kähler, A.K., Handsaker, R.E., Lindberg, J., Rose, S.A., Bakhoum, S.F., Chambert, K., Mick, E., Neale, B.M., Fromer, M., et al. (2014). Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. N. Engl. J. Med. *371*, 2477–2487.

Gersbach, C.A., Gaj, T., and Barbas, C.F. (2014). Synthetic Zinc Finger Proteins: The Advent of Targeted Gene Regulation and Genome Modification Technologies. Acc. Chem. Res. *47*, 2309–2318.

Ghiam, A.F., Cairns, R.A., Thoms, J., Dal Pra, A., Ahmed, O., Meng, A., Mak, T.W., and Bristow, R.G. (2012). IDH mutation status in prostate cancer. Oncogene *31*, 3826–3826.

Gilbert, L.A., Larson, M.H., Morsut, L., Liu, Z., Brar, G.A., Torres, S.E., Stern-Ginossar, N., Brandman, O., Whitehead, E.H., Doudna, J.A., et al. (2013). CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. Cell *154*, 442–451.

Gilbert, L.A., Horlbeck, M.A., Adamson, B., Villalta, J.E., Chen, Y., Whitehead, E.H., Guimaraes, C., Panning, B., Ploegh, H.L., Bassik, M.C., et al. (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell *159*, 647–661.

Glasspool, R.M., Brown, R., Gore, M.E., Rustin, G.J.S., McNeish, I.A., Wilson, R.H., Pledge, S., Paul, J., Mackean, M., Hall, G.D., et al. (2014). A randomised, phase II trial of the DNA-hypomethylating agent 5-aza-2′-deoxycytidine (decitabine) in combination with carboplatin vs carboplatin alone in patients with recurrent, partially platinum-sensitive ovarian cancer. Br. J. Cancer *110*, 1923–1929.

Globisch, D., Münzel, M., Müller, M., Michalakis, S., Wagner, M., Koch, S., Brückl, T., Biel, M., and Carell, T. (2010). Tissue Distribution of 5-Hydroxymethylcytosine and Search for Active Demethylation Intermediates. PLoS ONE *5*, e15367.

Gold, M., and Hurwitz, J. (1964). THE ENZYMATIC METHYLATION OF RIBONUCLEIC ACID AND DEOXYRIBONUCLEIC ACID. V. PURIFICATION AND PROPERTIES OF THE DEOXYRIBONUCLEIC ACID-METHYLATING ACTIVITY OF ESCHERICHIA COLI. J. Biol. Chem. *239*, 3858–3865.

Goll, M.G. (2006). Methylation of tRNAAsp by the DNA Methyltransferase Homolog Dnmt2. Science *311*, 395–398.

Good, C.R., Madzo, J., Patel, B., Maegawa, S., Engel, N., Jelinek, J., and Issa, J.-P.J. (2017). A novel isoform of TET1 that lacks a CXXC domain is overexpressed in cancer. Nucleic Acids Res. *45*, 8269–8281.

GORDIAN, E., RAMACHANDRAN, K., and SINGAL, R. (2009). Methylation Mediated Silencing of TMS1 in Breast Cancer and its Potential Contribution to Docetaxel Cytotoxicity. ANTICANCER Res. *29*, 3207–3210.

Gore, L., Triche, T.J., Farrar, J.E., Wai, D., Legendre, C., Gooden, G.C., Liang, W.S., Carpten, J., Lee, D., Alvaro, F., et al. (2017). A multicenter, randomized study of decitabine as epigenetic priming with induction chemotherapy in children with AML. Clin. Epigenetics *9*.

Gossen, M., and Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc. Natl. Acad. Sci. U. S. A. *89*, 5547–5551.

Gossen, M., Freundlieb, S., Bender, G., Muller, G., Hillen, W., and Bujard, H. (1995). Transcriptional activation by tetracyclines in mammalian cells. Science *268*, 1766–1769.

Graff, J.R., Herman, J.G., Lapidus, R.G., Chopra, H., Xu, R., Jarrard, D.F., Isaacs, W.B., Pitha, P.M., Davidson, N.E., and Baylin, S.B. (1995). E-cadherin expression is silenced by DNA hypermethylation in human breast and prostate carcinomas. Cancer Res. *55*, 5195–5199.

Green, A., and Beer, P. (2010). Somatic Mutations of *IDH1* and *IDH2* in the Leukemic Transformation of Myeloproliferative Neoplasms. N. Engl. J. Med. *362*, 369–370.

Greer, E.L., Blanco, M.A., Gu, L., Sendinc, E., Liu, J., Aristizábal-Corrales, D., Hsu, C.-H., Aravind, L., He, C., and Shi, Y. (2015). DNA methylation on N6-adenine in C. elegans. Cell *161*, 868–878.

Greger, V., Passarge, E., H**?**pping, W., Messmer, E., and Horsthemke, B. (1989). Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. Hum. Genet. *83*, 155–158.

Gregory, D.J., Mikhaylova, L., and Fedulov, A.V. (2012). Selective DNA demethylation by fusion of TDG with a sequence-specific DNA-binding domain. Epigenetics *7*, 344–349.

Griffiths, E.A., Choy, G., Redkar, S., Taverna, P., Azab, M., and Karpf, A.R. (2013). SGI-110. Drugs Future *38*, 535–543.

Gu, H., Smith, Z.D., Bock, C., Boyle, P., Gnirke, A., and Meissner, A. (2011a). Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. Nat. Protoc. *6*, 468–481.

Gu, K., Yang, B., Tian, D., Wu, L., Wang, D., Sreekala, C., Yang, F., Chu, Z., Wang, G.-L., White, F.F., et al. (2005). R gene expression induced by a type-III effector triggers disease resistance in rice. Nature *435*, 1122–1125.

Gu, T.-P., Guo, F., Yang, H., Wu, H.-P., Xu, G.-F., Liu, W., Xie, Z.-G., Shi, L., He, X., Jin, S., et al. (2011b). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature *477*, 606–610.

Guan, Z., Zhang, J., Song, S., and Dai, D. (2013). Promoter methylation and expression of TIMP3 gene in gastric cancer. Diagn. Pathol. *8*, 110.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A Chromatin Landmark and Transcription Initiation at Most Promoters in Human Cells. Cell *130*, 77–88.

Guerrero-Preston, R., Soudry, E., Acero, J., Orera, M., Moreno-López, L., Macía-Colón, G., Jaffe, A., Berdasco, M., Ili-Gangas, C., Brebi-Mieville, P., et al. (2011). NID2 and HOXA9 promoter hypermethylation as biomarkers for prevention and early detection in Oral Cavity Squamous Cell Carcinoma tissues and saliva. Cancer Prev. Res. Phila. Pa *4*, 1061–1072.

Guibert, S., Forné, T., and Weber, M. (2012). Global profiling of DNA methylation erasure in mouse primordial germ cells. Genome Res. *22*, 633–641.

Guilinger, J.P., Pattanayak, V., Reyon, D., Tsai, S.Q., Sander, J.D., Joung, J.K., and Liu, D.R. (2014). Broad Specificity Profiling of TALENs Results in Engineered Nucleases With Improved DNA Cleavage Specificity. Nat. Methods *11*, 429–435.

Guo, F., Li, X., Liang, D., Li, T., Zhu, P., Guo, H., Wu, X., Wen, L., Gu, T.-P., Hu, B., et al. (2014). Active and Passive Demethylation of Male and Female Pronuclear DNA in the Mammalian Zygote. Cell Stem Cell *15*, 447–459.

Guo, J.U., Su, Y., Zhong, C., Ming, G., and Song, H. (2011). Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell *145*, 423–434.

Guo, X., Wang, L., Li, J., Ding, Z., Xiao, J., Yin, X., He, S., Shi, P., Dong, L., Li, G., et al. (2015). Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. Nature *517*, 640–644.

Hackett, J.A., Zylicz, J.J., and Surani, M.A. (2012). Parallel mechanisms of epigenetic reprogramming in the germline. Trends Genet. *28*, 164–174.

Hajkova, P., Jeffries, S.J., Lee, C., Miller, N., Jackson, S.P., and Surani, M.A. (2010). Genome-Wide Reprogramming in the Mouse Germ Line Entails the Base Excision Repair Pathway. Science *329*, 78–82.

Halby, L., Champion, C., Sénamaud-Beaufort, C., Ajjan, S., Drujon, T., Rajavelu, A., Ceccaldi, A., Jurkowska, R., Lequin, O., Nelson, W.G., et al. (2012). Rapid Synthesis of New DNMT Inhibitors Derivatives of Procainamide. ChemBioChem *13*, 157–165.

Hansen, R.S., Wijmenga, C., Luo, P., Stanek, A.M., Canfield, T.K., Weemaes, C.M.R., and Gartler, S.M. (1999). The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. Proc. Natl. Acad. Sci. *96*, 14412–14417.

Hao, X., Luo, H., Krawczyk, M., Wei, W., Wang, W., Wang, J., Flagg, K., Hou, J., Zhang, H., Yi, S., et al. (2017). DNA methylation markers for diagnosis and prognosis of common cancers. Proc. Natl. Acad. Sci. U. S. A. *114*, 7414–7419.

Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature *405*, 486–489.

Hashimoto, H., Liu, Y., Upadhyay, A.K., Chang, Y., Howerton, S.B., Vertino, P.M., Zhang, X., and Cheng, X. (2012). Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. Nucleic Acids Res. *40*, 4841–4849.

Hata, K., Okano, M., Lei, H., and Li, E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development *129*, 1983–1993.

He, T., Zhang, M., Zheng, R., Zheng, S., Linghu, E., Herman, J.G., and Guo, M. (2017). Methylation of *SLFN11* is a marker of poor prognosis and cisplatin resistance in colorectal cancer. Epigenomics *9*, 849–862.

He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., et al. (2011). Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA. Science *333*, 1303–1307.

Hegi, M.E., Diserens, A.-C., Gorlia, T., Hamou, M.-F., de Tribolet, N., Weller, M., Kros, J.M., Hainfellner, J.A., Mason, W., Mariani, L., et al. (2005). MGMT gene silencing and benefit from temozolomide in glioblastoma. N. Engl. J. Med. *352*, 997–1003.

Hellebrekers, D.M.E.I., Lentjes, M.H.F.M., van den Bosch, S.M., Melotte, V., Wouters, K.A.D., Daenen, K.L.J., Smits, K.M., Akiyama, Y., Yuasa, Y., Sanduleanu, S., et al. (2009). GATA4 and GATA5 are Potential Tumor Suppressors and Biomarkers in Colorectal Cancer. Clin. Cancer Res. *15*, 3990–3997.

Heller, E.A., Cates, H.M., Peña, C.J., Sun, H., Shao, N., Feng, J., Golden, S.A., Herman, J.P., Walsh, J.J., Mazei-Robison, M., et al. (2014). Locus-Specific Epigenetic Remodeling Controls Addiction- and Depression-Related Behaviors (NN-A50213-T). Nat. Neurosci. *17*, 1720–1727.

Hellman, A., and Chess, A. (2007). Gene body-specific methylation on the active X chromosome. Science *315*, 1141–1143.

Hemerly, J.P., Bastos, A.U., and Cerutti, J.M. (2010). Identification of several novel non-p.R132 IDH1 variants in thyroid carcinomas. Eur. J. Endocrinol. *163*, 747–755.

Herranz, M. (2005). The novel DNA methylation inhibitor zebularine is effective against the development of murine T-cell lymphoma. Blood *107*, 1174–1177.

Hesson, L.B., Cooper, W.N., and Latif, F. (2007). The Role of RASSF1A Methylation in Cancer. Dis. Markers *23*, 73–87.

Heyn, H., and Esteller, M. (2012). DNA methylation profiling in the clinic: applications and challenges. Nat. Rev. Genet. *13*, 679–692.

Heyn, H., Vidal, E., Ferreira, H.J., Vizoso, M., Sayols, S., Gomez, A., Moran, S., Boque-Sastre, R., Guil, S., Martinez-Cardus, A., et al. (2016). Epigenomic analysis detects aberrant super-enhancer DNA methylation in human cancer. Genome Biol. *17*, 11.

Hill, P.W.S., Leitch, H.G., Requena, C.E., Sun, Z., Amouroux, R., Roman-Trufero, M., Borkowska, M., Terragni, J., Vaisvila, R., Linnett, S., et al. (2018). Epigenetic reprogramming enables the transition from primordial germ cell to gonocyte. Nature *555*, 392–396.

Hilton, I.B., D'Ippolito, A.M., Vockley, C.M., Thakore, P.I., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2015). Epigenome editing by a CRISPR/Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat. Biotechnol. *33*, 510–517.

Hinoue, T., Weisenberger, D.J., Lange, C.P.E., Shen, H., Byun, H.-M., Van Den Berg, D., Malik, S., Pan, F., Noushmehr, H., van Dijk, C.M., et al. (2012). Genome-scale analysis of aberrant DNA methylation in colorectal cancer. Genome Res. *22*, 271–282.

Hirasawa, R., Chiba, H., Kaneda, M., Tajima, S., Li, E., Jaenisch, R., and Sasaki, H. (2008). Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. Genes Dev. *22*, 1607–1616.

Holliday, R., and Pugh, J.E. (1975). DNA modification mechanisms and gene activity during development. Science *187*, 226–232.

Holm, T.M., Jackson-Grusby, L., Brambrink, T., Yamada, Y., Rideout, W.M., and Jaenisch, R. (2005). Global loss of imprinting leads to widespread tumorigenesis in adult mice. Cancer Cell *8*, 275–285.

Hon, G.C., Hawkins, R.D., Caballero, O.L., Lo, C., Lister, R., Pelizzola, M., Valsesia, A., Ye, Z., Kuan, S., Edsall, L.E., et al. (2012). Global DNA hypomethylation coupled to repressive chromatin domain formation and gene silencing in breast cancer. Genome Res. *22*, 246–258.

Hon, G.C., Rajagopal, N., Shen, Y., McCleary, D.F., Yue, F., Dang, M.D., and Ren, B. (2013). Adult tissue methylomes harbor epigenetic memory at embryonic enhancers. Nat. Genet. *45*, 1198–1206.

Hoque, M.O., Begum, S., Topaloglu, O., Chatterjee, A., Rosenbaum, E., Van Criekinge, W., Westra, W.H., Schoenberg, M., Zahurak, M., Goodman, S.N., et al. (2006). Quantitation of Promoter Methylation of Multiple Genes in Urine DNA and Bladder Cancer Detection. JNCI J. Natl. Cancer Inst. *98*, 996–1004.

Hotchkiss, R.D. (1948). The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. J. Biol. Chem. *175*, 315–332.

Hu, J., Lei, Y., Wong, W.-K., Liu, S., Lee, K.-C., He, X., You, W., Zhou, R., Guo, J.-T., Chen, X., et al. (2014). Direct activation of human and mouse Oct4 genes using engineered TALE and Cas9 transcription factors. Nucleic Acids Res. *42*, 4375–4390.

Huai, C., Jia, C., Sun, R., Xu, P., Min, T., Wang, Q., Zheng, C., Chen, H., and Lu, D. (2017). CRISPR/Cas9-mediated somatic and germline gene correction to restore hemostasis in hemophilia B mice. Hum. Genet. *136*, 875–883.

Huang, H., Jiang, X., Li, Z., Li, Y., Song, C.-X., He, C., Sun, M., Chen, P., Gurbuxani, S., Wang, J., et al. (2013). TET1 plays an essential oncogenic role in MLL-rearranged leukemia. Proc. Natl. Acad. Sci. U. S. A. *110*, 11994–11999.

Huang, Y., Chavez, L., Chang, X., Wang, X., Pastor, W.A., Kang, J., Zepeda-Martínez, J.A., Pape, U.J., Jacobsen, S.E., Peters, B., et al. (2014). Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. *111*, 1361–1366.

Huang, Y.-H., Su, J., Lei, Y., Brunetti, L., Gundry, M.C., Zhang, X., Jeong, M., Li, W., and Goodell, M.A. (2017). DNA epigenome editing using CRISPR-Cas SunTag-directed DNMT3A. Genome Biol. *18*, 176.

Huisman, C., van der Wijst, M.G.P., Schokker, M., Blancafort, P., Terpstra, M.M., Kok, K., van der Zee, A.G.J., Schuuring, E., Wisman, G.B.A., and Rots, M.G. (2016). Re-expression of Selected Epigenetically Silenced Candidate Tumor Suppressor Genes in Cervical Cancer by TET2-directed Demethylation. Mol. Ther. J. Am. Soc. Gene Ther. *24*, 536–547.

Humeniuk, R., Rosu-Myles, M., Fares, J., Koller, R., Bies, J., and Wolff, L. (2013). The role of tumor suppressor p15Ink4b in the regulation of hematopoietic progenitor cell fate. Blood Cancer J. *3*, e99–e99.

Hur, K., Cejas, P., Feliu, J., Moreno-Rubio, J., Burgos, E., Boland, C.R., and Goel, A. (2014). Hypomethylation of long interspersed nuclear element-1 (LINE-1) leads to activation of proto-oncogenes in human colorectal cancer metastasis. Gut *63*, 635–646.

Ibrahim, A.E.K., Arends, M.J., Silva, A.-L., Wyllie, A.H., Greger, L., Ito, Y., Vowler, S.L., Huang, T.H.-M., Tavare, S., Murrell, A., et al. (2011). Sequential DNA methylation changes are associated with DNMT3B overexpression in colorectal neoplastic progression. Gut *60*, 499–508.

Illingworth, R.S., Gruenewald-Schneider, U., Webb, S., Kerr, A.R.W., James, K.D., Turner, D.J., Smith, C., Harrison, D.J., Andrews, R., and Bird, A.P. (2010). Orphan CpG Islands Identify Numerous Conserved Promoters in the Mammalian Genome. PLOS Genet. *6*, e1001134.

Inoue, A., Shen, L., Dai, Q., He, C., and Zhang, Y. (2011). Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. Cell Res. *21*, 1670–1676.

Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES cell self-renewal, and ICM specification. Nature *466*, 1129–1133.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science *333*, 1300–1303.

Itzykson, R., Kosmider, O., Renneville, A., Morabito, M., Preudhomme, C., Berthon, C., Ades, L., Fenaux, P., Platzbecker, U., Gagey, O., et al. (2013). Clonal architecture of chronic myelomonocytic leukemias. Blood *121*, 2186–2198.

Iwagami, S., Baba, Y., Watanabe, M., Shigaki, H., Miyake, K., Ishimoto, T., Iwatsuki, M., Sakamaki, K., Ohashi, Y., and Baba, H. (2013). Line-1 Hypomethylation Is Associated With a Poor Prognosis Among Patients With Curatively Resected Esophageal Squamous Cell Carcinoma. Ann. Surg. *257*, 449–455.

Iyer, L.M., Tahiliani, M., Rao, A., and Aravind, L. (2009). Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. Cell Cycle Georget. Tex *8*, 1698–1710.

Jackson-Grusby, L., Laird, P.W., Magge, S.N., Moeller, B.J., and Jaenisch, R. (1997). Mutagenicity of 5-aza-2'-deoxycytidine is mediated by the mammalian DNA methyltransferase. Proc. Natl. Acad. Sci. U. S. A. *94*, 4681–4685.

Jain, D., Meydan, C., Lange, J., Claeys Bouuaert, C., Lailler, N., Mason, C.E., Anderson, K.V., and Keeney, S. (2017). rahu is a mutant allele of Dnmt3c, encoding a DNA methyltransferase homolog required for meiosis and transposon repression in the mouse male germline. PLoS Genet. *13*, e1006964.

Jankowska, A.M., Makishima, H., Tiu, R.V., Szpurka, H., Huang, Y., Traina, F., Visconte, V., Sugimoto, Y., Prince, C., O'Keefe, C., et al. (2011). Mutational spectrum analysis of chronic myelomonocytic leukemia includes genes associated with epigenetic regulation: UTX, EZH2, and DNMT3A. Blood *118*, 3932–3941.

Jatkoe, T.A., Karnes, R.J., Freedland, S.J., Wang, Y., Le, A., and Baden, J. (2015). A urine-based methylation signature for risk stratification within low-risk prostate cancer. Br. J. Cancer *112*, 802–808.

Jéhannin-Ligier, K., Dantony, E., Bossard, N., Molinié, F., Defossez, G., Daubisse-Marliac, L., Delafosse, P., Remontet, L., and Uhry, Z. (2017). Projection de l'incidence et de la mortalité en France métropolitaine en 2017 - Rapport technique - Ref : RAINCMETRO17 | Institut National Du Cancer.

Jeong, L.S., Buenger, G., McCormack, J.J., Cooney, D.A., Hao, Z., and Marquez, V.E. (1998). Carbocyclic Analogues of the Potent Cytidine Deaminase Inhibitor 1-(β- D -Ribofuranosyl)-1,2-dihydropyrimidin-2-one (Zebularine). J. Med. Chem. *41*, 2572–2578.

Jeziorska, D.M., Murray, R.J.S., De Gobbi, M., Gaentzsch, R., Garrick, D., Ayyub, H., Chen, T., Li, E., Telenius, J., Lynch, M., et al. (2017). DNA methylation of intragenic CpG islands depends on their transcriptional activity during differentiation and disease. Proc. Natl. Acad. Sci. *114*, E7526–E7535.

Ji, Q., Fischer, A.L., Brown, C.R., Eastlund, E.R., Dvash, T., Zhong, B., Gerber, M.A., Lyons, I., Knight, S.W., and Kreader, C.A. (2014). Engineered zinc-finger transcription factors activate OCT4 (POU5F1), SOX2, KLF4, c-MYC (MYC) and miR302/367. Nucleic Acids Res. *42*, 6158–6167.

Jia, D., Jurkowska, R.Z., Zhang, X., Jeltsch, A., and Cheng, X. (2007). Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. Nature *449*, 248–251.

Jin, B., Tao, Q., Peng, J., Soo, H.M., Wu, W., Ying, J., Fields, C.R., Delmas, A.L., Liu, X., Qiu, J., et al. (2008). DNA methyltransferase 3B (DNMT3B) mutations in ICF syndrome lead to altered epigenetic modifications and aberrant expression of genes regulating development, neurogenesis and immune function. Hum. Mol. Genet. *17*, 690–709.

Jin, S.-G., Zhang, Z.-M., Dunwell, T.L., Harter, M.R., Wu, X., Johnson, J., Li, Z., Liu, J., Szabó, P.E., Lu, Q., et al. (2016). Tet3 reads 5-carboxylcytosine through its CXXC domain and is a potential guardian against neurodegeneration. Cell Rep. *14*, 493–505.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013). RNA-programmed genome editing in human cells. ELife *2*, e00471.

Johnson, T.B., and Coghill, R.D. (1925). RESEARCHES ON PYRIMIDINES. C111. THE DISCOVERY OF 5-METHYL-CYTOSINE IN TUBERCULINIC ACID, THE NUCLEIC ACID OF THE TUBERCLE BACILLUS1. J. Am. Chem. Soc. *47*, 2838–2844.

Johnson, W.D., Harder, J.B., Naylor, J., McCormick, D.L., Detrisac, C.J., Glaze, E.R., and Tomaszewski, J.E. (2006). A pharmacokinetic/pharmacodynamic approach to evaluating the safety of zebularine in non-human primates. Cancer Res. *66*, 309–310.

Jones, P.A., and Taylor, S.M. (1980). Cellular differentiation, cytidine analogs and DNA methylation. Cell *20*, 85–93.

Joung, J., Konermann, S., Gootenberg, J.S., Abudayyeh, O.O., Platt, R.J., Brigham, M.D., Sanjana, N.E., and Zhang, F. (2017). Genome-scale CRISPR-Cas9 Knockout and Transcriptional Activation Screening. Nat. Protoc. *12*, 828–863.

Juergens, R.A., Wrangle, J., Vendetti, F.P., Murphy, S.C., Zhao, M., Coleman, B., Sebree, R., Rodgers, K., Hooker, C.M., Franco, N., et al. (2011). Combination Epigenetic Therapy Has Efficacy in Patients with Refractory Advanced Non Small Cell Lung Cancer. Cancer Discov. *1*, 598–607.

Julià, A., Absher, D., López-Lasanta, M., Palau, N., Pluma, A., Waite Jones, L., Glossop, J.R., Farrell, W.E., Myers, R.M., and Marsal, S. (2017). Epigenome-wide association study of rheumatoid arthritis identifies differentially methylated loci in B cells. Hum. Mol. Genet. *26*, 2803–2811.

Kafri, T., Ariel, M., Brandeis, M., Shemer, R., Urven, L., McCarrey, J., Cedar, H., and Razin, A. (1992). Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. Genes Dev. *6*, 705–714.

Kagiwada, S., Kurimoto, K., Hirota, T., Yamaji, M., and Saitou, M. (2013). Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. EMBO J. *32*, 340–353.

Kaina, B., Christmann, M., Naumann, S., and Roos, W.P. (2007). MGMT: Key node in the battle against genotoxicity, carcinogenicity and apoptosis induced by alkylating agents. DNA Repair *6*, 1079–1099.

Kaminskas, E. (2005). Approval Summary: Azacitidine for Treatment of Myelodysplastic Syndrome Subtypes. Clin. Cancer Res. *11*, 3604–3608.

Kanai, Y., Ushijima, S., Nakanishi, Y., Sakamoto, M., and Hirohashi, S. (2003). Mutation of the DNA methyltransferase (DNMT) 1 gene in human colorectal cancers. Cancer Lett. *192*, 75–82.

Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004). Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature *429*, 900–903.

Kantarjian, H., Issa, J.-P.J., Rosenfeld, C.S., Bennett, J.M., Albitar, M., DiPersio, J., Klimek, V., Slack, J., Castro, C. de, Ravandi, F., et al. (2006). Decitabine improves patient outcomes in myelodysplastic syndromes. Cancer *106*, 1794–1803.

Karon, M., Sieger, L., Leimbrock, S., Finklestein, J.Z., Nesbit, M.E., and Swaney, J.J. (1973). 5-Azacytidine: a new active agent for the treatment of acute leukemia. Blood *42*, 359–365.

Kato, Y., Kaneda, M., Hata, K., Kumaki, K., Hisano, M., Kohara, Y., Okano, M., Li, E., Nozaki, M., and Sasaki, H. (2007). Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. Hum. Mol. Genet. *16*, 2272–2280.

Kawano, H., Saeki, H., Kitao, H., Tsuda, Y., Otsu, H., Ando, K., Ito, S., Egashira, A., Oki, E., Morita, M., et al. (2014). Chromosomal Instability Associated with Global DNA Hypomethylation is Associated with the Initiation and Progression of Esophageal Squamous Cell Carcinoma. Ann. Surg. Oncol. *21*, 696–702.

Kearns, N.A., Pham, H., Tabak, B., Genga, R.M., Silverstein, N.J., Garber, M., and Maehr, R. (2015). Functional annotation of native enhancers with a Cas9 -histone demethylase fusion. Nat. Methods *12*, 401–403.

Keil, K.P., Abler, L.L., Mehta, V., Altmann, H.M., Laporta, J., Plisch, E.H., Suresh, M., Hernandez, L.L., and Vezina, C.M. (2014). DNA methylation of E-cadherin is a priming mechanism for prostate development. Dev. Biol. *387*, 142–153.

Kidambi, T.D., Blanco, A., Van Ziffle, J., and Terdiman, J.P. (2016). Constitutional MLH1 methylation presenting with colonic polyposis syndrome and not Lynch syndrome. Fam. Cancer *15*, 275–280.

Kim, S., Han, J., Kim, J., and Park, C. (2004). Maspin expression is transactivated by p63 and is critical for the modulation of lung cancer progression. Cancer Res. *64*, 6900–6905.

Kishikawa, S., Murata, T., Ugai, H., Yamazaki, T., and Yokoyama, K.K. (2003). Control elements of Dnmt1 gene are regulated in cell-cycle dependent manner. Nucleic Acids Res. Suppl. 2001 307–308.

Klann, T.S., Black, J.B., Chellappan, M., Safi, A., Song, L., Hilton, I.B., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2017). CRISPR–Cas9 epigenome editing enables high-throughput screening for functional regulatory elements in the human genome. Nat. Biotechnol. *35*, 561–568.

Klein, C.J., Botuyan, M.-V., Wu, Y., Ward, C.J., Nicholson, G.A., Hammans, S., Hojo, K., Yamanishi, H., Karpf, A.R., Wallace, D.C., et al. (2011). Mutations in DNMT1 cause hereditary sensory neuropathy with dementia and hearing loss. Nat. Genet. *43*, 595–600.

Kleinstiver, B.P., Prew, M.S., Tsai, S.Q., Topkar, V., Nguyen, N.T., Zheng, Z., Gonzales, A.P.W., Li, Z., Peterson, R.T., Yeh, J.-R.J., et al. (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature *523*, 481–485.

Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N., Zheng, Z., and Joung, J.K. (2016). High-fidelity CRISPR-Cas9 variants with undetectable genome-wide off-targets. Nature *529*, 490–495.

Klisovic, R.B., Stock, W., Cataland, S., Klisovic, M.I., Liu, S., Blum, W., Green, M., Odenike, O., Godley, L., Burgt, J.V., et al. (2008). A Phase I Biological Study of MG98, an Oligodeoxynucleotide Antisense to DNA Methyltransferase 1, in Patients with High-Risk Myelodysplasia and Acute Myeloid Leukemia. Clin. Cancer Res. *14*, 2444–2449.

Klose, R.J., and Bird, A.P. (2006). Genomic DNA methylation: the mark and its mediators. Trends Biochem. Sci. *31*, 89–97.

Klug, A. (2010). The discovery of zinc fingers and their development for practical applications in gene regulation and genome manipulation. Q. Rev. Biophys. *43*, 1–21.

Ko, M., Huang, Y., Jankowska, A.M., Pape, U.J., Tahiliani, M., Bandukwala, H.S., An, J., Lamperti, E.D., Koh, K.P., Ganetzky, R., et al. (2010). Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. Nature *468*, 839–843.

Ko, M., An, J., Bandukwala, H.S., Chavez, L., Äijö, T., Pastor, W.A., Segal, M.F., Li, H., Koh, K.P., Lähdesmäki, H., et al. (2013). Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. Nature *497*, 122–126.

Kobayashi, Y., Absher, D.M., Gulzar, Z.G., Young, S.R., McKenney, J.K., Peehl, D.M., Brooks, J.D., Myers, R.M., and Sherlock, G. (2011). DNA methylation profiling reveals novel biomarkers and important roles for DNA methyltransferases in prostate cancer. Genome Res. *21*, 1017–1027.

Kohli, R.M., and Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. Nature *502*, 472–479.

Konermann, S., Brigham, M.D., Trevino, A.E., Hsu, P.D., Heidenreich, M., Le Cong, Platt, R.J., Scott, D.A., Church, G.M., and Zhang, F. (2013). Optical control of mammalian endogenous transcription and epigenetic states. Nature *500*, 472–476.

Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., et al. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature *517*, 583–588.

Kong, H.K., Yoon, S., and Park, J.H. (2012a). The Regulatory Mechanism of the LY6K Gene Expression in Human Breast Cancer Cells. J. Biol. Chem. *287*, 38889–38900.

Kong, K.L., Kwong, D.L.W., Chan, T.H.-M., Law, S.Y.-K., Chen, L., Li, Y., Qin, Y.-R., and Guan, X.-Y. (2012b). MicroRNA-375 inhibits tumour growth and metastasis in oesophageal squamous cell carcinoma through repressing insulin-like growth factor 1 receptor. Gut *61*, 33–42.

Kosicki, M., Tomberg, K., and Bradley, A. (2018). Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. Nat. Biotechnol. *36*, 765–771.

Kosmider, O., Gelsi-Boyer, V., Ciudad, M., Racoeur, C., Jooste, V., Vey, N., Quesnel, B., Fenaux, P., Bastie, J.-N., Beyne-Rauzy, O., et al. (2009). TET2 gene mutation is a frequent and adverse event in chronic myelomonocytic leukemia. Haematologica *94*, 1676–1681.

Kosmider, O., Gelsi-Boyer, V., Slama, L., Dreyfus, F., Beyne-Rauzy, O., Quesnel, B., Hunault-Berger, M., Slama, B., Vey, N., Lacombe, C., et al. (2010). Mutations of IDH1 and IDH2 genes in early and accelerated phases of myelodysplastic syndromes and MDS/myeloproliferative neoplasms. Leukemia *24*, 1094–1096.

Krajnović, M., Radojković, M., Davidović, R., Dimitrijević, B., and Krtolica, K. (2013). Prognostic significance of epigenetic inactivation of p16, p15, MGMT and DAPK genes in follicular lymphoma. Med. Oncol. Northwood Lond. Engl. *30*, 441.

Krebs, A.R., Dessus-Babus, S., Burger, L., and Schübeler, D. (2014). High-throughput engineering of a mammalian genome reveals building principles of methylation states at CG rich regions. ELife *3*, e04094.

Kriaucionis, S., and Heintz, N. (2009). The nuclear DNA base, 5-hydroxymethylcytosine is present in brain and enriched in Purkinje neurons. Science *324*, 929–930.

Kuck, D., Caulfield, T., Lyko, F., and Medina-Franco, J.L. (2010). Nanaomycin A Selectively Inhibits DNMT3B and Reactivates Silenced Tumor Suppressor Genes in Human Cancer Cells. Mol. Cancer Ther. *9*, 3015–3023.

Kulis, M., Heath, S., Bibikova, M., Queirós, A.C., Navarro, A., Clot, G., Martínez-Trillos, A., Castellano, G., Brun-Heath, I., Pinyol, M., et al. (2012). Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. Nat. Genet. *44*, 1236–1242.

Kungulovski, G., and Jeltsch, A. (2016). Epigenome Editing: State of the Art, Concepts, and Perspectives. Trends Genet. *32*, 101–113.

Kungulovski, G., Nunna, S., Thomas, M., Zanger, U.M., Reinhardt, R., and Jeltsch, A. (2015). Targeted epigenome editing of an endogenous locus with chromatin modifiers is not stably maintained. Epigenetics Chromatin *8*, 12.

Kuramochi-Miyagawa, S., Watanabe, T., Gotoh, K., Totoki, Y., Toyoda, A., Ikawa, M., Asada, N., Kojima, K., Yamaguchi, Y., Ijiri, T.W., et al. (2008). DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. Genes Dev. *22*, 908–917.

Kuscu, C., Arslan, S., Singh, R., Thorpe, J., and Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. Nat. Biotechnol. *32*, 677–683.

Lamb, Y.N., and Dhillon, S. (2017). Epi proColon<Superscript>®</Superscript> 2.0 CE: A Blood-Based Screening Test for Colorectal Cancer. Mol. Diagn. Ther. *21*, 225–232.

de Lange, O., Wolf, C., Dietze, J., Elsaesser, J., Morbitzer, R., and Lahaye, T. (2014). Programmable DNA-binding proteins from Burkholderia provide a fresh perspective on the TALE-like repeat domain. Nucleic Acids Res. *42*, 7436–7449.

Larijani, M., Frieder, D., Sonbuchner, T., Bransteitter, R., Goodman, M., Bouhassira, E., Scharff, M., and Martin, A. (2005). Methylation protects cytidines from AID-mediated deamination. Mol. Immunol. *42*, 599–604.

Latha, K., Zhang, W., Cella, N., Shi, H.Y., and Zhang, M. (2005). Maspin Mediates Increased Tumor Cell Apoptosis upon Induction of the Mitochondrial Permeability Transition. Mol. Cell. Biol. *25*, 1737–1748.

Lauster, R., Trautner, T.A., and Noyer-Weidner, M. (1989). Cytosine-specific type II DNA methyltransferases. J. Mol. Biol. *206*, 305–312.

Lee, B.H., Yegnasubramanian, S., Lin, X., and Nelson, W.G. (2005). Procainamide Is a Specific Inhibitor of DNA Methyltransferase 1. J. Biol. Chem. *280*, 40749–40756.

Lee, E., Iskow, R., Yang, L., Gokcumen, O., Haseley, P., Luquette, L.J., Lohr, J.G., Harris, C.C., Ding, L., Wilson, R.K., et al. (2012a). Landscape of Somatic Retrotransposition in Human Cancers. Science *337*, 967–971.

Lee, H.-Z., Kwitkowski, V.E., Del Valle, P.L., Ricci, M.S., Saber, H., Habtemariam, B.A., Bullock, J., Bloomquist, E., Li Shen, Y., Chen, X.-H., et al. (2015). FDA Approval: Belinostat for the Treatment of Patients with Relapsed or Refractory Peripheral T-cell Lymphoma. Clin. Cancer Res. *21*, 2666–2670.

Lee, S.-J., Jang, H., and Park, C. (2012b). Maspin increases Ku70 acetylation and Bax-mediated cell death in cancer cells. Int. J. Mol. Med. *29*, 225–230.

Lee, S.M., Park, J.Y., and Kim, D.S. (2012c). Methylation of TMEFF2 Gene in Tissue and Serum DNA from Patients with Non-Small Cell Lung Cancer. Mol. Cells *34*, 171–176.

Legendre, C., Gooden, G.C., Johnson, K., Martinez, R.A., Liang, W.S., and Salhia, B. (2015). Whole-genome bisulfite sequencing of cell-free DNA identifies signature associated with metastatic breast cancer. Clin. Epigenetics *7*, 100.

Lei, H., Oh, S.P., Okano, M., Juttermann, R., Goss, K.A., Jaenisch, R., and Li, E. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. Development *122*, 3195–3205.

Lemaire, M., Momparler, L.F., Raynal, N.J.-M., Bernstein, M.L., and Momparler, R.L. (2009). Inhibition of cytidine deaminase by zebularine enhances the antineoplastic action of 5-aza-2′-deoxycytidine. Cancer Chemother. Pharmacol. *63*, 411–416.

Leonhardt, H., Page, A.W., Weier, H.-U., and Bestor, T.H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell *71*, 865–873.

Leung, D., Du, T., Wagner, U., Xie, W., Lee, A.Y., Goyal, P., Li, Y., Szulwach, K.E., Jin, P., Lorincz, M.C., et al. (2014). Regulation of DNA methylation turnover at LTR retrotransposons and imprinted loci by the histone methyltransferase Setdb1. Proc. Natl. Acad. Sci. *111*, 6690–6695.

Lev Maor, G., Yearim, A., and Ast, G. (2015). The alternative role of DNA methylation in splicing regulation. Trends Genet. *31*, 274–280.

Levine, J.J., Stimson-Crider, K.M., and Vertino, P.M. (2003). Effects of methylation on expression of TMS1/ASC in human breast cancer cells. Oncogene *22*, 3475–3488.

Ley, T.J., Ding, L., Walter, M.J., McLellan, M.D., Lamprecht, T., Larson, D.E., Kandoth, C., Payton, J.E., Baty, J., Welch, J., et al. (2010). DNMT3A Mutations in Acute Myeloid Leukemia. N. Engl. J. Med. *363*, 2424–2433.

Li, B.-Z., Huang, Z., Cui, Q.-Y., Song, X.-H., Du, L., Jeltsch, A., Chen, P., Li, G., Li, E., and Xu, G.-L. (2011a). Histone tails regulate DNA methylation by allosterically activating de novo methyltransferase. Cell Res. *21*, 1172–1181.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915–926.

Li, E., Beard, C., and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. Nature *366*, 362–365.

Li, J., Huang, Q., Zeng, F., Li, W., He, Z., Chen, W., Zhu, W., and Zhang, B. (2014). The Prognostic Value of Global DNA Hypomethylation in Cancer: A Meta-Analysis. PLoS ONE *9*, e106290.

Li, J., Hao, D., Wang, L., Wang, H., Wang, Y., Zhao, Z., Li, P., Deng, C., and Di, L. (2017). Epigenetic targeting drugs potentiate chemotherapeutic effects in solid tumor therapy. Sci. Rep. *7*, 4035.

Li, J.-Y., Lees-Murdock, D.J., Xu, G.-L., and Walsh, C.P. (2004). Timing of establishment of paternal methylation imprints in the mouse. Genomics *84*, 952–960.

Li, K., Pang, J., Cheng, H., Liu, W.-P., Di, J.-M., Xiao, H.-J., Luo, Y., Zhang, H., Huang, W.-T., Chen, M.-K., et al. (2015a). Manipulation of prostate cancer metastasis by locus-specific modification of the CRMP4 promoter region using chimeric TALE DNA methyltransferase and demethylase. Oncotarget *6*, 10030–10044.

Li, L., Ying, J., Li, H., Zhang, Y., Shu, X., Fan, Y., Tan, J., Cao, Y., Tsao, S.W., Srivastava, G., et al. (2012). The human cadherin 11 is a pro-apoptotic tumor suppressor modulating cell stemness through Wnt/β-catenin signaling and silenced in common carcinomas. Oncogene *31*, 3901–3912.

Li, L., Atef, A., Piatek, A., Ali, Z., Piatek, M., Aouida, M., Sharakuu, A., Mahjoub, A., Wang, G., Khan, S., et al. (2013a). Characterization and DNA-Binding Specificities of Ralstonia TAL-Like Effectors. Mol. Plant *6*, 1318–1330.

Li, L., Li, C., Mao, H., Du, Z., Chan, W.Y., Murray, P., Luo, B., Chan, A.T., Mok, T.S., Chan, F.K., et al. (2016). Epigenetic inactivation of the CpG demethylase TET1 as a DNA methylation feedback loop in human cancers. Sci. Rep. *6*, 26591.

Li, N., Ye, M., Li, Y., Yan, Z., Butcher, L.M., Sun, J., Han, X., Chen, Q., zhang, X., and Wang, J. (2010). Whole genome DNA methylation analysis based on high throughput sequencing technology. Methods *52*, 203–212.

Li, S., Garrett-Bakelman, F.E., Akalin, A., Zumbo, P., Levine, R., To, B.L., Lewis, I.D., Brown, A.L., D'Andrea, R.J., Melnick, A., et al. (2013b). An optimized algorithm for detecting and annotating regional differential methylation. BMC Bioinformatics *14*, S10.

Li, T., Wang, L., Du, Y., Xie, S., Yang, X., Lian, F., Zhou, Z., and Qian, C. (2018). Structural and mechanistic insights into UHRF1-mediated DNMT1 activation in the maintenance DNA methylation. Nucleic Acids Res. *46*, 3218–3231.

Li, W., Zhang, H., Guo, Q., Wu, X., Xu, Z., Dang, C., Xia, P., and Song, Y. (2015b). Detection of SNCA and FBN1 Methylation in the Stool as a Biomarker for Colorectal Cancer. Dis. Markers *2015*, 657570.

Li, X., Ito, M., Zhou, F., Youngson, N., Zuo, X., Leder, P., and Ferguson-Smith, A.C. (2008). A maternal-zygotic effect gene Zfp57 maintains both maternal and paternal imprints. Dev. Cell *15*, 547–557.

Li, Z., Cai, X., Cai, C.-L., Wang, J., Zhang, W., Petersen, B.E., Yang, F.-C., and Xu, M. (2011b). Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. Blood *118*, 4509–4518.

Liang, G., Chan, M.F., Tomigahara, Y., Tsai, Y.C., Gonzales, F.A., Li, E., Laird, P.W., and Jones, P.A. (2002). Cooperativity between DNA Methyltransferases in the Maintenance Methylation of Repetitive Elements. Mol. Cell. Biol. *22*, 480–491.

Liang, T.-J., Wang, H.-X., Zheng, Y.-Y., Cao, Y.-Q., Wu, X., Zhou, X., and Dong, S.-X. (2017). APC hypermethylation for early diagnosis of colorectal cancer: a meta-analysis and literature review. Oncotarget *8*, 46468–46479.

Lienert, F., Wirbelauer, C., Som, I., Dean, A., Mohn, F., and Schübeler, D. (2011). Identification of genetic elements that autonomously determine DNA methylation states. Nat. Genet. *43*, 1091–1097.

Lin, I.G., Han, L., Taghva, A., O'Brien, L.E., and Hsieh, C.-L. (2002). Murine De Novo Methyltransferase Dnmt3a Demonstrates Strand Asymmetry and Site Preference in the Methylation of DNA In Vitro. Mol. Cell. Biol. *22*, 704–723.

Lin, L., Liu, Y., Xu, F., Huang, J., Daugaard, T.F., Petersen, T.S., Hansen, B., Ye, L., Zhou, Q., Fang, F., et al. (2018). Genome-wide determination of on-target and off-target characteristics for RNA-guided DNA methylation by dCas9 methyltransferases. GigaScience *7*, 1–19.

LIN, Y.-L., GUI, S.-L., and MA, J.-G. (2015). Aberrant methylation of CDH11 predicts a poor outcome for patients with bladder cancer. Oncol. Lett. *10*, 647–652.

Linhart, H.G., Lin, H., Yamada, Y., Moran, E., Steine, E.J., Gokhale, S., Lo, G., Cantu, E., Ehrich, M., He, T., et al. (2007). Dnmt3b promotes tumorigenesis in vivo by gene-specific de novo methylation and transcriptional silencing. Genes Dev. *21*, 3110–3122.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.-M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315–322.

Lister, R., Mukamel, E.A., Nery, J.R., Urich, M., Puddifoot, C.A., Johnson, N.D., Lucero, J., Huang, Y., Dwork, A.J., Schultz, M.D., et al. (2013). Global Epigenomic Reconfiguration During Mammalian Brain Development. Science *341*, 1237905.

Liu, J., Nau, M.M., Zucman-Rossi, J., Powell, J.I., Allegra, C.J., and Wright, J.J. (1997). LINE-I element insertion at the t(11;22) translocation breakpoint of a desmoplastic small round cell tumor. Genes. Chromosomes Cancer *18*, 232–239.

Liu, L., Sun, L., Li, C., Li, X., Zhang, Y., Yu, Y., and Xia, W. (2015). Quantitative detection of methylation of FHIT and BRCA1 promoters in the serum of ductal breast cancer patients. Biomed. Mater. Eng. *26*, S2217–S2222.

Liu, P.-Q., Rebar, E.J., Zhang, L., Liu, Q., Jamieson, A.C., Liang, Y., Qi, H., Li, P.-X., Chen, B., Mendel, M.C., et al. (2001). Regulation of an Endogenous Locus Using a Panel of Designed Zinc Finger Proteins Targeted to Accessible Chromatin Regions: ACTIVATION OF VASCULAR ENDOTHELIAL GROWTH FACTOR A. J. Biol. Chem. *276*, 11323–11334.

Liu, S.-L., Sui, Y.-F., and Lin, M.-Z. (2016a). MiR-375 is epigenetically downregulated due to promoter methylation and modulates multi-drug resistance in breast cancer cells via targeting YBX1. Eur. Rev. Med. Pharmacol. Sci. *20*, 3223–3229.

Liu, X.S., Wu, H., Ji, X., Stelzer, Y., Wu, X., Czauderna, S., Shu, J., Dadon, D., Young, R.A., and Jaenisch, R. (2016b). Editing DNA methylation in the mammalian genome. Cell *167*, 233-247.e17.

Liu, X.S., Wu, H., Krzisch, M., Wu, X., Graef, J., Muffat, J., Hnisz, D., Li, C.H., Yuan, B., Xu, C., et al. (2018). Rescue of Fragile X Syndrome Neurons by DNA Methylation Editing of the FMR1 Gene. Cell *172*, 979-992.e6.

Liu, Y., Aryee, M.J., Padyukov, L., Fallin, M.D., Hesselberg, E., Runarsson, A., Reinius, L., Acevedo, N., Taub, M., Ronninger, M., et al. (2013). Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in Rheumatoid Arthritis. Nat. Biotechnol. *31*, 142–147.

Liu, Z., Xie, Z., Jones, W., Pavlovicz, R.E., Liu, S., Yu, J., Li, P., Lin, J., Fuchs, J.R., Marcucci, G., et al. (2009). Curcumin is a potent DNA hypomethylation agent. Bioorg. Med. Chem. Lett. *19*, 706–709.

Liutkevičiūtė, Z., Kriukienė, E., Ličytė, J., Rudytė, M., Urbanavičiūtė, G., and Klimašauskas, S. (2014). Direct Decarboxylation of 5-Carboxylcytosine by DNA C5- Methyltransferases. J. Am. Chem. Soc. *136*, 5884–5887.

Llinàs-Arias, P., and Esteller, M. (2017). Epigenetic inactivation of tumour suppressor coding and non-coding genes in human cancer: an update. Open Biol. *7*, 170152.

Loenarz, C., and Schofield, C.J. (2008). Expanding chemical biology of 2-oxoglutarate oxygenases. Nat. Chem. Biol. *4*, 152–156.

Lofton-Day, C., Model, F., DeVos, T., Tetzner, R., Distler, J., Schuster, M., Song, X., Lesche, R., Liebenberg, V., Ebert, M., et al. (2008). DNA Methylation Biomarkers for Blood-Based Colorectal Cancer Screening. Clin. Chem. *54*, 414–423.

Long, C., Amoasii, L., Mireault, A.A., McAnally, J.R., Li, H., Sanchez-Ortiz, E., Bhattacharyya, S., Shelton, J.M., Bassel-Duby, R., and Olson, E.N. (2016a). Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. Science *351*, 400–403.

Long, H.K., King, H.W., Patient, R.K., Odom, D.T., and Klose, R.J. (2016b). Protection of CpG islands from DNA methylation is DNA-encoded and evolutionarily conserved. Nucleic Acids Res. *44*, 6693–6706.

Longerich, S., Basu, U., Alt, F., and Storb, U. (2006). AID in somatic hypermutation and class switch recombination. Curr. Opin. Immunol. *18*, 164–174.

Lorincz, M.C., Dickerson, D.R., Schmitt, M., and Groudine, M. (2004). Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. Nat. Struct. Mol. Biol. *11*, 1068–1075.

Lorsbach, R.B., Moore, J., Mathew, S., Raimondi, S.C., Mukatira, S.T., and Downing, J.R. (2003). TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). Leukemia *17*, 637–641.

Losman, J.-A., and Kaelin, W.G. (2013). What a difference a hydroxyl makes: mutant IDH, (R)-2-hydroxyglutarate, and cancer. Genes Dev. *27*, 836–852.

Lu, C., Ward, P.S., Kapoor, G.S., Rohle, D., Turcan, S., Abdel-Wahab, O., Edwards, C.R., Khanin, R., Figueroa, M.E., Melnick, A., et al. (2012). IDH mutation impairs histone demethylation and results in a block to cell differentiation. Nature *483*, 474–478.

Lu, R., Wang, P., Parton, T., Zhou, Y., Chrysovergis, K., Rockowitz, S., Chen, W.-Y., Abdel-Wahab, O., Wade, P.A., Zheng, D., et al. (2016). Epigenetic Perturbations by Arg882-Mutated DNMT3A Potentiate Aberrant Stem Cell Gene-Expression Program and Acute Leukemia Development. Cancer Cell *30*, 92–107.

Lucifero, D., Mann, M.R.W., Bartolomei, M.S., and Trasler, J.M. (2004). Gene-specific timing and epigenetic memory in oocyte imprinting. Hum. Mol. Genet. *13*, 839–849.

Lui, E.L.H., Loo, W.T.Y., Zhu, L., Cheung, M.N.B., and Chow, L.W.C. (2005). DNA hypermethylation of TIMP3 gene in invasive breast ductal carcinoma. Biomed. Pharmacother. *59*, S363–S365.

Lujambio, A., Ropero, S., Ballestar, E., Fraga, M.F., Cerrato, C., Setien, F., Casado, S., Suarez-Gauthier, A., Sanchez-Cespedes, M., Gitt, A., et al. (2007). Genetic Unmasking of an Epigenetically Silenced microRNA in Human Cancer Cells. Cancer Res. *67*, 1424–1429.

Lyon, M.F. (1961). Gene Action in the X-chromosome of the Mouse (Mus musculus L.). Nature *190*, 372–373.

Ma, A.-N., Wang, H., Guo, R., Wang, Y.-X., Li, W., Cui, J., Wang, G., Hoffman, A.R., and Hu, J.-F. (2014). Targeted gene suppression by inducing de novo DNA methylation in the gene promoter. Epigenetics Chromatin *7*, 20.

Ma, H., Tu, L.-C., Naseri, A., Huisman, M., Zhang, S., Grunwald, D., and Pederson, T. (2016). Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. Nat. Biotechnol. *34*, 528–530.

Ma, H., Marti-Gutierrez, N., Park, S.-W., Wu, J., Lee, Y., Suzuki, K., Koski, A., Ji, D., Hayama, T., Ahmed, R., et al. (2017). Correction of a pathogenic gene mutation in human embryos. Nature *548*, 413–419.

Maass, N., Biallek, M., Rösel, F., Schem, C., Ohike, N., Zhang, M., Jonat, W., and Nagasaki, K. (2002). Hypermethylation and histone deacetylation lead to silencing of the maspin gene in human breast cancer. Biochem. Biophys. Res. Commun. *297*, 125–128.

Machado, J., Oliveira, C., Carvalho, R., Soares, P., Berx, G., Caldas, C., Seruca, R., Carneiro, F., and es, M.S.-S. (2001). E-cadherin gene (CDH1) promoter methylation as the second hit in sporadic di□use gastric carcinoma. Oncogene *20*, 1525–1528.

Macleod, D., Charlton, J., Mullins, J., and Bird, A.P. (1994). Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island. Genes Dev. *8*, 2282–2292.

Maeder, M.L., Linder, S.J., Cascio, V.M., Fu, Y., Ho, Q.H., and Joung, J.K. (2013a). CRISPR RNA–guided activation of endogenous human genes. Nat. Methods *10*, 977–979.

Maeder, M.L., Angstman, J.F., Richardson, M.E., Linder, S.J., Cascio, V.M., Tsai, S.Q., Ho, Q.H., Sander, J.D., Reyon, D., Bernstein, B.E., et al. (2013b). Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. Nat. Biotechnol. *31*, 1137–1142.

Maeder, M.L., Linder, S.J., Reyon, D., Angstman, J.F., Fu, Y., Sander, J.D., and Joung, J.K. (2013c). Robust, synergistic regulation of human gene expression using TALE activators. Nat. Methods *10*, 243–245.

Mahon, K.L., Qu, W., Devaney, J., Paul, C., Castillo, L., Wykes, R.J., Chatfield, M.D., Boyer, M.J., Stockler, M.R., Marx, G., et al. (2014). Methylated Glutathione S-transferase 1 (mGSTP1) is a potential plasma free DNA epigenetic marker of prognosis and response to chemotherapy in castrate-resistant prostate cancer. Br. J. Cancer *111*, 1802–1809.

Maiti, A., and Drohat, A.C. (2011). Thymine DNA Glycosylase Can Rapidly Excise 5-Formylcytosine and 5-Carboxylcytosine. J. Biol. Chem. *286*, 35334–35338.

Mak, A.N.-S., Bradley, P., Cernadas, R.A., Bogdanove, A.J., and Stoddard, B.L. (2012). The crystal structure of TAL effector PthXo1 bound to its DNA target. Science *335*, 716–719.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013a). RNA-Guided Human Genome Engineering via Cas9. Science *339*, 823–826.

Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., and Church, G.M. (2013b). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat. Biotechnol. *31*, 833–838.

Mann, B.S., Johnson, J.R., Cohen, M.H., Justice, R., and Pazdur, R. (2007). FDA Approval Summary: Vorinostat for Treatment of Advanced Primary Cutaneous T-Cell Lymphoma. The Oncologist *12*, 1247–1252.

Mardis, E.R., Ding, L., Dooling, D.J., Larson, D.E., McLellan, M.D., Chen, K., Koboldt, D.C., Fulton, R.S., Delehaunty, K.D., McGrath, S.D., et al. (2009). Recurring Mutations Found by Sequencing an Acute Myeloid Leukemia Genome. N. Engl. J. Med. *361*, 1058–1066.

Marzo, A.M.D., Marchi, V.L., Yang, E.S., Veeraswamy, R., Lin, X., and Nelson, W.G. (1999). Abnormal Regulation of DNA Methyltransferase Expression during Colorectal Carcinogenesis. Cancer Res. *59*, 3855–3860.

Mascarenhas, J., and Hoffman, R. (2012). Ruxolitinib: The First FDA Approved Therapy for the Treatment of Myelofibrosis. Clin. Cancer Res. *18*, 3008–3014.

Mastroeni, D., McKee, A., Grover, A., Rogers, J., and Coleman, P.D. (2009). Epigenetic Differences in Cortical Neurons from a Pair of Monozygotic Twins Discordant for Alzheimer's Disease. PLoS ONE *4*, e6617.

MASUDA, K., BANNO, K., YANOKURA, M., TSUJI, K., KOBAYASHI, Y., KISU, I., UEKI, A., YAMAGAMI, W., NOMURA, H., TOMINAGA, E., et al. (2012). Association of epigenetic inactivation of the WRN gene with anticancer drug sensitivity in cervical cancer cells. Oncol. Rep. *28*, 1146–1152.

Mathe, A., Wong-Brown, M., Locke, W.J., Stirzaker, C., Braye, S.G., Forbes, J.F., Clark, S.J., Avery-Kiejda, K.A., and Scott, R.J. (2016). DNA methylation profile of triple negative breast cancer-specific genes comparing lymph node positive patients to lymph node negative patients. Sci. Rep. *6*, 33435.

Mazzone, R., Zwergel, C., Mai, A., and Valente, S. (2017). Epi-drugs in combination with immunotherapy: a new avenue to improve anticancer efficacy. Clin. Epigenetics *9*, 59.

McDonald, J.I., Celik, H., Rois, L.E., Fishberger, G., Fowler, T., Rees, R., Kramer, A., Martens, A., Edwards, J.R., and Challen, G.A. (2016). Reprogrammable CRISPR/Cas9-based system for inducing site-specific DNA methylation. Biol. Open *5*, 866–874.

McGhee, J.D., and Ginder, G.D. (1979). Specific DNA methylation sites in the vicinity of the chicken beta-globin genes. Nature *280*, 419–420.

Meissner, A., Gnirke, A., Bell, G.W., Ramsahoye, B., Lander, E.S., and Jaenisch, R. (2005). Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. Nucleic Acids Res. *33*, 5868–5877.

Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., et al. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature *454*, 766–770.

Mendenhall, E.M., Williamson, K.E., Reyon, D., Zou, J.Y., Ram, O., Joung, J.K., and Bernstein, B.E. (2013). Locus-specific editing of histone modifications at endogenous enhancers. Nat. Biotechnol. *31*, 1133–1136.

Messerschmidt, D.M., Knowles, B.B., and Solter, D. (2014). DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. Genes Dev. *28*, 812–828.

Miglio, U., Mezzapelle, R., Paganotti, A., Veggiani, C., Mercalli, F., Mancuso, G., Gaudino, E., Rena, O., Buosi, R., and Boldorini, R. (2015). Frequency of O6-methylguanine-DNA methyltransferase promoter methylation in cytological samples from small cell lung cancer. Diagn. Cytopathol. *43*, 947–952.

Miki, Y., Nishisho, I., Horii, A., Miyoshi, Y., Utsunomiya, J., Kinzler, K.W., Vogelstein, B., and Nakamura, Y. (1992). Disruption of the APC Gene by a Retrotransposal Insertion of LI Sequence in a Colon Cancer. Cancer Res. *52*, 643–645.

Milani, L., Lundmark, A., Kiialainen, A., Nordlund, J., Flaegstad, T., Forestier, E., Heyman, M., Jonmundsson, G., Kanerva, J., Schmiegelow, K., et al. (2010). DNA methylation for subtype classification and prediction of treatment outcome in patients with childhood acute lymphoblastic leukemia. Blood *115*, 1214–1225.

Miller, J., McLachlan, A.D., and Klug, A. (1985). Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. EMBO J. *4*, 1609–1614.

Minor, E.A., Court, B.L., Young, J.I., and Wang, G. (2013). Ascorbate Induces Ten-Eleven Translocation (Tet) Methylcytosine Dioxygenase-mediated Generation of 5-Hydroxymethylcytosine. J. Biol. Chem. *288*, 13669–13674.

Mizuguchi, Y., Saiki, Y., Horii, A., and Fukushige, S. (2016). Targeted TET oxidase activity through methyl-CpG-binding domain extensively suppresses cancer cell proliferation. Cancer Med. *5*, 2522–2533.

Mizuno, S. -i. (2001). Expression of DNA methyltransferases DNMT1, 3A, and 3B in normal hematopoiesis and in acute and chronic myelogenous leukemia. Blood *97*, 1172–1179.

Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., and Soria, E. (2005). Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. J. Mol. Evol. *60*, 174–182.

Momparler, R.L., Rivard, G.E., and Gyger, M. (1985). Clinical trial on 5-AZA-2'-deoxycytidine in patients with acute leukemia. Pharmacol. Ther. *30*, 277–286.

MONK, M., BOUBELIK, M., and LEHNERT, S. (1987). Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. Development *99*, 371–382.

Moran-Crusio, K., Reavie, L., Shih, A., Abdel-Wahab, O., Ndiaye-Lobry, D., Lobry, C., Figueroa, M.E., Vasanthakumar, A., Patel, J., Zhao, X., et al. (2011). Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. Cancer Cell *20*, 11–24.

Morgan, H.D., Dean, W., Coker, H.A., Reik, W., and Petersen-Mahrt, S.K. (2004). Activation-induced Cytidine Deaminase Deaminates 5-Methylcytosine in DNA and Is Expressed in Pluripotent Tissues: IMPLICATIONS FOR EPIGENETIC REPROGRAMMING. J. Biol. Chem. *279*, 52353–52360.

Morgan, S.L., Mariano, N.C., Bermudez, A., Arruda, N.L., Wu, F., Luo, Y., Shankar, G., Jia, L., Chen, H., Hu, J.-F., et al. (2017). Manipulation of nuclear architecture through CRISPR-mediated chromosomal looping. Nat. Commun. *8*, 15993.

Morita, S., Noguchi, H., Horii, T., Nakabayashi, K., Kimura, M., Okamura, K., Sakai, A., Nakashima, H., Hata, K., Nakashima, K., et al. (2016). Targeted DNA demethylation in vivo using dCas9-peptide repeat and scFv-TET1 catalytic domain fusions. Nat. Biotechnol. *34*, 1060–1065.

Morselli, M., Pastor, W.A., Montanini, B., Nee, K., Ferrari, R., Fu, K., Bonora, G., Rubbi, L., Clark, A.T., Ottonello, S., et al. (2015). In vivo targeting of de novo DNA methylation by histone modifications in yeast and mouse. ELife *4*, e06205.

Muller, P.A.J., and Vousden, K.H. (2014). Mutant p53 in Cancer: New Functions and Therapeutic Opportunities. Cancer Cell *25*, 304–317.

Muramatsu, M., Sankaranand, V.S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N.O., and Honjo, T. (1999). Specific Expression of Activation-induced Cytidine Deaminase (AID), a Novel Member of the RNA-editing Deaminase Family in Germinal Center B Cells. J. Biol. Chem. *274*, 18470–18476.

Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., and Honjo, T. (2000). Class Switch Recombination and Hypermutation Require Activation-Induced Cytidine Deaminase (AID), a Potential RNA Editing Enzyme. Cell *102*, 553–563.

Mussolino, C., Morbitzer, R., Lütge, F., Dannemann, N., Lahaye, T., and Cathomen, T. (2011). A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. Nucleic Acids Res. *39*, 9283–9293.

Nabel, C.S., Jia, H., Ye, Y., Shen, L., Goldschmidt, H.L., Stivers, J.T., Zhang, Y., and Kohli, R.M. (2012). AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. Nat. Chem. Biol. *8*, 751–758.

Nagae, G., Isagawa, T., Shiraki, N., Fujita, T., Yamamoto, S., Tsutsumi, S., Nonaka, A., Yoshiba, S., Matsusaka, K., Midorikawa, Y., et al. (2011). Tissue-specific demethylation in CpG-poor promoters during cellular differentiation. Hum. Mol. Genet. *20*, 2710–2721.

Nakagawa, T., Lv, L., Nakagawa, M., Yu, Y., Yu, C., D'Alessio, A.C., Nakayama, K., Fan, H.-Y., Chen, X., and Xiong, Y. (2015). CRL4VprBP E3 Ligase Promotes Monoubiquitylation and Chromatin Binding of TET Dioxygenases. Mol. Cell *57*, 247–260.

Nakamura, M., Yonekawa, Y., Kleihues, P., and Ohgaki, H. (2001). Promoter hypermethylation of the RB1 gene in glioblastomas. Lab. Investig. J. Tech. Methods Pathol. *81*, 77–82.

Nakamura, T., Arai, Y., Umehara, H., Masuhara, M., Kimura, T., Taniguchi, H., Sekimoto, T., Ikawa, M., Yoneda, Y., Okabe, M., et al. (2007). PGC7/Stella protects against DNA demethylation in early embryogenesis. Nat. Cell Biol. *9*, 64–71.

Nakamura, T., Liu, Y.-J., Nakashima, H., Umehara, H., Inoue, K., Matoba, S., Tachibana, M., Ogura, A., Shinkai, Y., and Nakano, T. (2012). PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. Nature *486*, 415–419.

Nelson, C.E., Hakim, C.H., Ousterout, D.G., Thakore, P.I., Moreb, E.A., Rivera, R.M.C., Madhavan, S., Pan, X., Ran, F.A., Yan, W.X., et al. (2016). In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. Science *351*, 403–407.

Neri, F., Rapelli, S., Krepelova, A., Incarnato, D., Parlato, C., Basile, G., Maldotti, M., Anselmi, F., and Oliviero, S. (2017). Intragenic DNA methylation prevents spurious transcription initiation. Nature *543*, 72–77.

Neumann, M., Heesch, S., Schlee, C., Schwartz, S., Gokbuget, N., Hoelzer, D., Konstandin, N.P., Ksienzyk, B., Vosberg, S., Graf, A., et al. (2013). Whole-exome sequencing in adult ETP-ALL reveals a high rate of DNMT3A mutations. Blood *121*, 4749–4752.

Newman, E.M., Morgan, R.J., Kummar, S., Beumer, J.H., Blanchard, M.S., Ruel, C., El-Khoueiry, A.B., Carroll, M.I., Hou, J.M., Li, C., et al. (2015). A phase I, pharmacokinetic, and pharmacodynamic evaluation of the DNA methyltransferase inhibitor 5-fuoro-2'-deoxycytidine, administered with tetrahydrouridine. Cancer Chemother. Pharmacol. *75*, 537–546.

Nian, J., Sun, X., Ming, S., Yan, C., Ma, Y., Feng, Y., Yang, L., Yu, M., Zhang, G., and Wang, X. (2017). Diagnostic Accuracy of Methylated SEPT9 for Blood-based Colorectal Cancer Detection: A Systematic Review and Meta-Analysis. Clin. Transl. Gastroenterol. *8*, e216–e216.

Nihongaki, Y., Kawano, F., Nakajima, T., and Sato, M. (2015). Photoactivatable CRISPR-Cas9 for optogenetic genome editing. Nat. Biotechnol. *33*, 755–760.

Nishiyama, A., Yamaguchi, L., Sharif, J., Johmura, Y., Kawamura, T., Nakanishi, K., Shimamura, S., Arita, K., Kodama, T., Ishikawa, F., et al. (2013). Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. Nature *502*, 249–253.

Noia, J.M.D., and Neuberger, M.S. (2007). Molecular Mechanisms of Antibody Somatic Hypermutation. Annu. Rev. Biochem. *76*, 1–22.

Nosho, K., Shima, K., Irahara, N., Kure, S., Baba, Y., Kirkner, G.J., Chen, L., Gokhale, S., Hazra, A., Spiegelman, D., et al. (2009). DNMT3B Expression Might Contribute to CpG Island Methylator Phenotype in Colorectal Cancer. Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. *15*, 3663–3671.

Numoto, M., Yokoro, K., and Koshi, J. (1999). ZF5, which is a Kruppel-type transcriptional repressor, requires the zinc finger domain for self-association. Biochem. Biophys. Res. Commun. *256*, 573–578.

Nunna, S., Reinhardt, R., Ragozin, S., and Jeltsch, A. (2014). Targeted Methylation of the Epithelial Cell Adhesion Molecule (EpCAM) Promoter to Silence Its Expression in Ovarian Cancer Cells. PLoS ONE *9*, e87703.

Obata, T., Yanagidani, A., Yokoro, K., Numoto, M., and Yamamoto, S. (1999). Analysis of the consensus binding sequence and the DNA-binding domain of ZF5. Biochem. Biophys. Res. Commun. *255*, 528–534.

Oberlé, I., Rousseau, F., Heitz, D., Kretz, C., Devys, D., Hanauer, A., Boué, J., Bertheas, M.F., and Mandel, J.L. (1991). Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. Science *252*, 1097–1102.

Ogawa, O., Eccles, M.R., Szeto, J., McNoe, L.A., Yun, K., Maw, M.A., Smith, P.J., and Reeve, A.E. (1993). Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms' tumour. Nature *362*, 749–751.

O'Geen, H., Henry, I.M., Bhakta, M.S., Meckler, J.F., and Segal, D.J. (2015). A genome-wide analysis of Cas9 binding specificity using ChIP-seq and targeted sequence capture. Nucleic Acids Res. *43*, 3389–3404.

O'Geen, H., Ren, C., Nicolet, C.M., Perez, A.A., Halmai, J., Le, V.M., Mackay, J.P., Farnham, P.J., and Segal, D.J. (2017). dCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression. Nucleic Acids Res. *45*, 9901–9916.

Ohno, R., Nakayama, M., Naruse, C., Okashita, N., Takano, O., Tachibana, M., Asano, M., Saitou, M., and Seki, Y. (2013). A replication-dependent passive mechanism modulates DNA demethylation in mouse primordial germ cells. Development *140*, 2892–2903.

Ohtani-Fujita, N., Dryja, T.P., Rapaport, J.M., Fujita, T., Matsumura, S., Ozasa, K., Watanabe, Y., Hayashi, K., Maeda, K., Kinoshita, S., et al. (1997). Hypermethylation in the retinoblastoma gene is associated with unilateral, sporadic retinoblastoma. Cancer Genet. Cytogenet. *98*, 43–49.

Okada, M., Kanamori, M., Someya, K., Nakatsukasa, H., and Yoshimura, A. (2017). Stabilization of Foxp3 expression by CRISPR-dCas9-based epigenome editing in mouse primary T cells. Epigenetics Chromatin *10*, 24.

Okano, M., Xie, S., and Li, E. (1998a). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nat. Genet. *19*, 219–220.

Okano, M., Xie, S., and Li, E. (1998b). Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. Nucleic Acids Res. *26*, 2536–2540.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247–257.

Ooi, S.K.T., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.-P., Allis, C.D., et al. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature *448*, 714–717.

van der Oost, J., Westra, E.R., Jackson, R.N., and Wiedenheft, B. (2014). Unravelling the structural and mechanistic basis of CRISPR–Cas systems. Nat. Rev. Microbiol. *12*, 479–492.

Oostra, B.A., and Willemsen, R. (2009). FMR1: a gene with three faces. Biochim. Biophys. Acta *1790*, 467–477.

Oshiro, M.M., Watts, G.S., Wozniak, R.J., Junk, D.J., Munoz-Rodriguez, J.L., Domann, F.E., and Futscher, B.W. (2003). Mutant p53 and aberrant cytosine methylation cooperate to silence gene expression. Oncogene *22*, 3624–3634.

O'Sullivan Coyne, G., Chen, A., Kummar, S., Collins, J.M., Meehan, R.S., Suto, M., Rubinstein, L., Kinders, R., Moore, N., Parchment, R., et al. (2016). First-in-human trial of 4'-thio-2'-deoxycytidine (TdCyd) in patients with advanced solid tumors. Ann. Oncol. *27*, vi114–vi135.

Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M., and Shirakawa, M. (2009). Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX–DNMT3–DNMT3L domain. EMBO Rep. *10*, 1235–1241.

Paroush, Z., Keshet, I., Yisraeli, J., and Cedar, H. (1990). Dynamics of demethylation and activation of the α-actin gene in myoblasts. Cell *63*, 1229–1237.

Pathania, S., Nguyen, J., Hill, S.J., Scully, R., Adelmant, G.O., Marto, J.A., Feunteun, J., and Livingston, D.M. (2011). BRCA1 is required for postreplication repair after UV-induced DNA damage. Mol. Cell *44*, 235–251.

Patra, S.K., Patra, A., and Dahiya, R. (2001). Histone Deacetylase and DNA Methyltransferase in Human Prostate Cancer. Biochem. Biophys. Res. Commun. *287*, 705–713.

Pavletich, N., and Pabo, C. (1991). Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science *252*, 809–817.

Peng, D.-F., Kanai, Y., Sawada, M., Ushijima, S., Hiraoka, N., Kitazawa, S., and Hirohashi, S. (2006). DNA methylation of multiple tumor-related genes in association with overexpression of DNA methyltransferase 1 (DNMT1) during multistage carcinogenesis of the pancreas. Carcinogenesis *27*, 1160–1168.

Penny, G.D., Kay, G.F., Sheardown, S.A., Rastan, S., and Brockdorff, N. (1996). Requirement for Xist in X chromosome inactivation. Nature *379*, 131–137.

Perez-Pinera, P., Ousterout, D.G., Brunger, J.M., Farin, A.M., Glass, K.A., Guilak, F., Crawford, G.E., Hartemink, A.J., and Gersbach, C.A. (2013a). Synergistic and tunable human gene activation by combinations of synthetic transcription factors. Nat. Methods *10*, 239–242.

Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., et al. (2013b). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat. Methods *10*, 973–976.

Pfaffeneder, T., Hackner, B., Truß, M., Münzel, M., Müller, M., Deiml, C.A., Hagemeier, C., and Carell, T. (2011). The Discovery of 5-Formylcytosine in Embryonic Stem Cell DNA. Angew. Chem. Int. Ed. *50*, 7008–7012.

Pflueger, C., Tan, D., Swain, T., Nguyen, T., Pflueger, J., Nefzger, C., Polo, J.M., Ford, E., and Lister, R. (2018). A modular dCas9-SunTag DNMT3A epigenome editing system overcomes pervasive off-target activity of direct fusion dCas9-DNMT3A constructs. Genome Res. *28*, 1193–1206.

Pickhardt, P.J. (2016). Emerging stool-based and blood-based non-invasive DNA tests for colorectal cancer screening: The importance of cancer prevention in addition to cancer detection. Abdom. Radiol. N. Y. *41*, 1441–1444.

Plath, K., Fang, J., Mlynarczyk-Evans, S.K., Cao, R., Worringer, K.A., Wang, H., Cruz, C.C. de la, Otte, A.P., Panning, B., and Zhang, Y. (2003). Role of Histone H3 Lysine 27 Methylation in X Inactivation. Science *300*, 131–135.

Ponger, L., Duret, L., and Mouchiroud, D. (2001). Determinants of CpG Islands: Expression in Early Embryo and Isochore Structure. Genome Res. *11*, 1854–1860.

Popp, C., Dean, W., Feng, S., Cokus, S.J., Andrews, S., Pellegrini, M., Jacobsen, S.E., and Reik, W. (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by Aid deficiency. Nature *463*, 1101–1105.

Pradhan, M., Estève, P.-O., Chin, H.G., Samaranayke, M., Kim, G.-D., and Pradhan, S. (2008). CXXC Domain of Human DNMT1 Is Essential for Enzymatic Activity. Biochemistry *47*, 10000–10009.

Primeau, M., Gagnon, J., and Momparler, R.L. (2003). Synergistic antineoplastic action of DNA methylation inhibitor 5-AZA-2?-deoxycytidine and histone deacetylase inhibitor depsipeptide on human breast carcinoma cells. Int. J. Cancer *103*, 177–184.

Prissette, M., El-Maarri, O., Arnaud, D., Walter, J., and Avner, P. (2001). Methylation profiles of DXPas34 during the onset of X-inactivation. Hum. Mol. Genet. *10*, 31–38.

Prokhortchouk, A. (2001). The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. Genes Dev. *15*, 1613–1618.

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. Cell *152*, 1173–1183.

Qin, T., Si, J., Raynal, N.J.-M., Wang, X., Gharibyan, V., Ahmed, S., Hu, X., Jin, C., Lu, Y., Shu, J., et al. (2015). Epigenetic synergy between decitabine and platinum derivatives. Clin. Epigenetics *7*, 97.

Quivoron, C., Couronné, L., Della Valle, V., Lopez, C.K., Plo, I., Wagner-Ballon, O., Do Cruzeiro, M., Delhommeau, F., Arnulf, B., Stern, M.-H., et al. (2011). TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. Cancer Cell *20*, 25–38.

Raedler, L.A. (2016). Farydak (Panobinostat): First HDAC Inhibitor Approved for Patients with Relapsed Multiple Myeloma. Am. Health Drug Benefits *9*, 84–87.

Rai, K., Huggins, I.J., James, S.R., Karpf, A.R., Jones, D.A., and Cairns, B.R. (2008). DNA Demethylation in Zebrafish Involves the Coupling of a Deaminase, a Glycosylase, and Gadd45. Cell *135*, 1201–1212.

Raiber, E.-A., Beraldi, D., Ficz, G., Burgess, H.E., Branco, M.R., Murat, P., Oxley, D., Booth, M.J., Reik, W., and Balasubramanian, S. (2012). Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase. Genome Biol. *13*, R69.

Rainier, S., Johnson, L.A., Dobry, C.J., Ping, A.J., Grundy, P.E., and Feinberg, A.P. (1993). Relaxation of imprinted genes in human cancer. Nature *362*, 747–749.

Ralhan, R., and Kaur, J. (2007). Alkylating agents and cancer therapy. Expert Opin. Ther. Pat. *17*, 1061–1075.

Ramiro, A.R., and Barreto, V.M. (2015). Activation-induced cytidine deaminase and active cytidine demethylation. Trends Biochem. Sci. *40*, 172–181.

Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., et al. (2015). In vivo genome editing using Staphylococcus aureus Cas9. Nature *520*, 186–191.

Randhawa, G.S., Cui, H., Barletta, J.A., Strichman-Almashanu, L.Z., Talpaz, M., Kantarjian, H., Deisseroth, A.B., Champlin, R.C., and Feinberg, A.P. (1998). Loss of Imprinting in Disease Progression in Chronic Myelogenous Leukemia. Blood *91*, 3144–3147.

Rao, J.S., Keleshian, V.L., Klein, S., and Rapoport, S.I. (2012). Epigenetic modifications in frontal cortex from Alzheimer's disease and bipolar disorder patients. Transl. Psychiatry *2*, e132.

Rauch, T.A., Wu, X., Zhong, X., Riggs, A.D., and Pfeifer, G.P. (2009). A human B cell methylome at 100−base pair resolution. Proc. Natl. Acad. Sci. U. S. A. *106*, 671–678.

Razin, A., and Riggs, A. (1980). DNA methylation and gene function. Science *210*, 604–610.

Reese, K.J., Lin, S., Verona, R.I., Schultz, R.M., and Bartolomei, M.S. (2007). Maintenance of Paternal Methylation and Repression of the Imprinted H19 Gene Requires MBD3. PLoS Genet. *3*, e137.

Reik, W., Brown, K.W., Schneid, H., Le Bouc, Y., Bickmore, W., and Maher, E.R. (1995). Imprinting mutations in the Beckwith—Wiedemann syndrome suggested by an altered imprinting pattern in the *IGF2–H19* domain. Hum. Mol. Genet. *4*, 2379–2385.

Renard, I., Joniau, S., van Cleynenbreugel, B., Collette, C., Naômé, C., Vlassenbroeck, I., Nicolas, H., de Leval, J., Straub, J., Van Criekinge, W., et al. (2010). Identification and Validation of the Methylated TWIST1 and NID2 Genes through Real-Time Methylation-Specific Polymerase Chain Reaction Assays for the Noninvasive Detection of Primary Bladder Cancer in Urine Samples. Eur. Urol. *58*, 96–104.

Rettori, M.M., de Carvalho, A.C., Longo, A.L.B., de Oliveira, C.Z., Kowalski, L.P., Carvalho, A.L., and Vettore, A.L. (2013). TIMP3 and CCNA1 hypermethylation in HNSCC is associated with an increased incidence of second primary tumors. J. Transl. Med. *11*, 316.

Revy, P., Muto, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O., Catalan, N., Forveille, M., Dufourcq-Lagelouse, R., Gennery, A., et al. (2000). Activation-Induced Cytidine Deaminase (AID) Deficiency Causes the Autosomal Recessive Form of the Hyper-IgM Syndrome (HIGM2). Cell *102*, 565–575.

Ribeiro, A.F.T., Pratcorona, M., Erpelinck-Verschueren, C., Rockova, V., Sanders, M., Abbas, S., Figueroa, M.E., Zeilemaker, A., Melnick, A., Löwenberg, B., et al. (2012). Mutant DNMT3A: a marker of poor prognosis in acute myeloid leukemia. Blood *119*, 5824–5831.

Richel, D.J., Colly, L.P., Kluin-Nelemans, J.C., and Willemze, R. (1991). The antileukaemic activity of 5-Aza-2 deoxycytidine (Aza-dC) in patients with relapsed and resistant leukaemia. Br. J. Cancer *64*, 144–148.

Riggs, A.D. (1975). X inactivation, differentiation, and DNA methylation. Cytogenet. Cell Genet. *14*, 9–25.

Rivenbark, A.G., Stolzenburg, S., Beltran, A.S., Yuan, X., Rots, M.G., Strahl, B.D., and Blancafort, P. (2012). Epigenetic reprogramming of cancer cells via targeted DNA methylation. Epigenetics *7*, 350–360.

Rohde, C., Zhang, Y., Reinhardt, R., and Jeltsch, A. (2010). BISMA--fast and accurate bisulfite sequencing data analysis of individual clones from unique and repetitive sequences. BMC Bioinformatics *11*, 230.

Ropero, S. (2004). Epigenetic loss of the familial tumor-suppressor gene exostosin-1 (EXT1) disrupts heparan sulfate synthesis in cancer cells. Hum. Mol. Genet. *13*, 2753–2765.

Rosas, S.L.B., Koch, W., Westra, W., Jen, J., and Sidransky, D. (2001). Promoter Hypermethylation Patterns of p16, O6-Methylguanine-DNA- methyltransferase, and Death-associated Protein Kinase in Tumors and Saliva of Head and Neck Cancer Patients. Cancer Res. *61*, 939–942.

Rountree, M.R., Bachman, K.E., and Baylin, S.B. (2000). DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. Nat. Genet. *25*, 11.

Rowe, H.M., Kapopoulou, A., Corsinotti, A., Fasching, L., Macfarlan, T.S., Tarabay, Y., Viville, S., Jakobsson, J., Pfaff, S.L., and Trono, D. (2013). TRIM28 repression of retrotransposon-based enhancers is necessary to preserve transcriptional dynamics in embryonic stem cells. Genome Res. *23*, 452–461.

Ruiz-Magaña, M.J., Martínez-Aguilar, R., Lucendo, E., Campillo-Davo, D., Schulze-Osthoff, K., Ruiz-Ruiz, C., Ruiz-Magaña, M.J., Martínez-Aguilar, R., Lucendo, E., Campillo-Davo, D., et al. (2016). The antihypertensive drug hydralazine activates the intrinsic pathway of apoptosis and causes DNA damage in leukemic T cells. Oncotarget *7*, 21875–21886.

Ruzov, A., Tsenkina, Y., Serio, A., Dudnakova, T., Fletcher, J., Bai, Y., Chebotareva, T., Pells, S., Hannoun, Z., Sullivan, G., et al. (2011). Lineage-specific distribution of high levels of genomic 5-hydroxymethylcytosine in mammalian development. Cell Res. *21*, 1332–1342.

Sado, T. (2004). De novo DNA methylation is dispensable for the initiation and propagation of X chromosome inactivation. Development *131*, 975–982.

Sado, T., Fenner, M.H., Tan, S.-S., Tam, P., Shioda, T., and Li, E. (2000). X Inactivation in the Mouse Embryo Deficient for Dnmt1: Distinct Effect of Hypomethylation on Imprinted and Random X Inactivation. Dev. Biol. *225*, 294–303.

Saito, K., Kawakami, K., Matsumoto, I., Oda, M., Watanabe, G., and Minamoto, T. (2010). Long Interspersed Nuclear Element 1 Hypomethylation Is a Marker of Poor Prognosis in Stage IA Non-Small Cell Lung Cancer. Clin. Cancer Res. *16*, 2418–2426.

Saito, Y., Kanai, Y., Nakagawa, T., Sakamoto, M., Saito, H., Ishii, H., and Hirohashi, S. (2003). Increased protein expression of DNA methyltransferase (DNMT) 1 is significantly correlated with the malignant potential and poor prognosis of human hepatocellular carcinomas. Int. J. Cancer *105*, 527–532.

Saito, Y., Liang, G., Egger, G., Friedman, J.M., Chuang, J.C., Coetzee, G.A., and Jones, P.A. (2006). Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. Cancer Cell *9*, 435–443.

Saitou, M., Kagiwada, S., and Kurimoto, K. (2012). Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. Development *139*, 15–31.

San José-Enériz, E., Agirre, X., Rabal, O., Vilas-Zornoza, A., Sanchez-Arias, J.A., Miranda, E., Ugarte, A., Roa, S., Paiva, B., Estella-Hermoso de Mendoza, A., et al. (2017). Discovery of first-in-class reversible dual small molecule inhibitors against G9a and DNMTs in hematological malignancies. Nat. Commun. *8*.

Sanchez-Cespedes, M., Esteller, M., Wu, L., Nawroz-Danish, H., Yoo, G.H., Koch, W.M., Jen, J., Herman, J.G., and Sidransky, D. (2000). Gene Promoter Hypermethylation in Tumors and Serum of Head and Neck Cancer Patients. Cancer Res. *60*, 892–895.

Sanchez-Mut, J.V., Heyn, H., Vidal, E., Moran, S., Sayols, S., Delgado-Morales, R., Schultz, M.D., Ansoleaga, B., Garcia-Esparcia, P., Pons-Espinal, M., et al. (2016). Human DNA methylomes of neurodegenerative diseases show common epigenomic patterns. Transl. Psychiatry *6*, e718–e718.

Sansom, O.J., Maddison, K., and Clarke, A.R. (2007). Mechanisms of disease: methyl-binding domain proteins as potential therapeutic targets in cancer. Nat. Clin. Pract. Oncol. *4*, 305–315.

Santos, F., Hendrich, B., Reik, W., and Dean, W. (2002). Dynamic Reprogramming of DNA Methylation in the Early Mouse Embryo. Dev. Biol. *241*, 172–182.

Santos, F., Peat, J., Burgess, H., Rada, C., Reik, W., and Dean, W. (2013). Active demethylation in mouse zygotes involves cytosine deamination and base excision repair. Epigenetics Chromatin *6*, 39.

Sathyanarayana, U.G., Padar, A., Huang, C.X., Suzuki, M., Shigematsu, H., Bekele, B.N., and Gazdar, A.F. (2003a). Aberrant Promoter Methylation and Silencing of Laminin-5–Encoding Genes in Breast Carcinoma. Clin. Cancer Res. *9*, 6389–6394.

Sathyanarayana, U.G., Padar, A., Suzuki, M., Maruyama, R., Shigematsu, H., Hsieh, J.-T., Frenkel, E.P., and Gazdar, A.F. (2003b). Aberrant Promoter Methylation of Laminin-5–Encoding Genes in Prostate Cancers and Its Relationship to Clinicopathological Features. Clin. Cancer Res. *9*, 6395–6400.

Sathyanarayana, U.G., Maruyama, R., Padar, A., Suzuki, M., Bondaruk, J., Sagalowsky, A., Minna, J.D., Frenkel, E.P., Grossman, H.B., Czerniak, B., et al. (2004). Molecular Detection of Noninvasive and Invasive Bladder Tumor Tissues and Exfoliated Cells by Aberrant Promoter Methylation of Laminin-5 Encoding Genes. Cancer Res. *64*, 1425–1430.

Sato, S., Yoshida, W., Soejima, H., Nakabayashi, K., and Hata, K. (2011). Methylation dynamics of IG-DMR and Gtl2-DMR during murine embryonic and placental development. Genomics *98*, 120–127.

Saunderson, E.A., Stepper, P., Gomm, J.J., Hoa, L., Morgan, A., Allen, M.D., Jones, J.L., Gribben, J.G., Jurkowski, T.P., and Ficz, G. (2017). Hit-and-run epigenetic editing prevents senescence entry in primary breast cells from healthy donors. Nat. Commun. *8*, 1450.

Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc. Natl. Acad. Sci. U. S. A. *103*, 1412–1417.

Schaefer, M., Pollex, T., Hanna, K., Tuorto, F., Meusburger, M., Helm, M., and Lyko, F. (2010). RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. Genes Dev. *24*, 1590–1595.

Schermelleh, L., Spada, F., Easwaran, H.P., Zolghadr, K., Margot, J.B., Cardoso, M.C., and Leonhardt, H. (2005). Trapped in action: direct visualization of DNA methyltransferase activity in living cells. Nat. Methods *2*, 751–756.

Schiesser, S., Hackner, B., Pfaffeneder, T., Müller, M., Hagemeier, C., Truss, M., and Carell, T. (2012). Mechanism and Stem-Cell Activity of 5-Carboxycytosine Decarboxylation Determined by Isotope Tracing. Angew. Chem. Int. Ed. *51*, 6516–6520.

Schlesinger, Y., Straussman, R., Keshet, I., Farkash, S., Hecht, M., Zimmerman, J., Eden, E., Yakhini, Z., Ben-Shushan, E., Reubinoff, B.E., et al. (2007). Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for *de novo* methylation in cancer. Nat. Genet. *39*, 232–236.

Schmitz, M.L., and Baeuerle, P.A. (1991). The p65 subunit is responsible for the strong transcription activating potential of NF-kappa B. EMBO J. *10*, 3805–3817.

Schomacher, L., Han, D., Musheev, M.U., Arab, K., Kienhöfer, S., von Seggern, A., and Niehrs, C. (2016). Neil DNA glycosylases promote substrate turnover by Tdg during DNA demethylation. Nat. Struct. Mol. Biol. *23*, 116–124.

Schutsky, E.K., Nabel, C.S., Davis, A.K.F., DeNizio, J.E., and Kohli, R.M. (2017). APOBEC3A efficiently deaminates methylated, but not TET-oxidized, cytosine bases in DNA. Nucleic Acids Res. *45*, 7655–7665.

Scott, S.A., Lakshimikuttysamma, A., Sheridan, D.P., Sanche, S.E., Geyer, C.R., and DeCoteau, J.F. (2007). Zebularine inhibits human acute myeloid leukemia cell growth in vitro in association with p15INK4B demethylation and reexpression. Exp. Hematol. *35*, 263–273.

Scourzic, L., Mouly, E., and Bernard, O.A. (2015). TET proteins and the control of cytosine demethylation in cancer. Genome Med. *7*, 9.

Seelan, R.S., Mukhopadhyay, P., Pisano, M.M., and Greene, R.M. (2018). Effects of 5-Aza-2′-deoxycytidine (decitabine) on gene expression. Drug Metab. Rev. *50*, 193–207.

Seftor, R.E., Seftor, E.A., Sheng, S., Pemberton, P.A., Sager, R., and Hendrix, M.J. (1998). Maspin suppresses the invasive phenotype of human breast carcinoma. Cancer Res. *58*, 5681–5685.

Segura-Pacheco, B., Trejo-Becerril, C., Perez-Cardenas, E., and Taja-Chayeb, L. (2003). Reactivation of Tumor Suppressor Genes by the Cardiovascular Drugs Hydralazine and Procainamide and Their Potential Use in Cancer Therapy. Clin. Cancer Res. *9*, 1596–1603.

Seisenberger, S., Andrews, S., Krueger, F., Arand, J., Walter, J., Santos, F., Popp, C., Thienpont, B., Dean, W., and Reik, W. (2012). The Dynamics of Genome-wide DNA Methylation Reprogramming in Mouse Primordial Germ Cells. Mol. Cell *48*, 849–862.

Seshagiri, S., Stawiski, E.W., Durinck, S., Modrusan, Z., Storm, E.E., Conboy, C.B., Chaudhuri, S., Guan, Y., Janakiraman, V., Jaiswal, B.S., et al. (2012). Recurrent R-spondin fusions in colon cancer. Nature *488*, 660–664.

Shah, S.A., Erdmann, S., Mojica, F.J.M., and Garrett, R.A. (2013). Protospacer recognition motifs. RNA Biol. *10*, 891–899.

Shan, M., Yin, H., Li, J., Li, X., Wang, D., Su, Y., Niu, M., Zhong, Z., Wang, J., Zhang, X., et al. (2016). Detection of aberrant methylation of a six-gene panel in serum DNA for diagnosis of breast cancer. Oncotarget *7*, 18485–18494.

Shao, L. -j., Shi, H.Y., Ayala, G., Rowley, D., and Zhang, M. (2008). Haploinsufficiency of the Maspin Tumor Suppressor Gene Leads to Hyperplastic Lesions in Prostate. Cancer Res. *68*, 5143–5151.

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., et al. (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature *450*, 908–912.

Sharma, R.A., Plummer, R., Stock, J.K., Greenhalgh, T.A., Ataman, O., Kelly, S., Clay, R., Adams, R.A., Baird, R.D., Billingham, L., et al. (2016). Clinical development of new drug–radiotherapy combinations. Nat. Rev. Clin. Oncol. *13*, 627.

Shen, L., Wu, H., Diep, D., Yamaguchi, S., D'Alessio, A.C., Fung, A., Zhang, K., and Zhang, Y. (2013). Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. Cell *153*, 692–706.

Shen, L., Inoue, A., He, J., Liu, Y., Lu, F., and Zhang, Y. (2014). Tet3 and DNA Replication Mediate Demethylation of Both the Maternal and Paternal Genomes in Mouse Zygotes. Cell Stem Cell *15*, 459–471.

Shen, L., Shi, Q., and Wang, W. (2018). Double agents: genes with both oncogenic and tumor-suppressor functions. Oncogenesis *7*, 25.

Sheng, S., Carey, J., Seftor, E.A., Dias, L., Hendrix, M.J., and Sager, R. (1996). Maspin acts at the cell membrane to inhibit invasion and motility of mammary and prostatic cancer cells. Proc. Natl. Acad. Sci. *93*, 11669–11674.

Shenoy, N., Bhagat, T., Nieves, E., Stenson, M., Lawson, J., Choudhary, G.S., Habermann, T., Nowakowski, G., Singh, R., Wu, X., et al. (2017). Upregulation of TET activity with ascorbic acid induces epigenetic modulation of lymphoma cells. Blood Cancer J. *7*, e587.

Shi, H.Y., Liang, R., Templeton, N.S., and Zhang, M. (2002). Inhibition of Breast Tumor Progression by Systemic Delivery of the Maspin Gene in a Syngeneic Tumor Model. Mol. Ther. *5*, 755–761.

Shi, H.Y., Stafford, L.J., Liu, Z., Liu, M., and Zhang, M. (2007). Maspin controls mammary tumor cell migration through inhibiting Rac1 and Cdc42, but not the RhoA GTPase. Cell Motil. Cytoskeleton *64*, 338–346.

Shigaki, H., Baba, Y., Watanabe, M., Murata, A., Iwagami, S., Miyake, K., Ishimoto, T., Iwatsuki, M., and Baba, H. (2013). LINE-1 hypomethylation in gastric cancer, detected by bisulfite pyrosequencing, is associated with poor prognosis. Gastric Cancer *16*, 480–487.

Shimoda, N., Hirose, K., Kaneto, R., Izawa, T., Yokoi, H., Hashimoto, N., and Kikuchi, Y. (2014). No Evidence for AID/MBD4-Coupled DNA Demethylation in Zebrafish Embryos. PLOS ONE *9*, e114816.

Shirahata, A., and Hibi, K. (2014). Serum Vimentin Methylation as a Potential Marker for Colorectal Cancer. Anticancer Res. *34*, 4121–4125.

Shirahata, A., Sakuraba, K., Goto, T., Saito, M., Ishibashi, K., Kigawa, G., Nemoto, H., and Hibi, K. (2010). Detection of Vimentin (VIM) Methylation in the Serum of Colorectal Cancer Patients. Anticancer Res. *30*, 5015–5018.

Shirane, K., Toh, H., Kobayashi, H., Miura, F., Chiba, H., Ito, T., Kono, T., and Sasaki, H. (2013). Mouse Oocyte Methylomes at Base Resolution Reveal Genome-Wide Accumulation of Non-CpG Methylation and Role of DNA Methyltransferases. PLoS Genet. *9*, e1003439.

Shlush, L.I., Zandi, S., Mitchell, A., Chen, W.C., Brandwein, J.M., Gupta, V., Kennedy, J.A., Schimmer, A.D., Schuh, A.C., Yee, K.W., et al. (2014). Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. Nature *506*, 328–333.

Shukla, R., Upton, K.R., Muñoz-Lopez, M., Gerhardt, D.J., Fisher, M.E., Nguyen, T., Brennan, P.M., Baillie, J.K., Collino, A., Ghisletti, S., et al. (2013). Endogenous Retrotransposition Activates Oncogenic Pathways in Hepatocellular Carcinoma. Cell *153*, 101–111.

Siddique, A.N., Nunna, S., Rajavelu, A., Zhang, Y., Jurkowska, R.Z., Reinhardt, R., Rots, M.G., Ragozin, S., Jurkowski, T.P., and Jeltsch, A. (2013). Targeted Methylation and Gene Silencing of VEGF-A in Human Cells by Using a Designed Dnmt3a–Dnmt3L Single-Chain Fusion Protein with Increased DNA Methylation Activity. J. Mol. Biol. *425*, 479–491.

da Silva, V.L., Fonseca, A.F., Fonseca, M., da Silva, T.E., Coelho, A.C., Kroll, J.E., de Souza, J.E.S., Stransky, B., de Souza, G.A., and de Souza, S.J. (2017). Genome-wide identification of cancer/testis genes and their association with prognosis in a pan-cancer analysis. Oncotarget *8*, 92966–92977.

Simpkins, S.B., Bocker, T., Swisher, E.M., Mutch, D.G., Gersell, D.J., Kovatich, A.J., Palazzo, J.P., Fishel, R., and Goodfellow, P.J. (1999). MLH1 promoter methylation and gene silencing is the primary cause of microsatellite instability in sporadic endometrial cancers. Hum. Mol. Genet. *8*, 661–666.

Simpson, A.J.G., Caballero, O.L., Jungbluth, A., Chen, Y.-T., and Old, L.J. (2005). Cancer/testis antigens, gametogenesis and cancer. Nat. Rev. Cancer *5*, 615–625.

Singh, M., and Jadhav, H.R. (2018). Targeting non-small cell lung cancer with small-molecule EGFR tyrosine kinase inhibitors. Drug Discov. Today *23*, 745–753.

Skiriute, D., Vaitkiene, P., Saferis, V., Asmoniene, V., Skauminas, K., Deltuva, V.P., and Tamasauskas, A. (2012). MGMT, GATA6, CD81, DR4, and CASP8 gene promoter methylation in glioblastoma. BMC Cancer *12*, 218.

Slaymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X., and Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. Science *351*, 84–88.

Smallwood, S.A., Tomizawa, S., Krueger, F., Ruf, N., Carli, N., Segonds-Pichon, A., Sato, S., Hata, K., Andrews, S.R., and Kelsey, G. (2011). Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. Nat. Genet. *43*, 811–814.

Smets, M., Link, S., Wolf, P., Schneider, K., Solis, V., Ryan, J., Meilinger, D., Qin, W., and Leonhardt, H. (2017). DNMT1 mutations found in HSANIE patients affect interaction with UHRF1 and neuronal differentiation. Hum. Mol. Genet. *26*, 1522–1534.

Smith, Z.D., and Meissner, A. (2013). DNA methylation: roles in mammalian development. Nat. Rev. Genet. *14*, 204–220.

Smith, Z.D., Chan, M.M., Mikkelsen, T.S., Gu, H., Gnirke, A., Regev, A., and Meissner, A. (2012). A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature *484*, 339–344.

Snowden, A.W., Gregory, P.D., Case, C.C., and Pabo, C.O. (2002). Gene-Specific Targeting of H3K9 Methylation Is Sufficient for Initiating Repression In Vivo. Curr. Biol. *12*, 2159–2166.

Soengas, M.S., Capodieci, P., Polsky, D., Mora, J., Esteller, M., Opitz-Araya, X., McCombie, R., Herman, J.G., Gerald, W.L., and Lazebnik, Y.A. (2001). Inactivation of the apoptosis effector Apaf-1 in malignant melanoma. Nature *409*, 207–211.

Song, J., Rechkoblit, O., Bestor, T.H., and Patel, D.J. (2011). Structure of DNMT1-DNA Complex Reveals a Role for Autoinhibition in Maintenance DNA Methylation. Science *331*, 1036–1040.

Šorm, F., Píškala, A., Čihák, A., and Veselý, J. (1964). 5-Azacytidine, a new, highly effective cancerostatic. Experientia *20*, 202–203.

Spada, F., Haemmer, A., Kuch, D., Rothbauer, U., Schermelleh, L., Kremmer, E., Carell, T., Längst, G., and Leonhardt, H. (2007). DNMT1 but not its interaction with the replication machinery is required for maintenance of DNA methylation in human cells. J. Cell Biol. *176*, 565–571.

Sparago, A., Cerrato, F., Vernucci, M., Ferrero, G.B., Silengo, M.C., and Riccio, A. (2004). Microdeletions in the human H19 DMR result in loss of IGF2 imprinting and Beckwith-Wiedemann syndrome. Nat. Genet. *36*, 958–960.

Spencer, D.H., Russler-Germain, D.A., Ketkar, S., Helton, N.M., Lamprecht, T.L., Fulton, R.S., Fronick, C.C., O'Laughlin, M., Heath, S.E., Shinawi, M., et al. (2017). CpG Island Hypermethylation Mediated by DNMT3A Is a Consequence of AML Progression. Cell *168*, 801-816.e13.

Spiesbach, K., Tannapfel, A., Mössner, J., and Engeland, K. (2005). TAp63gamma can substitute for p53 in inducing expression of the maspin tumor suppressor. Int. J. Cancer *114*, 555–562.

Sproul, D., Kitchen, R.R., Nestor, C.E., Dixon, J.M., Sims, A.H., Harrison, D.J., Ramsahoye, B.H., and Meehan, R.R. (2012). Tissue of origin determines cancer-associated CpG island promoter hypermethylation patterns. Genome Biol. *13*, R84.

Spruijt, C.G., Gnerlich, F., Smits, A.H., Pfaffeneder, T., Jansen, P.W.T.C., Bauer, C., Münzel, M., Wagner, M., Müller, M., Khan, F., et al. (2013). Dynamic Readers for 5-(Hydroxy)Methylcytosine and Its Oxidized Derivatives. Cell *152*, 1146–1159.

Stadler, M.B., Murr, R., Burger, L., Ivanek, R., Lienert, F., Schöler, A., Wirbelauer, C., Oakeley, E.J., Gaidatzis, D., Tiwari, V.K., et al. (2011). DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature *480*, 490–495.

Stein, R., Razin, A., and Cedar, H. (1982). In vitro methylation of the hamster adenine phosphoribosyltransferase gene inhibits its expression in mouse L cells. Proc. Natl. Acad. Sci. U. S. A. *79*, 3418–3422.

Stepper, P., Kungulovski, G., Jurkowska, R.Z., Chandra, T., Krueger, F., Reinhardt, R., Reik, W., Jeltsch, A., and Jurkowski, T.P. (2017). Efficient targeted DNA methylation with chimeric dCas9–Dnmt3a–Dnmt3L methyltransferase. Nucleic Acids Res. *45*, 1703–1713.

Stewart, K.R., Veselovska, L., Kim, J., Huang, J., Saadeh, H., Tomizawa, S., Smallwood, S.A., Chen, T., and Kelsey, G. (2015). Dynamic changes in histone modifications precede de novo DNA methylation in oocytes. Genes Dev. *29*, 2449–2462.

Stolzenburg, S., Beltran, A.S., Swift-Scanlan, T., Rivenbark, A.G., Rashwan, R., and Blancafort, P. (2015). Stable oncogenic silencing in vivo by programmable and targeted de novo DNA methylation in breast cancer. Oncogene *34*, 5427–5435.

Su, Y., Fang, H., and Jiang, F. (2016). Integrating DNA methylation and microRNA biomarkers in sputum for lung cancer detection. Clin. Epigenetics *8*, 109.

Subramaniam, D., Thombre, R., Dhar, A., and Anant, S. (2014). DNA Methyltransferases: A Novel Target for Prevention and Therapy. Front. Oncol. *4*.

Sullivan, C.H., and Grainger, R.M. (1987). Delta-crystallin genes become hypomethylated in postmitotic lens cells during chicken development. Proc. Natl. Acad. Sci. U. S. A. *84*, 329–333.

Sun, D., Rou, W., Zhou, Y., Zhuo, T., Fan, X., Hu, X., and Zou, H. (2018a). Ectopic expression of the TAL effector AvrXa7 in Xanthomonas citri subsp. citri hinders citrus canker symptom formation by modulating transcriptional profile of citrus genes. Biochem. Biophys. Res. Commun. *502*, 479–485.

Sun, W., Ji, W., Hall, J.M., Hu, Q., Wang, C., Beisel, C.L., and Gu, Z. (2015). Self-Assembled DNA Nanoclews for the Efficient Delivery of CRISPR–Cas9 for Genome Editing. Angew. Chem. Int. Ed. *54*, 12029–12033.

Sun, W., Triche, T., Malvar, J., Gaynon, P., Sposto, R., Yang, X., Bittencourt, H., Place, A.E., Messinger, Y., Fraser, C., et al. (2018b). A phase 1 study of azacitidine combined with chemotherapy in childhood leukemia: a report from TACL consortium. Blood *131*, 1145–1148.

Sun, Z., Wu, Y., Ordog, T., Baheti, S., Nie, J., Duan, X., Hojo, K., Kocher, J.-P., Dyck, P.J., and Klein, C.J. (2014). Aberrant signature methylome by DNMT1 hot spot mutation in hereditary sensory and autonomic neuropathy 1E. Epigenetics *9*, 1184–1193.

Swets, M., Zaalberg, A., Boot, A., van Wezel, T., Frouws, M.A., Bastiaannet, E., Gelderblom, H., van de Velde, C.J.H., and Kuppen, P.J.K. (2016). Tumor LINE-1 Methylation Level in Association with Survival of Patients with Stage II Colon Cancer. Int. J. Mol. Sci. *18*, 36.

Szurek, B., Rossier, O., Hause, G., and Bonas, U. (2002). Type III-dependent translocation of the Xanthomonas AvrBs3 protein into the plant cell. Mol. Microbiol. *46*, 13–23.

Tabebordbar, M., Zhu, K., Cheng, J.K.W., Chew, W.L., Widrick, J.J., Yan, W.X., Maesner, C., Wu, E.Y., Xiao, R., Ran, F.A., et al. (2016). In vivo gene editing in dystrophic mouse muscle and muscle stem cells. Science *351*, 407–411.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930–935.

Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S., and Vale, R.D. (2014). A protein tagging system for signal amplification in gene expression and fluorescence imaging. Cell *159*, 635–646.

Tang, X., Khuri, F.R., Lee, J.J., Kemp, B.L., Liu, D., Hong, W.K., and Mao, L. (2000). Hypermethylation of the death-associated protein (DAP) kinase promoter and aggressiveness in stage I non-small-cell lung cancer. J. Natl. Cancer Inst. *92*, 1511–1516.

Tanić, M., and Beck, S. (2017). Epigenome-wide association studies for cancer biomarker discovery in circulating cell-free DNA: technical advances and challenges. Curr. Opin. Genet. Dev. *42*, 48–55.

Taylor, S.M., and Jones, P.A. (1979). Multiple new phenotypes induced in 10T1/2 and 3T3 cells treated with 5-azacytidine. Cell *17*, 771–779.

Tefferi, A., Levine, R., Lim, K.-H., Abdel-Wahab, O., Lasho, T., Patel, J., Finke, C., Mullally, A., Li, C.-Y., Pardanani, A., et al. (2009). Frequent TET2 mutations in systemic mastocytosis: clinical, KITD816V and FIP1L1-PDGFRA correlates. Leukemia *23*, 900–904.

Temple, I., and Shield, J. (2002). Transient neonatal diabetes, a disorder of imprinting. J. Med. Genet. *39*, 872–875.

Teng, I.-W., Hou, P.-C., Lee, K.-D., Chu, P.-Y., Yeh, K.-T., Jin, V.X., Tseng, M.-J., Tsai, S.-J., Chang, Y.-S., Wu, C.-S., et al. (2011). Targeted Methylation of Two Tumor Suppressor Genes Is Sufficient to Transform Mesenchymal Stem Cells into Cancer Stem/Initiating Cells. Cancer Res. *71*, 4653–4663.

Teoh, S.S.Y., Vieusseux, J., Prakash, M., Berkowicz, S., Luu, J., Bird, C.H., Law, R.H.P., Rosado, C., Price, J.T., Whisstock, J.C., et al. (2014). Maspin is not required for embryonic development or tumour suppression. Nat. Commun. *5*, 3164.

Terasawa, K., Sagae, S., Toyota, M., Tsukada, K., Ogi, K., Satoh, A., Mita, H., Imai, K., Tokino, T., and Kudo, R. (2004). Epigenetic Inactivation of ***TMS1/ASC*** in Ovarian Cancer. Clin. Cancer Res. *10*, 2000–2006.

Thakore, P.I., D'Ippolito, A.M., Song, L., Safi, A., Shivakumar, N.K., Kabadi, A.M., Reddy, T.E., Crawford, G.E., and Gersbach, C.A. (2015). Highly Specific Epigenome Editing by CRISPR/Cas9 Repressors for Silencing of Distal Regulatory Elements. Nat. Methods *12*, 1143–1149.

Tham, C., Chew, M., Soong, R., Lim, J., Ang, M., Tang, C., Zhao, Y., Ong, S.Y.K., and Liu, Y. (2014a). Postoperative serum methylation levels of TAC1 and SEPT9 are independent predictors of recurrence and survival of patients with colorectal cancer. Cancer *120*, 3131–3141.

Thijssen, P.E., Ito, Y., Grillo, G., Wang, J., Velasco, G., Nitta, H., Unoki, M., Yoshihara, M., Suyama, M., Sun, Y., et al. (2015). Mutations in CDCA7 and HELLS cause immunodeficiency–centromeric instability–facial anomalies syndrome. Nat. Commun. *6*, 7870.

Timp, W., Bravo, H.C., McDonald, O.G., Goggins, M., Umbricht, C., Zeiger, M., Feinberg, A.P., and Irizarry, R.A. (2014). Large hypomethylated blocks as a universal defining epigenetic alteration in human solid tumors. Genome Med. *6*, 61.

Tomizawa, S., Kobayashi, H., Watanabe, T., Andrews, S., Hata, K., Kelsey, G., and Sasaki, H. (2011). Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. Dev. Camb. Engl. *138*, 811–820.

Tsuji, S., Futaki, S., and Imanishi, M. (2013). Creating a TALE protein with unbiased 5′-T binding. Biochem. Biophys. Res. Commun. *441*, 262–265.

Tuorto, F., Liebers, R., Musch, T., Schaefer, M., Hofmann, S., Kellner, S., Frye, M., Helm, M., Stoecklin, G., and Lyko, F. (2012). RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis. Nat. Struct. Mol. Biol. *19*, 900–905.

Turcan, S., Rohle, D., Goenka, A., Walsh, L.A., Fang, F., Yilmaz, E., Campos, C., Fabius, A.W.M., Lu, C., Ward, P.S., et al. (2012). IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. Nature *483*, 479–483.

Valton, J., Dupuy, A., Daboussi, F., Thomas, S., Maréchal, A., Macmaster, R., Melliand, K., Juillerat, A., and Duchateau, P. (2012). Overcoming Transcription Activator-like Effector (TALE) DNA Binding Domain Sensitivity to Cytosine Methylation. J. Biol. Chem. *287*, 38427–38432.

Van Neste, L., Herman, J.G., Otto, G., Bigley, J.W., Epstein, J.I., and Van Criekinge, W. (2012). The Epigenetic promise for prostate cancer diagnosis. The Prostate *72*, 1248–1261.

Van Tongelen, A., Loriot, A., and De Smet, C. (2017). Oncogenic roles of DNA hypomethylation through the activation of cancer-germline genes. Cancer Lett. *396*, 130–137.

Vardimon, L., Kressmann, A., Cedar, H., Maechler, M., and Doerfler, W. (1982). Expression of a cloned adenovirus gene is inhibited by in vitro methylation. Proc. Natl. Acad. Sci. U. S. A. *79*, 1073–1077.

Varley, K.E., Gertz, J., Bowling, K.M., Parker, S.L., Reddy, T.E., Pauli-Behn, F., Cross, M.K., Williams, B.A., Stamatoyannopoulos, J.A., Crawford, G.E., et al. (2013). Dynamic DNA methylation across diverse human cell lines and tissues. Genome Res. *23*, 555–567.

Veland, N., Hardikar, S., Zhong, Y., Gayatri, S., Dan, J., Strahl, B.D., Rothbart, S.B., Bedford, M.T., and Chen, T. (2017). The Arginine Methyltransferase PRMT6 Regulates DNA Methylation and Contributes to Global DNA Hypomethylation in Cancer. Cell Rep. *21*, 3390–3397.

Velasco, G., Grillo, G., Touleimat, N., Ferry, L., Ivkovic, I., Ribierre, F., Deleuze, J.-F., Chantalat, S., Picard, C., and Francastel, C. (2018). Comparative methylome analysis of ICF patients identifies heterochromatin loci that require ZBTB24, CDCA7 and HELLS for their methylated state. Hum. Mol. Genet. *27*, 2409–2424.

Veluchamy, A., Lin, X., Maumus, F., Rivarola, M., Bhavsar, J., Creasy, T., O'Brien, K., Sengamalay, N.A., Tallon, L.J., Smith, A.D., et al. (2013). Insights into the role of DNA methylation in diatoms by genome-wide profiling in Phaeodactylum tricornutum. Nat. Commun. *4*, 2091.

Vidal, E., Sayols, S., Moran, S., Guillaumet-Adkins, A., Schroeder, M.P., Royo, R., Orozco, M., Gut, M., Gut, I., Lopez-Bigas, N., et al. (2017). A DNA methylation map of human cancer at single base-pair resolution. Oncogene *36*, 5648–5657.

Vilkaitis, G., Suetake, I., Klimašauskas, S., and Tajima, S. (2005). Processive Methylation of Hemimethylated CpG Sites by Mouse Dnmt1 DNA Methyltransferase. J. Biol. Chem. *280*, 64–72.

Villar-Garea, A., Fraga, M.F., Espada, J., and Esteller, M. (2003). Procaine Is a DNA-demethylating Agent with Growth-inhibitory Effects in Human Cancer Cells. Cancer Res. *63*, 4984–4989.

Vogler, W.R., Miller, D.S., and Keller, J.W. (1976). 5-Azacytidine (NSC 102816): a new drug for the treatment of myeloblastic leukemia. Blood *48*, 331–337.

Vojta, A., Dobrinić, P., Tadić, V., Bočkor, L., Korać, P., Julg, B., Klasić, M., and Zoldoš, V. (2016). Repurposing the CRISPR-Cas9 system for targeted DNA methylation. Nucleic Acids Res. *44*, 5615–5628.

Vukobrat-Bijedic, Z., Husic-Selimovic, A., Sofic, A., Bijedic, N., Bjelogrlic, I., Gogov, B., and Mehmedovic, A. (2013). Cancer Antigens (CEA and CA 19-9) as Markers of Advanced Stage of Colorectal Carcinoma. Med. Arch. *67*, 397–401.

Wahlberg, P., Lundmark, A., Nordlund, J., Busche, S., Raine, A., Tandre, K., Rönnblom, L., Sinnett, D., Forestier, E., Pastinen, T., et al. (2016). DNA methylome analysis of acute lymphoblastic leukemia cells reveals stochastic *de novo* DNA methylation in CpG islands. Epigenomics *8*, 1367–1387.

Waitkus, M.S., Diplas, B.H., and Yan, H. (2018). Biological Role and Therapeutic Potential of IDH Mutations in Cancer. Cancer Cell *34*, 186–195.

Walsh, C.P., Chaillet, J.R., and Bestor, T.H. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat. Genet. *20*, 116–117.

Walter, M.J., Ding, L., Shen, D., Shao, J., Grillot, M., McLellan, M., Fulton, R., Schmidt, H., Kalicki-Veizer, J., O'Laughlin, M., et al. (2011). Recurrent DNMT3A Mutations in Patients with Myelodysplastic Syndromes. Leukemia *25*, 1153–1158.

Wang, C., Gu, Y., Zhang, K., Xie, K., Zhu, M., Dai, N., Jiang, Y., Guo, X., Liu, M., Dai, J., et al. (2016). Systematic identification of genes with a cancer-testis expression pattern in 19 cancer types. Nat. Commun. *7*, 10499.

Wang, J., Sasco, A.J., Fu, C., Xue, H., Guo, G., Hua, Z., Zhou, Q., Jiang, Q., and Xu, B. (2008). Aberrant DNA Methylation of P16, MGMT, and hMLH1 Genes in Combination with MTHFR C677T Genetic Polymorphism in Esophageal Squamous Cell Carcinoma. Cancer Epidemiol. Biomarkers Prev. *17*, 118–125.

Wang, J., Xia, Y., Li, L., Gong, D., Yao, Y., Luo, H., Lu, H., Yi, N., Wu, H., Zhang, X., et al. (2013). Double restriction-enzyme digestion improves the coverage and accuracy of genome-wide CpG methylation profiling by reduced representation bisulfite sequencing. BMC Genomics *14*, 11.

Wang, L., Zhang, J., Duan, J., Gao, X., Zhu, W., Lu, X., Yang, L., Zhang, J., Li, G., Ci, W., et al. (2014a). Programming and inheritance of parental DNA methylomes in mammals. Cell *157*, 979–991.

Wang, S.E., Narasanna, A., Whitell, C.W., Wu, F.Y., Friedman, D.B., and Arteaga, C.L. (2007). Convergence of p53 and Transforming Growth Factor β (TGFβ) Signaling on Activating Expression of the Tumor Suppressor Gene *maspin* in Mammary Epithelial Cells. J. Biol. Chem. *282*, 5661–5669.

Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014b). Genetic Screens in Human Cells Using the CRISPR-Cas9 System. Science *343*, 80–84.

Ward, A.K., Mellor, P., Smith, S.E., Kendall, S., Just, N.A., Vizeacoumar, F.S., Sarker, S., Phillips, Z., Alvi, R., Saxena, A., et al. (2016). Epigenetic silencing of CREB3L1 by DNA methylation is associated with high-grade metastatic breast cancers with poor prognosis and is prevalent in triple negative breast cancers. Breast Cancer Res. BCR *18*, 12.

Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Pääbo, S., Rebhan, M., and Schübeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat. Genet. *39*, 457–466.

Weiner, G.J. (2015). Building better monoclonal antibody-based therapeutics. Nat. Rev. Cancer *15*, 361–370.

Weiss, G., Schlegel, A., Kottwitz, D., König, T., and Tetzner, R. (2017). Validation of the SHOX2/PTGER4 DNA Methylation Marker Panel for Plasma-Based Discrimination between Patients with Malignant and Nonmalignant Lung Disease. J. Thorac. Oncol. *12*, 77–84.

Weissmann, S., Alpermann, T., Grossmann, V., Kowarsch, A., Nadarajah, N., Eder, C., Dicker, F., Fasan, A., Haferlach, C., Haferlach, T., et al. (2012). Landscape of TET2 mutations in acute myeloid leukemia. Leukemia *26*, 934–942.

Whitehurst, A.W. (2014). Cause and Consequence of Cancer/Testis Antigen Activation in Cancer. Annu. Rev. Pharmacol. Toxicol. *54*, 251–272.

Wienholz, B.L., Kareta, M.S., Moarefi, A.H., Gordon, C.A., Ginno, P.A., and Chédin, F. (2010). DNMT3L Modulates Significant and Distinct Flanking Sequence Preference for DNA Methylation by DNMT3A and DNMT3B In Vivo. PLoS Genet. *6*, e1001106.

Wiesner, T., Lee, W., Obenauf, A.C., Ran, L., Murali, R., Zhang, Q.F., Wong, E.W.P., Hu, W., Scott, S.N., Shah, R.H., et al. (2015). Alternative transcription initiation leads to expression of a novel ALK isoform in cancer. Nature *526*, 453–457.

Wijesinghe, P., and Bhagwat, A.S. (2012). Efficient deamination of 5-methylcytosines in DNA by human APOBEC3A, but not by AID or APOBEC3G. Nucleic Acids Res. *40*, 9206–9217.

Wilks, A., Seldran, M., and Jost, J.P. (1984). An estrogen-dependent demethylation at the 5' end of the chicken vitellogenin gene is independent of DNA synthesis. Nucleic Acids Res. *12*, 1163–1177.

Williams, K., Christensen, J., Pedersen, M.T., Johansen, J.V., Cloos, P.A.C., Rappsilber, J., and Helin, K. (2011). Tet1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature *473*, 343–348.

Wilson, V.L., Jones, P.A., and Momparler, R.L. (1983). Inhibition of DNA Methylation in L1210 Leukemic Cells by 5-Aza-2'-deoxycytidine as a Possible Mechanism of Chemotherapeutic Action. Cancer Res. *43*, 3493–3496.

Winkelmann, J., Lin, L., Schormair, B., Kornum, B.R., Faraco, J., Plazzi, G., Melberg, A., Cornelio, F., Urban, A.E., Pizza, F., et al. (2012). Mutations in DNMT1 cause autosomal dominant cerebellar ataxia, deafness and narcolepsy. Hum. Mol. Genet. *21*, 2205–2210.

Winquist, E., Knox, J., Ayoub, J.-P., Wood, L., Wainman, N., Reid, G.K., Pearce, L., Shah, A., and Eisenhauer, E. (2006). Phase II trial of DNA methyltransferase 1 inhibition with the antisense oligonucleotide MG98 in patients with metastatic renal carcinoma: A National Cancer Institute of Canada Clinical Trials Group investigational new drug study. Invest. New Drugs *24*, 159–167.

Wiznerowicz, M., Jakobsson, J., Szulc, J., Liao, S., Quazzola, A., Beermann, F., Aebischer, P., and Trono, D. (2007). The Kruppel-associated Box Repressor Domain Can Trigger de Novo Promoter Methylation during Mouse Early Embryogenesis. J. Biol. Chem. *282*, 34535–34541.

Wolff, E.M., Byun, H.-M., Han, H.F., Sharma, S., Nichols, P.W., Siegmund, K.D., Yang, A.S., Jones, P.A., and Liang, G. (2010). Hypomethylation of a LINE-1 Promoter Activates an Alternate Transcript of the MET Oncogene in Bladders with Cancer. PLoS Genet. *6*, e1000917.

Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C.J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W., and Walter, J. (2011). 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nat. Commun. *2*, 241.

Wozniak, R.J., Klimecki, W.T., Lau, S.S., Feinstein, Y., and Futscher, B.W. (2006). 5-Aza-2'-deoxycytidine-mediated reductions in G9A histone methyltransferase and histone H3 K9 di-methylation levels are linked to tumor suppressor gene reactivation. Oncogene *26*, 77–90.

Wozniak, R.J., Klimecki, W.T., Lau, S.S., Feinstein, Y., and Futscher, B.W. (2007). 5-Aza-2'-deoxycytidine-mediated reductions in G9A histone methyltransferase and histone H3 K9 di-methylation levels are linked to tumor suppressor gene reactivation. Oncogene *26*, 77–90.

Wu, X., and Zhang, Y. (2017). TET-mediated active DNA demethylation: mechanism, function and beyond. Nat. Rev. Genet. *18*, 517–534.

Wu, H., D'Alessio, A.C., Ito, S., Xia, K., Wang, Z., Cui, K., Zhao, K., Sun, Y.E., and Zhang, Y. (2011a). Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature *473*, 389–393.

Wu, T., Giovannucci, E., Welge, J., Mallick, P., Tang, W.-Y., and Ho, S.-M. (2011b). Measurement of GSTP1 promoter methylation in body fluids may complement PSA screening: a meta-analysis. Br. J. Cancer *105*, 65–73.

Wu, X., Scott, D.A., Kriz, A.J., Chiu, A.C., Hsu, P.D., Dadon, D.B., Cheng, A.W., Trevino, A.E., Konermann, S., Chen, S., et al. (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat. Biotechnol. *32*, 670–676.

Wu, Y., Strawn, E., Basir, Z., Halverson, G., and Guo, S. (2007). Aberrant expression of deoxyribonucleic acid methyltransferases DNMT1, DNMT3A, and DNMT3B in women with endometriosis. Fertil. Steril. *87*, 24–32.

Wu, Y., Alvarez, M., Slamon, D.J., Koeffler, P., and Vadgama, J.V. (2010). Caspase 8 and maspin are downregulated in breast cancer cells due to CpG site promoter methylation. BMC Cancer *10*, 32.

Wu, Y., Zhou, H., Fan, X., Zhang, Y., Zhang, M., Wang, Y., Xie, Z., Bai, M., Yin, Q., Liang, D., et al. (2015). Correction of a genetic disease by CRISPR-Cas9-mediated gene editing in mouse spermatogonial stem cells. Cell Res. *25*, 67–79.

Xie, Q., Bai, Q., Zou, L.-Y., Zhang, Q.-Y., Zhou, Y., Chang, H., Yi, L., Zhu, J.-D., and Mi, M.-T. (2014). Genistein inhibits DNA methylation and increases expression of tumor suppressor genes in human breast cancer cells: Genistein Inhibits DNA Methylation. Genes. Chromosomes Cancer *53*, 422–431.

Xie, W., Barr, C.L., Kim, A., Yue, F., Lee, A.Y., Eubanks, J., Dempster, E.L., and Ren, B. (2012). Base-resolution Analyses of Sequence and Parent-of-Origin Dependent DNA Methylation in the Mouse Genome. Cell *148*, 816–831.

Xiong, Z., and Laird, P.W. (1997). COBRA: a sensitive and quantitative DNA methylation assay. Nucleic Acids Res. *25*, 2532–2534.

Xiong, T., Meister, G.E., Workman, R.E., Kato, N.C., Spellberg, M.J., Turker, F., Timp, W., Ostermeier, M., and Novina, C.D. (2017). Targeted DNA methylation in human cells using engineered dCas9-methyltransferases. Sci. Rep. *7*, 6732.

Xu, G.-L., Bestor, T.H., Hsieh, C.-L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J.J., and quignot, E.V.-P. (1999). Chromosome instability and immunode®ciency syndrome caused by mutations in a DNA methyltransferase gene. Nature *402*, 187–191.

Xu, X., Tao, Y., Gao, X., Zhang, L., Li, X., Zou, W., Ruan, K., Wang, F., Xu, G., and Hu, R. (2016a). A CRISPR-based approach for targeted DNA demethylation. Cell Discov. *2*, 16009.

Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A., Zheng, L., Zhang, H., Huang, S., et al. (2011). Genome-wide Regulation of 5hmC, 5mC and Gene Expression by Tet1 Hydroxylase in Mouse Embryonic Stem Cells. Mol. Cell *42*, 451–464.

Xu, Y., Xu, C., Kato, A., Tempel, W., Abreu, J.G., Bian, C., Hu, Y., Hu, D., Zhao, B., Cerovina, T., et al. (2012). Tet3 CXXC Domain and Dioxygenase Activity Cooperatively Regulate Key Genes for Xenopus Eye and Neural Development. Cell *151*, 1200–1213.

Xue, J.-H., Xu, G.-F., Gu, T.-P., Chen, G.-D., Han, B.-B., Xu, Z.-M., Bjørås, M., Krokan, H.E., Xu, G.-L., and Du, Y.-R. (2016). Uracil-DNA Glycosylase UNG Promotes Tet-mediated DNA Demethylation. J. Biol. Chem. *291*, 731–738.

Yamaguchi, S., Hong, K., Liu, R., Shen, L., Inoue, A., Diep, D., Zhang, K., and Zhang, Y. (2012). Tet1 controls meiosis by regulating meiotic gene expression. Nature *492*, 443–447.

Yamaguchi, S., Shen, L., Liu, Y., Sendler, D., and Zhang, Y. (2013). Role of Tet1 in erasure of genomic imprinting. Nature *504*, 460–464.

Yamamoto, N., Nakayama, T., Kajita, M., Miyake, T., Iwamoto, T., Kim, S.J., Sakai, A., Ishihara, H., Tamaki, Y., and Noguchi, S. (2012). Detection of aberrant promoter methylation of GSTP1, RASSF1A, and RARβ2 in serum DNA of patients with breast cancer by a newly established one-step methylation-specific PCR assay. Breast Cancer Res. Treat. *132*, 165–173.

Yamashita, Y., Yuan, J., Suetake, I., Suzuki, H., Ishikawa, Y., Choi, Y.L., Ueno, T., Soda, M., Hamada, T., Haruta, H., et al. (2010). Array-based genomic resequencing of human leukemia. Oncogene *29*, 3723–3731.

Yan, H., Parsons, W., and Jin, G. (2009). IDH1 and IDH2 Mutations in Gliomas. N. Engl. J. Med. *360*, 765–773.

Yan, W., Wu, K., Herman, J.G., Brock, M.V., Zhou, Y., Lu, Y., Zhang, Z., Yang, Y., and Guo, M. (2014). Epigenetic silencing of DACH1 induces the invasion and metastasis of gastric cancer by activating TGF-β signalling. J. Cell. Mol. Med. *18*, 2499–2511.

Yang, B., and White, F.F. (2004). Diverse Members of the AvrBs3/PthA Family of Type III Effectors Are Major Virulence Determinants in Bacterial Blight Disease of Rice. Mol. Plant. Microbe Interact. *17*, 1192–1200.

Yang, Y., and Bedford, M.T. (2013). Protein arginine methyltransferases and cancer. Nat. Rev. Cancer *13*, 37–50.

Yang, F., Babak, T., Shendure, J., and Disteche, C.M. (2010). Global survey of escape from X inactivation by RNA-sequencing in mouse. Genome Res. *20*, 614–622.

Yang, H., Liu, Y., Bai, F., Zhang, J.-Y., Ma, S.-H., Liu, J., Xu, Z.-D., Zhu, H.-G., Ling, Z.-Q., Ye, D., et al. (2013). Tumor development is associated with decrease of TET gene expression and 5-methylcytosine hydroxylation. Oncogene *32*, 663–669.

Yang, Q.-W., Liu, S., Tian, Y., Salwen, H.R., Chlenski, A., Weinstein, J., and Cohn, S.L. (2003). Methylation-associated Silencing of the Thrombospondin-1 Gene in Human Neuroblastoma. Cancer Res. *63*, 6299–6310.

Ye, H., Rong, Z., and Lin, Y. (2017). Live cell imaging of genomic loci using dCas9-SunTag system and a bright fluorescent protein. Protein Cell *8*, 853–855.

Yi, J.M., Dhir, M., Van Neste, L., Downing, S.R., Jeschke, J., Glöckner, S.C., de Freitas Calmon, M., Hooker, C.M., Funes, J.M., Boshoff, C., et al. (2011). Genomic and epigenomic integration identifies a prognostic signature in colon cancer. Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. *17*, 1535–1545.

Yin, R., Mao, S.-Q., Zhao, B., Chong, Z., Yang, Y., Zhao, C., Zhang, D., Huang, H., Gao, J., Li, Z., et al. (2013). Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. J. Am. Chem. Soc. *135*, 10396–10403.

Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., Das, P.K., Kivioja, T., Dave, K., Zhong, F., et al. (2017). Impact of cytosine methylation on DNA binding specificities of human transcription factors. Science *356*, eaaj2239.

Yoo, C.B., Jeong, S., Egger, G., Liang, G., Phiasivongsa, P., Tang, C., Redkar, S., and Jones, P.A. (2007). Delivery of 5-Aza-2'-Deoxycytidine to Cells Using Oligodeoxynucleotides. Cancer Res. *67*, 6400–6408.

Yoo, C.B., Chuang, J.C., Byun, H.-M., Egger, G., Yang, A.S., Dubeau, L., Long, T., Laird, P.W., Marquez, V.E., and Jones, P.A. (2008). Long-term Epigenetic Therapy with Oral Zebularine Has Minimal Side Effects and Prevents Intestinal Tumors in Mice. Cancer Prev. Res. (Phila. Pa.) *1*, 233–240.

Yu, M., Hon, G.C., Szulwach, K.E., Song, C.-X., Zhang, L., Kim, A., Li, X., Dai, Q., Shen, Y., Park, B., et al. (2012). Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. Cell *149*, 1368–1380.

Yuan, W., Chen, J., Shu, Y., Liu, S., Wu, L., Ji, J., Liu, Z., Tang, Q., Zhou, Z., Cheng, Y., et al. (2017). Correlation of DAPK1 methylation and the risk of gastrointestinal cancer: A systematic review and meta-analysis. PLOS ONE *12*, e0184959.

Zalatan, J.G., Lee, M.E., Almeida, R., Gilbert, L.A., Whitehead, E.H., La Russa, M., Tsai, J.C., Weissman, J.S., Dueber, J.E., Qi, L.S., et al. (2015). Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds. Cell *160*, 339–350.

Zamudio, N., Barau, J., Teissandier, A., Walter, M., Borsos, M., Servant, N., and Bourc'his, D. (2015). DNA methylation restrains transposons from adopting a chromatin signature permissive for meiotic recombination. Genes Dev. *29*, 1256–1270.

Zetsche, B., Volz, S.E., and Zhang, F. (2015a). A split-Cas9 architecture for inducible genome editing and transcription modulation. Nat. Biotechnol. *33*, 139–142.

Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S., Joung, J., van der Oost, J., Regev, A., et al. (2015b). Cpf1 is a single RNA-guided endonuclease of a Class 2 CRISPR-Cas system. Cell *163*, 759–771.

Zhang, F., Cong, L., Lodato, S., Kosuri, S., Church, G.M., and Arlotta, P. (2011a). Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat. Biotechnol. *29*, 149–153.

Zhang, G., Huang, H., Liu, D., Cheng, Y., Liu, X., Zhang, W., Yin, R., Zhang, D., Zhang, P., Liu, J., et al. (2015a). N6-Methyladenine DNA Modification in Drosophila. Cell *161*, 893–906.

Zhang, H., Zhang, X., Clark, E., Mulcahey, M., Huang, S., and Shi, Y.G. (2010). TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. Cell Res. *20*, 1390–1393.

Zhang, H.-H., Huang, B., Cao, Y.-H., Li, Q., and Xu, H.-F. (2017a). Role of 5-Aza-CdR in mitomycin-C chemosensitivity of T24 bladder cancer cells. Oncol. Lett. *14*, 5652–5656.

Zhang, J., Yang, P.L., and Gray, N.S. (2009). Targeting cancer with small molecule kinase inhibitors. Nat. Rev. Cancer *9*, 28–39.

Zhang, M., Maass, N., Magit, D., and Sager, R. (1997). Transactivation through Ets and Ap1 transcription sites determines the expression of the tumor-suppressing gene maspin. Cell Growth Differ. Mol. Biol. J. Am. Assoc. Cancer Res. *8*, 179–186.

Zhang, W., Shi, H.Y., and Zhang, M. (2005). Maspin overexpression modulates tumor cell apoptosis through the regulation of Bcl-2 family proteins. BMC Cancer *5*, 50.

Zhang, W., Zhou, L., Ding, S.-M., Xie, H.-Y., Xu, X., Wu, J., Chen, Q.-X., Zhang, F., Wei, B.-J., Eldin, A.T., et al. (2011b). Aberrant methylation of the CADM1 promoter is associated with poor prognosis in hepatocellular carcinoma treated with liver transplantation. Oncol. Rep. *25*, 1053–1062.

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S.W.-L., Chen, H., Henderson, I.R., Shinn, P., Pellegrini, M., Jacobsen, S.E., et al. (2006). Genome-wide High-Resolution Mapping and Functional Analysis of DNA Methylation in Arabidopsis. Cell *126*, 1189–1201.

Zhang, Y., Chen, F., Sun, Y., Zhou, S., Li, T., and Chen, R. (2011c). Effects of DNMT1 silencing on malignant phenotype and methylated gene expression in cervical cancer cells. J. Exp. Clin. Cancer Res. CR *30*, 98.

Zhang, Y., Liu, L., Guo, S., Song, J., Zhu, C., Yue, Z., Wei, W., and Yi, C. (2017b). Deciphering TAL effectors for 5-methylcytosine and 5-hydroxymethylcytosine recognition. Nat. Commun. *8*, 901.

Zhang, Z., Wu, E., Qian, Z., and Wu, W.-S. (2015b). A multicolor panel of TALE-KRAB based transcriptional repressor vectors enabling knockdown of multiple gene targets. Sci. Rep. *4*, 7338.

Zhao, J., Sun, B.K., Erwin, J.A., Song, J.-J., and Lee, J.T. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X-chromosome. Science *322*, 750–756.

Zhao, S.-L., Zhu, S.-T., Hao, X., Li, P., and Zhang, S.-T. (2011). Effects of DNA methyltransferase 1 inhibition on esophageal squamous cell carcinoma: The role of DNMT1 in ESCC. Dis. Esophagus *24*, 601–610.

Zhao, Z., Wu, Q., Cheng, J., Qiu, X., Zhang, J., and Fan, H. (2010). Depletion of DNMT3A Suppressed Cell Proliferation and Restored PTEN in Hepatocellular Carcinoma Cell. J. Biomed. Biotechnol. *2010*, 10.

Zhao, Z., Chen, L., Dawlaty, M.M., Pan, F., Weeks, O., Zhou, Y., Cao, Z., Shi, H., Wang, J., Lin, L., et al. (2015). Combined loss of Tet1 and Tet2 promotes B-cell, but not myeloid malignancies in mice. Cell Rep. *13*, 1692–1704.

Zheng, J., Xiong, D., Sun, X., Wang, J., Hao, M., Ding, T., Xiao, G., Wang, X., Mao, Y., Fu, Y., et al. (2012). Signification of Hypermethylated in Cancer 1 (HIC1) as Tumor Suppressor Gene in Tumor Progression. Cancer Microenviron. *5*, 285–293.

Zheng, Y., Zhang, H., Wang, Y., Li, X., Lu, P., Dong, F., Pang, Y., Ma, S., Cheng, H., Hao, S., et al. (2016). Loss of Dnmt3b accelerates MLL-AF9 leukemia progression. Leukemia *30*, 2373–2384.

Zhou, D., Tang, W., Wang, W., Pan, X., An, H.-X., and Zhang, Y. (2016). Association between aberrant APC promoter methylation and breast cancer pathogenesis: a meta-analysis of 35 observational studies. PeerJ *4*, e2203.

Zhou, L., Cheng, X., Connolly, B.A., Dickman, M.J., Hurd, P.J., and Hornby, D.P. (2002). Zebularine: A Novel DNA Methylation Inhibitor that Forms a Covalent Complex with DNA Methyltransferases. J. Mol. Biol. *321*, 591–599.

Zhou, W., Dinh, H.Q., Ramjan, Z., Weisenberger, D.J., Nicolet, C.M., Shen, H., Laird, P.W., and Berman, B.P. (2018). DNA methylation loss in late-replicating domains is linked to mitotic cell division. Nat. Genet. *50*, 591–602.

Zhou, X., Vink, M., Klaver, B., Berkhout, B., and Das, A.T. (2006). Optimization of the Tet-On system for regulated gene expression through viral evolution. Gene Ther. *13*, 1382–1390.

Zhu, J.-K. (2009). Active DNA Demethylation Mediated by DNA Glycosylases. Annu. Rev. Genet. *43*, 143–166.

Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. Nat. Genet. *39*, 61–69.

Ziller, M.J., Gu, H., Müller, F., Donaghey, J., Tsai, L.T.-Y., Kohlbacher, O., De Jager, P.L., Rosen, E.D., Bennett, D.A., Bernstein, B.E., et al. (2013). Charting a dynamic DNA methylation landscape of the human genome. Nature *500*, 477–481.

Zou, Z., Anisowicz, A., Hendrix, M.J., Thor, A., Neveu, M., Sheng, S., Rafidi, K., Seftor, E., and Sager, R. (1994). Maspin, a serpin with tumor-suppressing activity in human mammary epithelial cells. Science *263*, 526–529.